Screening for Depression:

November, 2012

McMaster Evidence Review and Synthesis Centre:

Homa Keshavarz, Donna Fitzpatrick-Lewis, David Streiner, Maureen Rice, Parminder Raina

McMaster University, Hamilton Ontario Canada

Project staff: Mary Gauld, L. Peirson

CTF Lead:
M. Joffres

CTF Members:
K. Pottie, J. Dickinson, G. Lewin,
E. Shaw, M. Tonelli

PHAC Scientific Officers:
Alejandra Jaramillo, Lesley Dunfield

Abstract

Background: Depression is a complex mental illness that is associated with disability and reduced quality of life for the person with the disorder, as well as posing a substantial societal burden. Prevalence of depression in the Canadian population has been estimated to vary from 5 to 8.2 percent annually.

Purpose: The purpose of this evidence review is to provide the Canadian Task Force on Preventive Health Care (CTFPHC) with the most recent literature to develop evidence-based recommendations on screening for depression. The aim of this evidence report is to evaluate the literature on the effectiveness of screening for depression in adults, including major depressive disorder (MDD) and related disorders and the impact on critical and important outcomes.

Data Sources: To identify the literature on screening for depression, the following electronic databases were searched: Medline, EMBASE, PsycINFO, Cochrane Central, and Cochrane Database of Systematic Reviews from 1994 to May 23, 2012. The search was quite broad in nature with the only limitations being date, human subjects, and English or French language. In addition, a grey literature search using a number of keyword terms for depression and screening was undertaken focusing on Canadian sources.

Study Selection: We used randomized controlled trials, observational studies, and systematic reviews with evidence for the clinical benefit or potential harms of screening to address key questions and contextual questions.

Data Abstraction: The titles and abstracts were each reviewed by two members of the synthesis team. Articles marked for inclusion by either team member went on to full text rating. Full text inclusion, data abstraction, and quality assessment were done by two people. All disagreements were resolved through discussions with the synthesis team and inclusion results were reviewed by a third person. The strength of evidence was determined based on the GRADE system of rating quality of evidence using GRADEPro® software. The exception to this process were studies related to the contextual questions of costs, performance indicators, patient preferences, subpopulations, and grey literature, for which abstraction was done by one person and evidence was not rated using the GRADE system.

Results: Five studies met the inclusion criteria for this review. No studies on community-based screening for depression in the general population met the inclusion criteria. Five cluster control studies in Japanese regions with high suicide rates were included in this report to quantify the effect of community-based depression screening (CDS) with followup on the completed suicide risk for residents aged 60 and over. The result of ratio of the rate ratios of meta-analysis from all five included studies demonstrate that the implementation of CDS program had a protective effect on the overall incidence of completed suicide among elderly population (RRR 0.5, 95% CI, 0.32 to 0.78; p=0.002).

When gender was considered, the ratio of the risk ratio showed a significant reduction effect in suicide in women (RRR=0.37 95% CI, 0.21 to 0.66; p=0.0006) but not in men (RRR=0.67, 95% CI, 0.35 to 1.27; p=0.22). The overall GRADE rating applied to this evidence is very low quality. There was no evidence that met the inclusion criteria concerning harms of screening for depression.

Limitations: The studies that were reviewed here evaluated the effectiveness of the community-based depression screening programs which incorporated screening for depression, follow-up with mental health care or psychiatric treatment, and health education in the community setting in rural Japan with higher than average rates of suicide. As such, the observed reduction in suicide rates or recovery from depression cannot be attributed solely to the screening component of these programs. The findings of this review are affected by the limitations of the included literature. The search was limited to papers written in English or French. There is the potential that we have missed the opportunity to analyze data from papers written in other languages.

Conclusions: The ultimate goal of screening for depression is to decrease the incidence of and mortality from this disease. Although the scope of our review included outcomes beyond incidence of or mortality from depression there is a very limited research evidence allowing conclusions to be drawn on the effectiveness of screening for depression in the general or high risk populations.

Contents

Figures

Tables

Appendices

Chapter 1. Introduction

Purpose

Depression is a complex mental illness that is associated with disability and reduced quality of life for the person with the disorder, as well as posing a substantial societal burden. The purpose of this evidence report is to evaluate the literature on the effectiveness of screening for depression in adults, including major depressive disorder (MDD) and related disorders (such as dysthymia and subsyndromal depression) and the impact on critical and important outcomes. This review will provide the Canadian Task Force on Preventive Health Care (CTFPHC) with the most recent literature to develop evidence-based recommendations on screening for depression

Condition Background

Definition

Depression is common in adults and is characterized by chronic, recurrent episodes that have significant impact on disability and mortality. Major depressive disorder (MDD) is the occurrence of one or more major depressive episodes (MDE) and can be classified from mild to severe. An MDE is defined as a period of at least 2 weeks that is characterized either by depressed mood and/or markedly diminished interest or pleasure in all, or almost all, activities in addition to at least four other symptoms.[1] When considering the less severe forms of depression, dysthymic disorder is characterized by a chronically depressed mood and at least two other depressive symptoms, that occur most of the day, more days than not, for at least 2 years. DSM-IV includes specifiers that can be used to further describe the characteristics of MDE or dysthymic disorder, such as whether an episode of depression includes psychosis or occurs in the postpartum period.[1]

Prevalence and burden of disease

MDD is a leading cause of disability across the world.[2,3] It is the second leading medical cause of long-term disability and the fourth leading cause of global burden of disease, and is predicted to become the second leading cause (next to heart disease) by 2020. According to the World Health Organization (WHO), at 4.5 percent, major depression is among the leading causes of disability-adjusted life years (DALY) exceeding ischemic heart disease, stroke, cancer (3.9%, 3.3% and 3.5% respectively).[4]

Prevalence of depression in Canada has been estimated to vary from 5 to 8.2 percent annually.[5,6] Kessler reported estimates of 6.6 percent annual prevalence of depression in the United States for adults.[7] These are higher than European prevalence rates of 3.9 percent annually.[8]

As cognitive problems are a hallmark of MDD, its negative effect is especially destructive in the knowledge-based economy prominent in western societies.[9]The economic burden of depressive disorders in the United States (U.S.)is estimated to be $83.1 billion (USD).[10] A 2001 Health Canada study estimated that productivity losses due to depression at $4.5 billion (CDN) and the annual direct and indirect cost to be in excess of $14 billion (CDN).[11] The National Comorbidity

Survey Replication study in the U.S. found that about 60 percent of respondents with an MDE in the past year reported severe or very severe role impairment.[7] Parental depression has a negative effect on the development of their children and on family dynamics[12,13] and intergenerational effects may amplify the impact of depression on population health.[14]

The burden of MDD is amplified by the typically chronic and recurrent nature of this disorder.[15-17] In the Canadian population MDD is associated with a wide variety of adverse outcomes such as suicide,[18,19] elevated chronic disease incidence,[20-23] a negative impact on chronic disease risk factors such as obesity,[24,25] and smoking,[26,27] increased injury risk,[28] diminished participation in preventive health activities,[29] and negative employment transitions.[30]

Despite increases in provision of treatment for people with depression,[31] a reduction in prevalence has not yet been discernable in those countries where before–after comparisons have been feasible.[32,33] This may be in part because a substantial number of people with depression remain untreated or receive inadequate treatment.[34] Subsyndromal depression, although poorly recognized, is at least as common as MDD and linked to poor quality of life.[35]

Etiology and natural history

There are several competing theories about the etiology of MDD. These include genetic predisposition, biochemical imbalances, endocrine and neurophysiological dysfunction, psychological, and/or social processes and factors. The evidence is limited for each of these factors, but from a healthcare perspective, the observations that some physical illnesses (e.g., hypothyroidism, Cushing's syndrome, cardiac disease, diabetes, etc.,) are frequently associated with depression tends to support the physical or endocrine role in the development of depression.[36] The recent aggressive exploration of brain function with neuroimaging, has provided some evidence to suggest that depression may be related to brain structure and function.[37] Similarly, psychological assessment findings provide some evidence for the role of cognitive and emotional processes as important factors in depression.[38]

The Netherlands Mental Health Survey and Incidence Study assessed episode duration in individuals with new-onset episodes. Although 50 percent of people recovered within 3 months, the recovery rate flattened over time, and the authors estimated that approximately 20 percent would have episodes lasting longer than 24 months.[39] Some estimates suggest that at least 10 percent of patients have persistent or chronic depression.[7]

Traditionally, depression has been thought of as a time-limited disorder lasting on average 4 to 9 months and that complete recovery is an attainable outcome following treatment. However, it is apparent that incomplete recovery (residual symptoms) and relapse are common. At least 50 percent of people following their first episode of major depression will go on to have at least one more episode and after the second and third episodes, the risk of further relapse rises to 70 percent and 90 percent respectively.[40]

Consequences if left untreated

Untreated depression exerts a negative impact on physical health; it reduces adherence to medical treatment,[42] reduces participation in preventive activities,[43] and increases the likelihood

of risk factors such as obesity,[44] smoking,[45] and sedentary lifestyles.[46] MDD may be associated with immune dysfunction,[47-50] cardiovascular disease,[23,51,52] endocrine and neurological diseases and a general increase in chronic disease incidence.[20] The relationship between depression and medical illness cuts both ways. While people with chronic illnesses have substantially higher rates of depression than the general population, according to the Mood Disorder Society of Canada, people with depression, compared to the general population, have higher risk levels for stroke (2.6 times), epilepsy (4-6 times), Alzheimer's Disease (1.71-2.67 times), and cancer (1.35-1.88 times)[53] People who are depressed are 4 times more likely to have a heart attack than the general population and heart patients who are depressed are 4 times more likely to die in the 6 months after a heart attack.[53] Mortality rates are high: approximately 4 percent of people with a mood disorder die by their own hand and about two thirds of suicides are preceded by depression.[34]

Depression also has a negative impact on occupational functioning. In one study, depressed workers had significantly greater performance deficits than control workers with regard to performing interpersonal tasks, time management, output and physical tasks.[54] When depressed workers were compared to workers with rheumatoid arthritis, the depressed employees were almost 5 times more likely to become unemployed than those with arthritis.[55] Depressed employees are also more likely to become unemployed or miss time at work than physically ill employees.[56]

An American study of quality of life (QOL) impairment found that 63 percent of respondents with MDD had severe impairment in QOL, while 85 percent of those with MDD and dysthymic disorder and 56 percent of those with dysthymic disorder had QOL impairment in the severe range.[57]

Risk factors

The factors most commonly associated with depression are gender, age, and family history of depression, and chronic physical illness.[58-61] The most consistent finding in cross-national studies of risk factors for depression is that women are at higher risk than men.[62] In Canada women are twice as likely as men to develop depression in their lifetime,[53]although these gender differences tend to decrease with age.[5] People under 20 years old have the highest rate of depression symptoms and age of onset of the illness is most commonly late adolescence or early adulthood,[53,62] while the prevalence decreases with age.[63,64] Family history of depression is frequently cited as a risk factor for depression, and one Canadian study of the cumulative incidence of major depressive episodes (MDE) over 6 years found that family history of MDE was the strongest independent predictor of MDE (HR = 2.01, 95% CI: 1.51, 2.68).[64] They also found that both female gender or having one or more chronic conditions were strongly associated with MDE (HR = 1.72, 95% CI: 1.36, 2.18 and HR = 1.54, 95% CI: 1.18-2.00 respectively) while age (entered as a continuous variable, per year) was negatively associated with MDE (HR = 0.98, 95% CI: 0.97, 0.99). Finally, they found that income, employment status, marital status, educational levels, ethnicity and geographic location were not significantly associated with MDE. The two most controversial results in this study relate to income and marital status; prevalence studies regularly cite low income and single marital status (including widowed, separated or divorced) as being strongly associated with depression.[62,63,65]

One risk factor not considered in many prevalence or incidence studies of depression is aboriginal heritage. According to the Mood Disorders Society of Canada,[53] 30.4 percent of Aboriginal Canadians living on reserves have had at least one MDE compared to 10-12 percent of the general population, and 26.1 percent have chronic depression compared to 2.7 percent of the general population.[41] In addition, off-reserve Aboriginal people are 1.5 times more likely to experience depression than the general population.[53]

Rationale for screening

Depression is an illness that has several factors that favour screening as well as several that do not. As discussed in previous sections, depression is a fairly common illness with a substantial societal burden and there are widely agreed upon high risk groups (women, youth, people with chronic illness or a family history of depression). In addition, depression is frequently under-diagnosed in primary practice. The WHO Psychological Problems in General Health Care study[66] released in 1996, reported that primary care physicians diagnosed only 42 percent of adult patients with major depression.

Potential benefits of screening for depression in adults include improved detection of MDD, dysthymia, and subsyndromal depression which can lead to earlier treatment. Treatment of MDD in adults is thought to result in improved outcomes such as quality of life, work, and minimized risk of suicide.[59] This review was designed to determine which of these benefits are supported by evidence.

Factors that argue against screening include the fact that in up to 50 percent of people depression resolves without treatment within 3 months.[39] Most organizations or guidelines that do recommend screening do so only in the case where there are systems in place to ensure effective treatment and followup.[67,68] In addition, screening instruments have a low positive predictive value, meaning that many who screen positive do not have depression.[69,70]Although a previous review found no literature specifically evaluating harms associated with screening for depression and related disorders,[58] those persons screening positive for depression who do not have the disorder may be exposed to stigmatization, further psychological testing, as well as unnecessary psychological and pharmacological treatment regimes.

Interventions/treatments of persons diagnosed with MDD, dysthymia or subsyndromal depression

Patients with positive findings for depression and related disorders may be recommended for treatment which consists primarily of pharmacological interventions, psychological therapies, or some combination of the two. Alternative treatments may include exercise, light therapy, acupuncture or herbal remedies and supplements. Many non-pharmacological therapies have not been validated for their efficacy.[71]

In primary care, the range of interventions offered following diagnosis may extend from close monitoring of mild episodes without immediate treatment (watchful waiting), through guided

self-management,[72] brief psychological or behavioural interventions, pharmacological management, and, if needed, referral to more specialized services or hospital admission. Recently, a stepped care approach to treatment has gained prominence. In a stepped care approach, more intensive treatments (e.g., second and third line treatments) are applied to persons who have not, or may not benefit from first-line treatments (which tend to be monotherapy type interventions considered to be simpler in scope), or for those who can be accurately predicted not to benefit from such treatment.[71] Often low intensity in the first step, these initial interventions may include watchful waiting, non-pharmacological treatments (including psychological interventions), or monotherapy drug interventions. The aim in this stepped approach is to reserve high intensity and high resource interventions for those who do not respond to more benign therapies. There is always the risk, however, that the preference for first step approaches may lead to initial inadequate management of those with more severe symptoms or higher needs. Recently, NICE guidelines from the United Kingdom, where a stepped care approach is advocated, have taken this risk into account by considering the symptom severity at the start.[59] The overall effectiveness of this approach has yet to be determined.

Phases of treatment of major depressive disorder

Based on the work of Kupfer et al.,[40] there are three phases of treatment for MDD: acute, continuation, and maintenance. Acute treatment is aimed at the elimination of symptoms of depression and restoration of psychosocial functioning. Continuation is a prolongation of treatment from 4 to 9 months, such that the episode of depression is considered completely resolved. For the continuation phase, the treatment aims to return patients to baseline function and quality of life and to prevent recurrence of symptoms. For the maintenance phase, the treatment goal is to prevent recurrence of new episodes of MDD.

The target goal for acute treatment should be remission, which is defined as a resolution of depressive symptoms. Response to treatment (usually defined as at least a 50 percent reduction in symptom levels)[73] may not be sufficient as a target outcome because residual depressive symptoms are risk factors for relapse and negative predictors of long-term outcome.[74]

Previous Review and CTFPHC Recommendations

There have been two sets of recommendations by the CTFPHC that have evaluated the value of screening for depression.[67,75] The recommendations from the most recent CTFPHC are detailed in Table 1.

Table 1. Recommendations for Screening for Depression by the CTFPHC for adults (from MacMillan 2004)[67]

Manoeuvre	Effectiveness	Level of Evidence	Recommendation^
Screening adults for depression in settings with integrated feedback and treatment	There is evidence that screening improves the accuracy of identifying depressed patients. In those studies where an integrated system of screening and follow-up was available, there was improvement in	Level I, good, fair (Pignone et al., 2002; Katzelnick et al., 2000; Rost et al., 2001, Wells et al.,	The CTFPHC concludes that there is fair evidence to recommend screening adults for depression in those primary care settings that have integrated programs for feedback to patients and access to case

systems*	depressive symptoms (Pignone et al., 2002)	2000)	management or mental health care (**B Recommendation**)
Screening adults for depression in settings without integrated feedback and treatment systems*	There is evidence that screening improves the accuracy of identifying depressed patients. In those studies without integrated feedback and treatment systems, there were no improvements in depressive symptoms (Pignone et al.,2002).	Level I, good (systematic review of RCTs) (Pignone et al., 2002)	The CTFPHC concludes that there is insufficient evidence to recommend for or against screening adults for depression in primary care settings where effective follow-up and treatment* are not available (**I recommendation**).

* screening programs integrated with both feedback to the clinician regarding depression status, as well as a system for managing treatment (antidepressant medications and psychotherapeutic interventions). Trials that included access to case management or mental health care as part of the system of care were particularly effective in reducing depressive symptoms. Since integrated screening and feedback/treatment systems are not the norm in Canadian primary care practice, clinicians are encouraged to advocate for these.

^ Grading scheme for the recommendations is based on the Task Force criteria. The CTFPHC concludes that there is fair evidence to recommend screening adults for depression in primary care settings since screening improves health outcomes when linked to effective follow-up and treatment* (**B recommendation**). The CTFPHC concludes that there is insufficient evidence to recommend for or against screening adults for depression in primary care settings where effective follow- up and treatment* are not available (**I recommendation**).

In addition to the previous Canadian Task Force recommendations, there have been two related sets of recommendations[58,68] for screening for depression in adults based on systematic reviews[58,60,68,77,78] sponsored by the United States Preventive Services Task Force (USPSTF) in 2002 and 2009.

In 2009, the USPSTF task force updated their 2002 recommendations for both adults and adolescents. In 2002 the USPSTF recommended "screening adults for depression in clinical practices that have systems in place to assure accurate diagnosis, effective treatment, and follow-up".[68] In 2009 the Task Force updated these recommendations to specify that the support systems had to be "staff-assisted depression care supports".[58] In both cases the guidelines were given a B grade meaning that the USPSTF recommends the service and that there is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. In 2009 the Task Force also "recommended against routinely screening adults for depression when staff-assisted depression care supports are not in place." They noted that "there may be considerations that support screening for depression in an individual patient." [58]

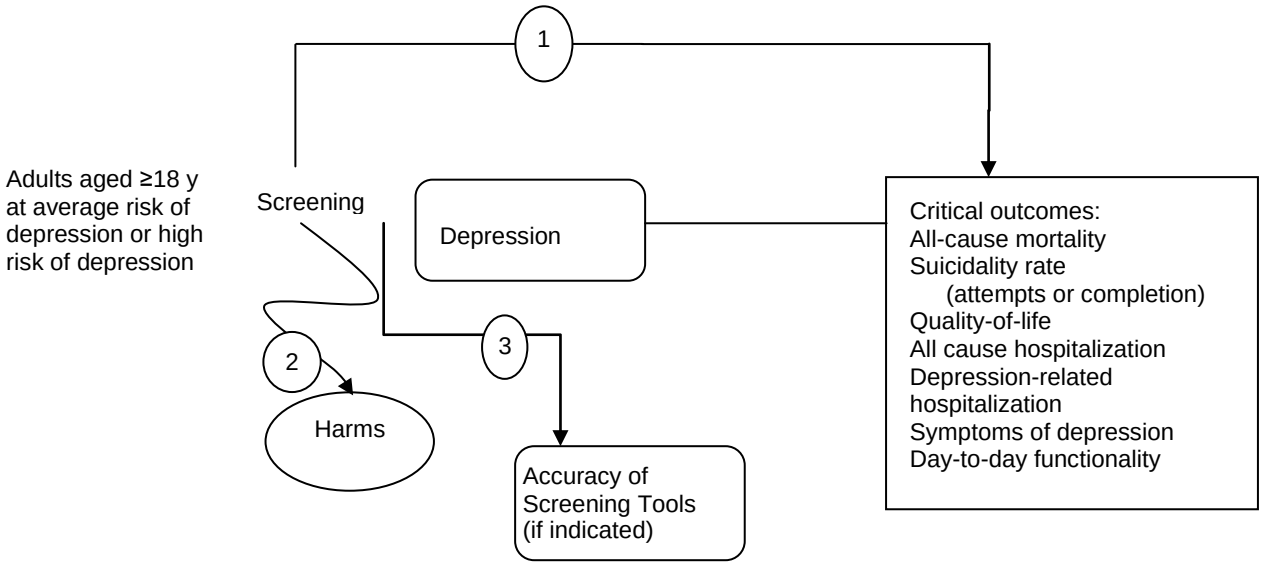
The National Institute for Health & Clinical Evidence (NICE) in the U.K. considered screening for depression in the general population but decided that it should only be undertaken for specific high-risk populations.[79] These guidelines, first put forth in 2004 have been integrated into the primary care system. More recent guidelines have focused on case identification in primary care and on the instruments used for that purpose.[69]

Chapter 2. Methods

Analytic Framework and Key Questions.

The analytic framework (Figure 1) shows the key questions to be addressed in the evidence review.

Figure 1. Analytic Framework for Screening for Depression



Key Questions

Stage 1: Evaluating the evidence for benefits or harms of screening for depression

Two key questions (KQ) are included in the first stage of this review as follows:

KQ1: What is the evidence for the benefit of screening for depression in:

a. asymptomatic adults 18 years of age or over from the general population in (i) primary care or (ii) other outpatient settings to improve critical outcomes?

b. adults at high risk for depression, in (i) primary care, (ii) other outpatient settings, or (iii) specialty clinic setting to improve critical outcomes?

KQ2: What is the evidence for the harms of screening for depression in

a. asymptomatic adults 18 years of age or over not at high risk for depression in (i) primary care or (ii) other outpatient settings

b. adults at high risk for depression in (i) primary care, (ii) other outpatient settings, or (iii) specialty clinics?

Stage 2: Evaluating the evidence for accuracy of tools to detect depression in primary care settings

KQ3a. What are the depression screening tool(s) that are most effective (accurate) in diagnosing or detecting depression in adult patients in primary care settings?

KQ3b. What is the effectiveness of short screening questions tools (ultra short = 1-4 items and taking less than 2 minutes to complete; short = 5-14 items and 2 to 5 minutes) compared with long screening tools (≥15 items and more than 5 minutes) to screen for depression in primary care settings?

The first stage, which has been completed, looked at the evidence for benefits and harms of screening in adults in both the general population in primary care and in selected patient groups selected by the Depression Screening Working Group that are considered to be at high risk in primary care, and other outpatient settings.

Stage 2 would be undertaken only if the evidence pointed to the benefit of screening for at least some of the population under investigation. This stage would have evaluated the accuracy of tools to detect depression in the primary care setting. Based on the findings of Stage 1, it was decided that evidence was not sufficient to complete Stage 2 of the review.

Contextual questions

Additional contextual questions (CQ) in unselected and high risk adult populations in primary care, outpatient and specialty clinic settings previously identified include:

CQ1. What is the evidence concerning the optimal interval of screening for depression?
CQ2. What is the cost-effectiveness of screening for depression?
CQ3. What are the patient preferences and values regarding screening?
CQ4. Are there subgroups of the Canadian population who have a higher prevalence of depression or for whom it would be difficult to implement screening programs? Subgroup analysis that explores issues of burden of disease, screening rates and special implementation issues include:
- Aboriginal
- Rural or remote-dwelling populations
- other ethnic groups

CQ5. What are patient preferences and values for treatment interventions (antidepressants and/or psychotherapy) for depression?
CQ6. What are the benefits and harms associated with the treatment (antidepressants and/or psychotherapy) for depression?

Originally, consideration was given to updating the O'Connor et al., (2009) review[58] which served as the basis for the 2009 USPSTF recommendations for depression screening in adults. However, the review excluded high risk populations that were to be included in this review in order to address KQ1b as outlined below. Based on that the Depression Working Group decided to undertake a de novo review.

Search Strategy

For Stage 1 of the review, the following electronic databases were searched: Medline, EMBASE, PsycINFO, Cochrane Central and Cochrane Database of Systematic Reviews from 1994 to May 23, 2012 The search was quite broad in nature with the only limitations being date, human subjects and English or French language. In addition, a grey literature search using a number of keyword terms for depression and screening was undertaken focusing on Canadian sources. This included both site specific searching (see Appendix A) and a general Google search limited to "pages from Canada".

For the contextual questions on depression screening, it was not necessary to undertake a separate search as any results would be a subset of the results of the search for KQ1 and KQ2. Articles addressing these questions were identified as part of the screening process for the key questions. For contextual questions, three separate searches were conducted in the databases mentioned above: 1) systematic reviews on depression treatment, 2) systematic reviews on adverse events associated with treatment, and 3) patient preferences and values regarding treatment for depression. All these searches were limited to the last 5 years, human subjects and English or French language. Detailed search strategies can be found in Appendix A.

Study Selection

KQ1: Eligible studies included adults ≥18 from unselected primary care populations or, for KQ1b, high risk groups. The intervention of interest was routine screening as a normal part of care and any comparative study design was accepted as long as it involved a screen versus no-screen comparison. To be included the study setting had to be in primary care or, in the case of high risk groups, specialty clinics and the study had to have at least one of the outcomes of interest (see Table 2).

Table 2. Depression Screening Working Group Ranked Outcomes of Interest

Outcome	Importance
All cause mortality	Critical (9)
Suicidal completion	Critical (9)
Suicide attempt	Critical (8)
Quality of life	Critical (8)
All cause hospitalization rate	Critical (7)
Depression-related hospitalization	Critical (7)
Improved symptoms of depression	Critical (7)
Improved day-to-day functionality	Critical (7)
Lost time at work/school	Important (6)
Suicidal ideation	Important (4)
Impact on lifestyle behaviour (alcohol abuse, smoking, drugs, gambling, etc.)	Important (4)

KQ2: The inclusion criteria for KQ2 were the same as for KQ1 with two exceptions: 1) all study types except for case studies were eligible, and 2) outcomes of interest were harms related to screening (see Table 3).

The Depression Screening Working Group rated each of the outcomes and potential harms of screening using the GRADE Process. GRADE suggests a nine point scale (1 to 9) to judge the importance of the outcomes and harms. The upper end of the scale, rankings of 7 to 9, identifies outcomes of critical importance for clinical decision making. Rankings of 4 to 6 represent outcomes that are important but not critical, while rankings of 1 to 3 are items that are deemed to be of limited importance to decision making or to patients. The outcomes of harms associated with depression screening resulted in the rankings presented in Table 3.

Table 3. Depression Screening Working Group Ranked Harms Associated with Screening

Harm	Importance
Mortality rate	Critical (9)
Suicidality rate (attempt)	Critical (8)
Hospitalization rate	Critical (7)
Cost	Important (5)
Labeling: intermediate harm, false positives and negatives, stigma, anxiety, decreased day to day functionality, increased symptoms	Important (5)

Contextual Questions: Studies on the contextual questions on screening were included if they were relevant to the questions posed. In the case of contextual questions on treatment, studies were included if they were systematic reviews, RCTs and observational studies.

External Review

Before the review began, the protocol was internally reviewed by the Depression Screening Working Group, which includes members of the CTFPHC and Public Health Agency Staff. The revised protocol was sent to external reviewers with expertise in review methodology or depression (see Appendix B). Revisions were made after feedback was received.

Quality Assessment, Data Abstraction and Analysis

Each title and abstract was reviewed by two trained screeners and disagreements were resolved by a third screener, any article marked for inclusion by either team member went on to full text rating. Full text inclusion and quality assessment were each done by two people and data abstraction was done by one person and checked by another. All disagreements were resolved through discussion. The inclusion results were reviewed by a third person. The exception to this process were studies related to the contextual questions of costs, performance indicators, patient preferences, and subpopulations, for which abstraction was done by one person and evidence was not rated using the GRADE system. Five quasi-experimental studies with pre- and post-

implementation design were quality appraised with the Newcastle-Ottawa Scale (NOS)[80] measuring the domains: selection of study groups, comparability of study groups, and means of ascertaining exposure or outcome. The NOS uses a 'star system' to score studies (maximum score is nine stars).

The strength of evidence was determined based on the GRADE system of rating quality of evidence using GRADEPro software.[81-83] This system of grading evidence has been widely used and has been endorsed by over 40 major organizations including the World Health Organization (WHO), Centers for Disease Control and Prevention (CDC), and the Agency for Healthcare Research and Quality (AHRQ).[81] The GRADE system rates the quality of a body of evidence as high, moderate, low or very low; each of the four levels reflects a different assessment of the likelihood that further research will impact the estimate of effect (e.g., high quality: further research is unlikely to change confidence in the estimate of effect). [81] A GRADE quality rating is based on an assessment of five conditions: (1) limitations in study designs (risk of bias), (2) inconsistency (heterogeneity) in the direction and/or size of the estimates of effect, (3) indirectness of the body of evidence to the populations, interventions, comparators and/or outcomes of interest, (4) imprecision of results (few events/observations, wide confidence intervals), and (5) indications of reporting or publication bias. Grouped RCTs begin with a high quality rating which may be downgraded if there are serious or very serious concerns across the studies related to one or more of the five conditions. All groups of observational (e.g., case-control and cohort) studies begin with a low quality rating which may be further downgraded based on assessments of the same five criteria. All other types of evidence are assigned a very low quality rating. For this review, key data were entered into the GRADEPro software along with the quality assessment ratings to produce two analytic products, the GRADE Evidence Profile Tables and the GRADE Summary of Findings Tables .

Review Manager 5.1 was used for meta-analysis. Pooled relative risk (RR) was used to summarize the effect of intervention on suicide (dichotomous outcome). A Random Effect assumption (Inverse Variance Weighting method) was used to calculate pooled estimates and the corresponding confidence intervals.[84] Homogeneity between studies was assessed using the χ^2 test and the I^2 statistic. Heterogeneity was considered significant if the p-value was less than 0.1.

Chapter 3. Results

Summary of the Literature Search

Our search located 14,226 potentially relevant unique citations (Figure 2). At title and abstract screening, 12,694 were excluded. A total of 1,532 papers were retrieved and were assessed on inclusion criteria. A total of 1,527 papers were excluded at this level because they did not meet our inclusion critieria for Key Questions 1 or 2.

Evaluating the evidence for benefits or harms of screening for depression

Two key questions are included in the first stage of this review as follows:

KQ1: What is the evidence for the benefit of screening for depression in:

a. asymptomatic adults 18 years of age or over from the general population not at high risk for depression in (i) primary care or (ii) other outpatient settings to improve critical outcomes?

b. adults at high risk for depression in (i) primary care, (ii) other outpatient settings, or (iii) specialty clinic setting to improve critical outcomes?

General population/primary care

No studies of screening for the depression in the general population as a whole met the inclusion criteria of this review. Five primary studies were the only eligible evidence that was identified on elderly populations. These quasi-experimental studies, with the same initial author (Oyama), were conducted in rural regions of Japan with high suicide rates in the elderly (range from 49.6 to 418.4/100,000 in women and 113 to 326/100,000 in men),[85-89] and targeted the residents aged 60 and over. Oyama, et al., (1978 to 2006)[85-89] developed a universal suicide prevention program, which included a screening component adapted from the WHO World Mental Health Survey.[90] This included screening for depression, followup with mental health care or psychiatric treatment, and psychoeducation in the community setting. The duration of studies varied from 4 to 20 years with different stages. The overall aim of these studies was to evaluate the effectiveness of the community-based depression screening (CDS) program in both the short- and long-term. A pre- and post-implementation design, using an intervention community and a control community with similar demographics, was used in all five studies. The studies offered a range of programs to residents, including educational health workshops. Two of the five studies[88,89] also provided group social and recreational activities (see details in Appendix C Evidence Tables). In all studies, more than 60 percent of men and more than 80 percent of women in the targeted residents (aged ≥60) participated in the program during the implementation.

The five studies implemented almost the same program, providing a two-stepped screening and followup process for depression. In the first step, the older residents of the selected communities were called to participate in the educational health workshop on the signs 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 Self-rating Depression Scale (SDS),[91] or the Geriatric Depression Scale five-item (GDS-5).[92] The SDS was originally developed as a self-administered questionnaire, but since participants in the present program were elderly, the examiner read the questions aloud to those who could not fill in the questionnaire sheets themselves. Those who did not attend the workshop were contacted the following day and asked to participate in the program. Examiners then visited those who agreed to participate, and conducted the program following the same procedures. There were several examiners, including psychiatrists and public health nurses (PHNs).

In the second step, a mental health assessment was carried out by a PHN on enrolled participants with positive screening results on the SDS. Japanese translated schedules of a standardized assessment of patients with depressive disorders were used[93,94] and a clinical decision was made about whether a psychiatrist's medical examination was necessary. Throughout the interview, if the participants were suspected of having depression, they were given a clinical decision as to whether to refer to a psychiatrist or to continue to the PHN's followup interview, and were then re-examined. The final decision on treatment was made few months after the first contact.

Statistical Analysis

Data was presented in the papers as pre (baseline) and post (implementation) analysis for both intervention and control groups. Two out of the five papers[86,88] had two control groups where one paper[86] used adjusted odds ratios. The remaining three papers all had one control group. Four out of the five papers presented data using adjusted incidence rate ratios (IRR). Because data was given as pre and post for both intervention and control groups and not all papers calculated IRR, the ratio of rate ratios (RRR) was calculated for each group. Ratio of rate ratios is the ratio of the post- to pre-rate ratio in the intervention area divided by the corresponding post- to pre-rate ratio in the control area (see explanation Appendix D).

A weighted intervention effect was calculated across studies using data for overall population and stratified for age and gender.Assuming that any changes to the population at risk in the intervention area are the same as those in the control area, a RRR of less than 1.0 shows the reduction in the suicide IRR in the intervention area to that predicted from the IRR in the control area. Standard errors for logarithms of rate ratios and 95% CIs for rate ratios were calculated assuming that the number of events in each area in each period followed a Poisson distribution. The generic inverse variance method was used with a random effects meta-analysis model, since all studies were done by the same team/author working the same research design. Heterogeneity between studies was evaluated using a chi-squared test; there was considered to be significant heterogeneity when the p-value was less than or equal to 0.05.[95]

Effect of community-based screening on completed suicide risk

The meta-analysis of the target population involved 70,053 person-years and 65 suicide victims in intervention groups compared to 113,324 person-years and 145 suicide victims in the control groups during the implementation period. These studies reported six gender- and age-specific target population groups (age group 65 to 74, 75 to 84, and ≥85), with the exception of one study[89] that had different age groups (60 to 69, 70 to 79, ≥80). All five studies provided

sufficient data stratified by age, gender, and time periods for baseline and program implementation.

Outcomes of individual studies and a summary of meta-analysis results are shown in Figure 3 All the studies[85-89] demonstrated a statistically significant reduction in the number of completed suicides after implementation of the CDS program (RRR=0.5, 95% CI, 0.32 to 0.78; p=0.002). There was no significant heterogeneity among these studies (I^2=21%, χ^2=5.04; p=0.28). The outcome measure was an IRR based on binary data (i.e., suicide/no suicide that was calculated in both implementation and control before and after the intervention). There was no significant heterogeneity among these studies in either men or women, (I^2=21%, χ^2=5.07; p=0.28) and (I^2=0%, χ^2=1.41; p=0.84), respectively. Publication bias could not be assessed given the small number of included studies.

The difference between pooled incidence rate ratios and the corresponding 95% CI for completed suicide were calculated using the generic inverse variance weighting method for total number of men and women (Appendix C). The RRR of the data from all five included studies[85-89] suggested that the CDS program had a protective effect on the overall IRR in the elderly, (RRR=0.50, 95% CI, 0.32 to 0.78; p=0.002) (Figure 4) . The RRR also showed a risk reduction in suicide of women (RRR=0.37, 95% CI, 0.21 to 0.66; p=0.0006), whereas in men the effect was not significant (RRR=0.67, 95% CI, 0.35 to 1.27; p=0.22) (Figure 3).

Subgroup analysis

We considered subgroup analysis based on population characteristics. We carried out prespecified subgroup analyses by age groups (65-74, 75-84, and 85 or older) (Figure 5) and by gender and age groups (i.e men and women in age groups 65-74, 75-84, and 85 or older) (Figures 6-7).
Data were also pooled from the five studies reporting suicide rates for subgroups of similar age groups. As outlined above, four out of the five studies had similar age groups[85-88] and the other had a slightly different age group.[89]

To compare pooled results from all five of the studies with the pooled results of only the four studies with the same age groups, we carried out two separate pooled analyses. We did not find significant differences between the two analyses in terms of heterogeneity in all age groups in both men and women. We calculated the RRR for pre- and post-data in both the intervention and control groups for each specific age group and by gender and specific age group from the data in each study (Appendix C). Outcomes of individual studies and a summary of meta-analyses results for each age group and for each age group in both women and men are shown in Figures 6-7. Meta-analysis stratified by age groups showed a significant reduction effect on suicide in elderly at ages between 65 to 74 years (RRR=0.49, 95% CI, 0.26 to 0.94; p=0.03) and elderly at ages bwetween 75 to 84 years (RRR=0.44, 95% CI, 0.22 to 0.88; p=0.02) (Figure 5).
Subgroup meta-analysis showed a non-significant reduction effect on suicide in men across all age groups (RRR=0.74, 95% CI, 0.44 to 1.24; p=0.25) (Figure 6). There was a statistically significant reduction of completed suicide only in women at ages between 75 to 84 years (RRR=0.37, 95% CI, 0.17 to 0.81; p=0.01) (Figure 7).

Quality Assessment

Agreement was reached on the ratings by two independent reviewers. Using the NOS for Cohort Studies,[96] the body of evidence received a rating of seven stars out of a possible nine stars. In terms of the selection criteria, two out of a possible four stars would be awarded. The exposed cohorts would be considered truly or somewhat representative of the average elderly person in the community (star awarded). The nonexposed cohorts were drawn from a different source (reference communities with similar sociodemographic characteristics), thus no star was awarded for this item. Similarly, no star was awarded for the ascertainment of exposure (written self-reports/surveys). The second star in this category was awarded for the demonstration that the outcome of interest (i.e., suicide) was not present at the start of the study. With respect to the comparability criterion, this group of studies satisfies both items (comparability of cohorts on the basis of ages and several sociodemographic variables) and would be awarded two stars. The five studies received all three stars available for the outcome category. The assessment of the outcome (suicide) was based on local public health department registers of suicide episodes with diagnoses based on ICD-9 or ICD-10. The studies all used an adequate followup period (10 years) to allow for the outcome of interest to occur. The studies demonstrated adequate followup of the cohorts. (Table 4)

GRADE *Rating*

According to the GRADE system for assessing quality, observational evidence (including cohort designs) begins with a LOW rating . We have downgraded the evidence for indirectness given that the included studies all looked at elderly, rural Japanese populations which are unlikely to be representative of Canadians with an average or high risk for depression. We also downgraded the evidence because the use of community-based depression screening (CDS) programs which incorporated education and treatment means the result cannot be attributed solely to the screening component of these programs. Thus the overall GRADE rating applied to this evidence is VERY LOW QUALITY (Tables 5-14).

High risk population

Initially, the Depression Working Group selected only the 5 high risk groups in the key questions, however it was determined that some risk groups were not represented in that list. As a result the scope of the review was extended to include any risk factor. We re-reviewed our evidence base but did not find any evidence that met our inclusion criteria for any high risk group.

KQ2: What is the evidence for the harms of screening for depression in:

a. asymptomatic adults 18 years of age or over not at high risk for depression in (i) primary care or (ii) other outpatient settings?

b. adults at high risk for depression in (i) primary care, (ii) other outpatient settings, or (iii) specialty clinics?

Harms of Screening

No studies were identified that addressed the harms of depression screening.

Contextual questions

CQ1: What is the evidence concerning the optimal interval of screening for depression?

There is very little evidence available on the optimal interval of screening for depression. In their 2009 review the USPSTF concluded that the optimal interval for screening is "unknown".[107]

CQ2. What is the cost-effectiveness of screening for depression?

In their cost-utility analysis, Valenstein et al.,[108] compared screening for depression with no screening in a hypothetical cohort of 40-year old primary care patients. The lifelong time horizon was used and costs were reported in US dollars. However, a specific reference year for costs was not stated. Both costs and benefits (expressed as quality adjusted life years (QALYs)) were discounted at an annual rate of 3 percent. The analysis was taken from health care payer and societal perspectives. Patients were screened for depression using a self-administered questionnaire followed by an assessment by a nurse or primary care provider. The evaluation was based upon a non-stationary Markov model representing eight major health states of: 1) never depressed; 2) history of depression, in remission; 3) history of depression, still in treatment; 4) significant depressive symptoms; 5) significant depressive symptoms, in treatment; 6) major depression; 7) major depression, in treatment, and 8) deceased. The length of each transition cycle was considered to be 3 months. The model inputs including the estimate of prevalence and incidence of major depression and related symptoms, sensitivity and specificity of screening instruments, practice patterns, treatment outcomes, and the utility decrements associated with depressive illness were derived from published literature. In their base-case analysis the authors compared the cost-effective ness of screening strategies to no screening and estimated the expected cost of annual screening for depression, to be $192,444/QALY and $225,467/QALY from the societal and third party payer perspectives, respectively. The costs of one time and periodic (every 5 years) screening strategies for society were estimated to be $32,053/QALY and 50,988/QALY, respectively. The results of one-way sensitivity analyses showed that cost-utility ratios were most sensitive to the prevalence of major depression, the costs of screening, rates of treatment initiation, and remission rates following treatment. Based in the results of a Monte Carlo sensitivity analysis the authors reported that only 2.2 percent of annual screening interventions had cost-utility ratios less than the commonly used benchmark of $50,000/QALY. One time and periodic screening strategies had cost utility ratios less than $50,000/QALY in 35 percent and 16 percent of the simulations, respectively. Based on the above mentioned results, the authors concluded that only one time screening for depression was cost-effective. They believed that improvements in the treatment of depression may increase the cost-effectiveness of screening interventions.

CQ3. What are the patient preferences and values regarding screening?

Much of the research on acceptability of depression screening focuses on perinatal women and tests particular tools, most commonly the EPDS. In a nonrandomized study, Leigh and Milgrom, (2007)[109] contacted 407 women who had been administered the EPDS, of whom 84 had scored above the cut-off. All of the women, depressed or not, found the questionnaire acceptable, and none were upset or distressed by the questions.

Similarly, Buist et al., (2006)[110] surveyed 860 women in Australia from six cities and four towns who had been routinely screened both antenatally and postnatally using the EPDS. They report that women found the EPDS easy to complete (93.4%) and 85 percent experienced no discomfort in completing it. Another Australian study[111] of 479 postnatal women using a 5 point Likert scale that ranged from "Not Comfortable" to "Very Comfortable" found that 81.2 percent of the women rated the experience as "Comfortable" or "Very Comfortable" and only 4.3 percent as "Not Comfortable". A total of 97 percent of respondents felt that screening was desirable.

One study addressed the question of patients' preferences with regard to how they are screened for depression. Allen, Cull, and Sharpe (2003)[112] asked cancer patients, who had been screened using the SCID in a telephone interview, if they found it helpful, and their preferred mode of questioning. Most of the patients (85%) found it helpful, and 17 percent found it distressing. The study does not explain why some individuals found the experience distressing but does note that most of these people (84%) also said that if they had known in advance how the questioning would make them feel, they would still have taken part in the interview. About half (47%) would have preferred a face-to-face interview.

A qualitative study of depression screening among ethnically diverse mothers in Boston looked at the acceptability of screening by pediatricians at well-child visits. The 42 women interviewed had to speak English, Spanish or Vietnamese and be the mother of a child 4 years of age or younger attending a well-child visit at one of the community health centres. They were screened using the PHQ-2. The women were generally receptive to being screened in the pediatric setting and overall they preferred that it be conducted by the pediatrician rather than support staff. The identified barriers to this screening were cultural "beliefs and norms" that discourage them for sharing their depressive symptoms and general concerns about stigmatization. They also expressed concerns about being reported to child protective services and losing custody of their children. They suggested a number of ways to improve acceptability of this screening. They mentioned the importance of the screening not being conducted in the presence of a third party such as their husband, child or interpreter. The women discussed that it was important that they be provided with the context of the screening particularly with regard to the reasons the questions were being asked and how they would benefit from answering them (e.g. available community support or treatment options). They also said they wanted reassurance that they weren't being singled out (that this screening was a common procedure) and that the pediatrician show concern about their overall health and wellness.

Finally, research done among Australian Aboriginal and Torres Strait Islanders points to the importance of particular screening tools being validated for use in specific populations and being modified where necessary.[113,114] Esler et al., (2007) conducted four focus groups in an urban Aboriginal health centre. Participants were asked given a copy of the PHQ-9 and asked to discuss the tool in terms of acceptability of the screening process, the individual questions and the rating scale. Three themes developed regarding acceptability: appropriate wording of the questions, particularly with regard to education levels and proficiency in English; the role of the family in the process, and; the relationship between the person administering the tool and the person being screened. Two other areas of concern were that the results could be confounded by social disadvantage and physical comorbidities and, the absence of a question that addresses

anger as a depression symptom (the authors note this last issue requires further study). The authors argue that this screening instrument was not acceptable among this population without modification.

It is difficult to generalize about patient preferences regarding depression screening given the variety of populations included in this literature, however, there are a couple of common themes. Overall people involved in these studies felt that screening was important and a majority of them felt the instruments tested were acceptable; however, screening needs to be done in a manner that is culturally sensitive. This includes ensuring that the context and purpose of the screening is clear and that the person being screened has an acceptable level of trust in the person doing the screening (or in the mode of screening). It is also important that the tool used for screening be validated in the population of interest.

CQ4. Are there subgroups of the Canadian population who have a higher prevalence of depression or for whom it would be difficult to implement screening programs? Subgroup analysis that explores issues of burden of disease, screening rates and special implementation issues include:

- Aboriginal
- Rural or remote-dwelling populations
- Other ethnic groups

Aboriginal

The prevalence of depression is significantly higher among Aboriginal Canadians that it is for the general Canadian population. A longitudinal survey done in 2005 found that 30 percent of First Nations people felt sad or depressed for 2 or more weeks that year,[115] while the rate among the general population is estimated to be 5 to 8 percent[5] In addition up to 44 years of age, suicide or self-inflicted harm is the leading cause of death among First Nations youth and adults.[115] Suicide rates among the Inuit are even higher than for First Nations at 6 to 11 times that national average.[116] A national study of mental health in Canada points out that despite high suicide rates, the numbers of Inuit suffering major depression using standard scales in 2001 was 3.1 percent, well below the national average.[4] The authors suggest that one possible explanation for this discrepancy may be that standard scales used in the general population may not be valid in Inuit culture.

Surveys suggest that Aboriginal Canadians seek help for mental health problems at a greater rate than Canadians overall.[4] While 8 percent of the Canadian population sought help from a professional regarding their mental/emotional health in 2000, the corresponding figure was as high as 17 percent among some Aboriginal groups.

One study[117] examined clinical utility and validity of the PDSS and the EPDS with First Nations and Metis women in Saskatchewan. The relatively lower prevalence of postpartum depression (17%) found may reflect that the sample came from an Aboriginal health centre that offers supportive, culturally appropriate perinatal care. The author cautions against generalizing these results, pointing out that the results may have been different if women from more remote, northern communities were included. The authors conclude that both the EPDS and the PDSS

demonstrated utility as screening tools in this population, but that the EPDS has better predictive power (25% PPV and 1% NPV over and above the PDSS).

Concerns about the overall poor health status of Aboriginal Canadians relative to the general Canadian population, have led to the development of a wide range of programs being developed at both the provincial and federal level. The Government of Canada's Aboriginal Portal website (http://www.aboriginalcanada.gc.ca/acp/site.nsf/eng/ao31174.html) has a list of over 50 programs and services relating to aboriginal health, including mental health issues. There was also an Aboriginal Health Transition Fund which dedicated $200 million over 5 years (2006-2011) to addressing the gap between Aboriginal and non-Aboriginal health status, including mental health and depression.

Rural/Remote

A study of women's health equity in Ontario[118] found that while rural residents were more likely to be hospitalized for depression, urban dwellers accounted for proportionately higher OHIP costs for mental health care. They also found that geographic patterns of use reflected supply of services more than they did need. The study did not find differences in prevalence of probable depression between rural and urban areas. While this study was specific to Ontario, a study of unmet need for depression treatment in Atlantic Canada[119] also found that while there was no significant rural/urban differences in depression prevalence, rural dwellers were significantly less likely to receive treatment for depression in either primary or specialty care.
Fisher and Copenhaver, (2006)[120] compared the results of five screening and survey tools validated for rural settings: Primary Care Evaluation of Mental Disorders, Mood Module (PRIME-MD), Short Form 36 (SF-36), Montgomery-Asberg Depression Rating Scale (MADRS), Mini-Mental State Exam (MMSE), and Global Assessment Score (GAS). All were applied by nurses to assess the mental health of older adults in rural public housing facilities in Pennsylvania. Besides social isolation and low income, this sample also was characterized by a high percentage of illiteracy and low education levels. The majority (95%) of participants were White women. Although the authors reported the results were "context bound", they were unable to form general conclusions about which tool was more useful due to small sample size, lack of randomization, self-selection, high rate of refusal to participate (47%), and self-reporting nature of the instruments.

Although evidence is limited, it appears that there is no significant rural/urban difference in terms of prevalence of depression, access to treatment for depression is less available in rural/remote areas of the country.

Newcomer/Immigrants

In looking at immigrant health, Ali et al., (2004)[121] reported that immigrants had significantly lower rates of depression than the general Canadian population with newly arrived immigrants having the lowest rates and long-term immigrants having rates closer to the general population.

They also found that immigrants from Asia had lower rates of depression than those from Europe and North America.

One important exception to this is postpartum women. In a study of newcomers to Canada, Stewart et al., (2008)[122] found that immigrant, asylum-seeking and refugee women had higher rates of postpartum depression than Canadian-born women (35.1%, 31.1%, 25.7% and 8.1% respectively) using the EPDS.

Smith et al., (2007)[65] explored the interaction of gender and income with the protective factor of recent immigration using data collected using the Composite International Diagnostic Interview Short Form (CIDI-SF) during the Canadian Community Health Survey (CCHS). Previous research has established that that depression is more prevalent is among immigrants that have become acculturated to Canada. This study found a differential effect by gender and income for recently arrived immigrants. Low-income male recent immigrants appear to have lower rates of depression than middle-high income men. Low-income recent women immigrants have a higher prevalence of depression than middle-high income women. This situation highlights the need for closer examination of various socio-demographic factors such as living alone, single parent status, and receiving government financial support. The findings are limited by being cross-sectional rather than longitudinal. CIDI-SF has not been validated as a detection tool and CIDI-SF may over-estimate prevalence among recent immigrants as they may be under-represented due to language and other cultural barriers to participation. Caution should be exercised in assuming homogeneity of the sample. Low-income may reflect transitional status, lack of familiarity with life in Canada and with the health care system. Also, social ties/capital and acquired language abilities may confer mental health benefits.[122]

Ethnicity/Race

There is little evidence to suggest that ethnicity itself is an important factor in prevalence rates of depression in Canada. In Atlantic Canada, Starkes et al., (2005)[119] found no significant connection between ethnicity and probable depression. Lin et al., (2009)[118] report that while some studies on depression and primary care have found racial disparities, this is not consistent suggesting that disparities vary across practice settings and that these care gaps can be closed.

Huang et al., (2006)[123] studied the PHQ-9 with the four largest racial/ethnic groups in the United States (African American, Chinese American, Latino and non-Hispanic White). The mean and total scores and the factor structure suggest that it can be used without adjustment in diverse populations.

Other

Lin et al., (2009)[118] in an Ontario-based study, found that there were gender and income differences in the prevalence of depression. Data obtained from the Canadian Community Health Survey (CCHS), Cycle 1.1.was used to assess the percentage of Ontarians aged 15 and older with probable depression in 2000/01. Depression was measured during this cycle of the CCHS using the Composite International Diagnostic Interview-short Form (CIDI-SF) for Major Depression. The CIDI-SF is a series of questions used to calculate the predicted probability of major depressive episodes occurring within the year preceding the interview. Respondents whose

predicted probability score was 0.9 or greater were considered to have probable depression.[118] It should be noted that the authors recognize that the CIDI-SF is not a fully validated measure and that the prevalence reported in their report " is not precise and very likely somewhat overestimates the actual population prevalence" (p.94). Using this measure Lin et al. found that 7.4 percent of people aged 15 and older in Ontario in 2000/2001 met the criteria for having probable depression. Women were found to be twice as likely as men to have probable depression (9.8 percent versus 4.9 percent, respectively). The prevalence of depression by level of income varied for women and men. Rates of depression were highest among those residing in the lowest income neighbourhoods (11.8 percent of women and 5.5 percent of men) compared to those residing in the highest-income neighbourhoods (8.5 percent of women and 4.2 percent of men). Perhaps of greater concern, only 40 percent of the Ontarians identified with probable depression using the CIDI-SF (41 percent of women and 37 percent of men) had at least one physician visit for depression within the year following the CCHS interview. The rate of physician visits for depression was highest among those aged 45 to 64 and lowest among those aged 15 to 24. The gender and income findings show some consistency with results of an earlier study done in Atlantic Canada. Starkes et al., (2005)[119] found that individuals who were female, widowed, separated or divorced, low income or with two or more co-morbid conditions were at greatest risk for depression. They also found that only 40 percent of respondents with probable depression screened using the CIDI-SF, (score 9 to 11) reported having had a consultation with a general practitioner or mental health specialist.

CQ5. What are patient preferences and values for treatment interventions antidepressants and/or psychotherapy) for depression?

A number of studies asked patients and non-patients what forms of treatment they preferred, and some of these studies included the option of no treatment at all. There is no clear pattern in the results, and the findings are often contradictory. In a German study of depressed patients, Lowe et al., (2006),[124] found that 25 percent of people wanted no intervention at all. Similarly, Dobscha, Corson, and Gerrity (2007)[125] found that the same proportion of VA out-patients preferred "watchful waiting;" and in Givens et al., (2007)[126] about one-quarter of people with elevated scores on the CES-D did not want any form of therapy. A smaller proportion of people wanted "watchful waiting" rather than active treatment (16%) in the Johnson et al., (2006)[127] study of depressed patients in primary care.

A relatively large number of studies asked both depressed and non-depressed people to choose among various treatment options. In almost all cases, people preferred psychotherapy or CBT combined with medication.[128-131] Exceptions were Raue et al., 2009[132] and Dobscha et al., 2007[125] who found only 17 percent wanted a combination of treatments.

When given a choice between therapy (psychotherapy or CBT) and medication, most studies found a strong preference for nonpharmacological treatment[124,128,131-138]. However, there were exceptions. Van Voorhees et al., 2005[139] found that among young adults who scored high on the CES-D, an equal number would accept medications or counseling (although about one-fourth would not accept a GP's diagnosis of depression at all); and Leykin et al., 2007[140] found slightly more people wanting medication rather than therapy. Givens et al., 2007[126] found that whites had a preference for medication over therapy, but that this was reversed among ethnic minority groups.

Even among those who were on medications or who would contemplate using drugs if they were depressed, a large proportion of people had reservations about becoming dependent on them[141] or their side effects.[130,142]

Stecker and Alvidrez (2007)[143] looked at what a small number (N = 21) of primary care depressed patients actually did. Although 90 percent thought that psychotherapy would be effective, only 21 percent actually initiated therapy, and only one person (3%) attended more than one session. Gum et al., (2006)[137] though, found that with depressed elderly patients, collaborative care improved access to treatment from 33 percent to 74 percent. The evidence is weak, but it appears that among people with depression about 75 percent want some form of therapy, but the follow-through is quite poor. Collaborative care may improve this. Collaborative care includes of interventions of varying intensity, ranging from simple interventions such as telephone communication to encourage compliance with medication to more complex interventions that involve intensive follow-up and incorporate a form of structured psychosocial intervention, but this current review did not focus on that factor.[144]

CQ6. What are the benefits and harms associated with the treatment (antidepressants and/or psychotherapy) for depression?

For the general population, three systematic reviews helped to inform the question of the benefits and harms of treatment for depression and are discussed here. One Cochrane review[145] examined the efficacy and tolerability of anti-depressant drugs in adult depression patients younger than age 65 administered in a primary care setting; and one study that meta-analyzed the psychological treatment of late-life depression.[146] The third review looked at the relative benefits and harms of second generation antidepressants (bupropion, citalopram, duloxetine, escitalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, sertraline, trazodone, and venlafaxine) for the treatment of depression in adults.[147]

The Cochrane review included 14 RCTs with 16 comparisons of which 10 examined tricyclic antidepressants (TCAs); two examined serotonin reuptake inhibitors (SSRIs), and two that examined both; all against placebos. Total participants were 1,364 in the treatment group and 919 in the control groups. All the included studies were of short duration (6 to 8 weeks). Pooled estimates for benefit (identified as from any response to remission) of TCAs over placebo is RR. 1.24 (95% CI, 1.11 to 138) in favour of the TCAs. The harms for this group were measured based on withdrawals due to adverse events RR 2.14 (95%CI, 1.41 to 3.26). The review also assessed the clinical benefit of SSRIs (2 medications Escitalopram and Citalopram) versus placebo with the resulting effective of benefit of the medication RR 1.28 (95% CI, 1.15 to 1.43). Withdrawals due to adverse events in these patient groups on SSRIs were 2.05 (95% CI, 1.11 to 3.75). The authors concluded that both TCAs and SSRIs are effective for depression treated in primary care.

The review by Gartlehner et al., (2009)[147] served as the background paper for guidelines from the American College of Physicians. After looking at over 200 included studies, the overall conclusion of the review was that, on the basis of efficacy and effectiveness, current evidence does not support choosing one second generation antidepressant over another. The most commonly reported minor adverse events were constipation, diarrhea, dizziness, headache and insomnia. On average at least one adverse events was experienced by 61 percent of participants

in included efficacy studies. The review found that the adverse events profiles of the second generation antidepressants were similar although there were differences in incidence rates of specific adverse events. The review also looked at rates for severe adverse events: sexual dysfunction, suicidality, seizures cardiovascular events, hyponatremia, hepatotoxicity, and serotonin syndrome. The most common severe adverse event was sexual dysfunction which was experienced by more than 50 percent of patients. The review found "moderate" evidence (using a modified GRADE approach) that bupropion had the lowest rates of sexual dysfunction of the drugs examined while paroxetine had the highest. There was insufficient evidence on the other severe adverse events to draw conclusions about comparative risk although there was "weak" evidence from observational studies that bupropion may increase risk for seizures, venlafaxine may increase risk for cardiovascular events and nefazodone might increase risk for hepatoxicity.

The effectiveness of psychological intervention in the treatment of depression in older adults was the focus of a 2006 review.[146] Seventeen comparison studies were included in which the comparison was psychological treatment to control groups (wait lists and usual care). The review authors analyzed psychological treatment versus antidepressant medications. Their analysis found a non-significant difference between the two treatment types with a mean effect size -0.01 (95% CI -0.26-0.24). In the studies that combined psychological treatment with antidepressant medication the mean effect was 0.50 (95% CI 0.13-0.87) in favour of combining treatment. When comparing different psychological treatment the review found no clear difference. The authors note that the treatment types were so different they could not pool or provide further analysis beyond stating that as a result of this difference, it was not possible to decide the effectiveness of one psychological intervention versus another. The authors also state that limitations of their review were the small number of includes studies and small effect size. In spite of this limitation they felt the evidence sufficiently supported the use of psychological treatment as a first step in treating depression in older adults.

Indirect Evidence on the benefit of treatment for depression
Rationale

With the lack of direct evidence on the effectiveness of screening for depression in all groups and especially in high risk populations the Evidence Review Synthesis Cente was commissioned by the Depression Working Group (WG) to search, retrieve and summarize systematic review data on the benefits of treating depression in high risk groups (people with chronic illness, chronic pain, perinatal women, people with alcohol or substance misuse disorders, and Aboriginal people) identified as groups of interest by the WG.

Methods

Quantitative systematic reviews offer concise summary of evidence about treatment effectiveness for depression. Searches of Medline, EMBASE, PsycINFO and Cochrane from 2001 to January 2012 were conducted. Titles, abstracts, and full text articles were reviewed by two blinded screeners using three inclusion criteria: 1) systematic review with meta-analysis, 2) pre-defined high risk adult population, and 3) psychological and/or pharmacological treatment intervention. Quality of all included reviews was independently evaluated by two screeners using the Assessment of Multiple Systematic Reviews (AMSTAR) instrument.

The search retrieved 2,040 titles and abstracts, In order to manage this large volume of literature, we decided to begin with reviews with meta-analysis of patients that included response and remission and that were published in the last 2 years. From that literature we located 20 high quality reviews for the groups of people with chronic disease, chronic pain or substance abuse. Finding no high quality reviews for perinatal women in the 2-year time frame, we looked back to January 2006 and located two high quality reviews. There were no reviews that assessed depression treatment in aboriginal people. A total of 21 systematic reviews were retained. We present results of those meta-analyses with statistically significant treatment effects. Results of all meta-analyses are included in the evidence tables (see Tables 12-16).

Results

A wide range of interventions were reported by these reviews with multiple meta-analyses for efficacy of treatment involving various interventions, treatment periods, and diagnostic instruments. Interventions were classified into two main categories: pharmacological and psychological treatments. The definition of an adequate response to treatment is is generally accepted as a 50 percent decrease in symptom severity.[73] Remission from depression is defined as being free or nearly free of symptoms for the current episode.[73] Response and/or remission were presented as an outcome of interest as relative risk (RR), odds ratio (OR),mean difference (MD) and standardized mean difference (SMD). One challenge in summarizing results of these reviews is that outcomes were not reported similarly. For those that report similar statistics (e.g., OR, RR, SMD, etc.) we report the range of these values.

Chronic illnesses

We identified 15 reviews[148-162] on the treatment of depression in people with chronic disease. Of these, six included populations with cardiovascular disease (CVD).[148-153] Three of these compared psychological treatment with usual care[150,152,153], two compared pharmacological treatment with placebo[149,151], and the remaining study compared both pharmacological and psychological treatment with placebo.[148] All six of the CVD reviews reported statistically significant effects of medication or psychological treatment on depression related outcomes. Psychotherapy treatments for depression in populations with CVD did result in small to moderate improvements in depression outcomes (response); SMD Range: -0.21 to 0.65.[150,153]Pharmacological treatment resulted in higher remission rates :OR 1.80 (1.18 to 2.74);[149] MD 2.27 (0.60 to 3.94)[149] For treatment response, antidepressants were superior to placebo for decreasing depressive symptoms: weighted mean difference 1.41 (0.53 to 2.29);[149] OR 1.72 (1.17 to 2.54);[149] SMD: -0.24 (-0.38 to -0.09).[148] See Table 12 for details.

There were 9 reviews[154-162] on depression treatment involving populations with chronic diseases, such as cancer and diabetes mellitus, all showing statistically significant benefit of treatment. Psychotherapy was reported to result in remission: RR = 2.0 (1.1 to 3.5).[159] and response: RR = 1.5 (1.1 to 2.1);[159] SMD -0.83 to -0.16;[162] d= 0.37 to 0.80[157,159] Pharmacotherapy was reported to result in remission: RR 0.81 to 0.88,[156] and response: OR 2.08 to 2.13:[154] RR 0.83 to 0.89:[156] SMD -0.51 to -0.27.[156,160] Table 13 shows further detail.

Chronic pain

In people with chronic pain two Cochrane reviews[163,164] indicated that neither psychological treatment SMD=-0.11 (-0.67, 0.44), nor pharmacological treatment SMD =0.06 (-0.29, 0.40) had a statistically significant result. See Table 14 for details.

Alcohol and substance abuse disorder

For people with alcohol and substance abuse problems the results are mixed. One review of psychological interventions versus placebo[165,166] indicates no effect whereas for those receiving pharmacological treatment there was a positive effect: RR 2.03 (1.17 to 3.53)[165] See Table 15 for details.

Perinatal women

In perinatal women one Cochrane review[167] and one non-Cochrane systematic review[168] indicated that psychological treatment have a statistically significant effect on depression outcomes. These systematic reviews concluded that the interventions under study were useful in the treatment or prevention of perinatal depression. Psychological interventions resulted in response: RR 0.48 to 0.80;[168] d: 0.61 (0.37 to 0.85).[167] See Table 16 for details.

Aboriginal people

We did not find any systematic reviews providing evidence on effectiveness of treatment for depression in aboriginal population.

Limitations/Cautions

There are several limitations to this data. The search, appraisal and synthesis were undertaken in a expedited review of systematic reviews with meta-analysis. In this section we have only reported on the benefit of treatment not the harms. We purposely chose to report only the one outcome that was common in all the included reviews. Many of the included reviews contained other patient important outcomes that have not been reported here. Finally we did not analyze the reviews for overlap of included studies so we do not know how many reviews reported the results of the same primary studies.

Chapter 4. Discussion

We examined the USPSTF review [58] to determine if the CTFPHC Depression guideline group should update that review. It was determined that the the USPSTF review did not include data on high risk groups therefore this current review is de novo. For the question of the benefit of screening we found no direct evidence for the population as a whole, rather we have included five studies conducted by the same primary researcher in the elderly in rural Japan. The USPSTF included one paper Williams et al [169] as a screen versus no screen comparison. Our team concluded that this paper it did not meet our inclusion criteria because the goal of this study was to determine the difference between two screening tools and all participants underwent a diagnostic interview which meant that the participants were preselected for depression and not an asymptomatic population.

Five studies met the inclusion criteria for this review; however, the results provide limited evidence on the effectiveness of screening for depression in the general population or high risk groups. We found no studies on harms of screening for depression that met our inclusion criteria. These results are consistent with previous guidelines and evidence reviews. The USPSTF 2009[58] found no evidence for the 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 NICE[170] and SIGN[171]; neither recommend screening of asymptomatic people in the general population. The NICE guideline for people with chronic illness recommend that physicians remain alert to the possibility of depression [79] and another for perinatal women[172] recommended screening postpartum women, yet those recommendations are based on the indirect evidence of the benefit of treatment rather than the direct evidence of the effectiveness of screening or case finding for depression.

We included five quasi-experimental pre- and post-implementation design studies, with an intervention community and neighboring community controls in an elderly population in Japan.[85-89] The principle investigator, Oyama, was consistent for all five studies. The intervention communities received screening for depression, followup with mental health care or psychiatric treatment, and psychoeducation in the community setting. The results of the meta-analysis in intervention areas compared to control areas suggest that the implementation of a community-based suicide intervention program, including programs involving community-based depression screening (CDS) and health education lead to statistically significant reductions in suicide only among women ages 75 to 84.. There was no statistically significant effect of screening on the overall suicide rate of elderly men. These studies had a quasi-experimental design, and as such are prone to bias and confounding. However, all five studies reported similar demographics, including similar socioeconomic characteristics between the intervention and control communities. Time-dependent confounding (e.g., regression towards the mean) inherent in the before-and-after design of these studies presents another limitation. However, these studies dealt with this regression effect to some degree because of the statistically unchanged risks in the control population, indicating that risk reductions in the target population were still considerably greater than secular trends.

The generalizability of the finding of the Oyama studies should be viewed with caution as Japan has a national suicide rate much higher than Canada or the United States. In the case of elderly

women in the age group that showed benefit, the Japanese suicide rate is over 7 times higher than the Canadian rate (23.4 versus 3.3 per 100,000 respectively).[173] In addition, the regions of study in Japan had average rates of suicide much higher than even the Japanese average. [85-89]

Limitations

The findings of this review are affected by the limitations of the included literature. We limited our search to papers written in English or French. There is the potential that we have missed the opportunity to analyze data from papers written in other languages. The studies that were reviewed here evaluated the effectiveness of the community-based depression screening programs which incorporated screening for depression, follow-up with mental health care or psychiatric treatment, and health education in the community setting in rural Japan with higher than average rates of suicide. As such, the observed reduction in suicide rates or recovery from depression cannot be attributed solely to the screening component of these programs.

Conclusion

The ultimate goal of screening for depression is to decrease morbidity and mortality related to this disease. There is a very limited research evidence from which to draw any conclusions on the effectiveness of screening for depression in the general or high risk populations.

Reference List

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed. Washington, DC: American Psychiatric Association; 1994.

2. Murray CJ and Lopez AD. Evidence-based health policy--lessons from the Global Burden of Disease Study. Science. 1996; 274(5288):740-3. PM:8966556.

3. Wakefield JC, Schmitz MF, First MB, and Horwitz AV. Extending the bereavement exclusion for major depression to other losses: evidence from the National Comorbidity Survey. Arch Gen Psychiatry. 2007; 64(4):433-40. PM:17404120.

4. Government of Canada. The human face of mental health and mental illness in Canada. HP5-19/2006E. Ottawa, ON: Minister of Public Works and Government Services Canada; 2006. Available at: http://www.phac-aspc.gc.ca/publicat/human-humain06/index-eng.php.

5. Patten SB, Wang JL, Williams JV, Currie S, Beck CA, Maxwell CJ, and El Guebaly N. Descriptive epidemiology of major depression in Canada. Can J Psychiatry. 2006; 51(2):84-90. PM:16989107.

6. Vasiliadis H, Lesage A, Adair C, Wang PS, and Kessler RC. Do Canada and the United States differ in prevalence of depression and utilization of services? Psychiatr Serv. 2007; 58(1):63-71. PM:17215414.

7. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, Rush AJ, Walters EE, Wang PS, and National Comorbidity SR. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA. 2003; 289(23):3095-105. PM:12813115.

8. 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, Lepine 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, Martinez-Alonso M, Matschinger H, Mazzi F, Morgan Z, Morosini P, Palacin C, Romera B, Taub N, and Vollebergh WA. Prevalence of mental disorders in Europe: Results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatr Scand. 2004; (420):21-7. PM:15128384.

9. The Standing Senate Committee on Social Affairs SaT. Out of the shadows at last: transforming mental health, mental illness and addiction services in Canada. Ottawa: Government of Canada; 2006. Available at: http://www.parl.gc.ca/39/1/parlbus/commbus/senate/com-e/soci-e/rep-e/rep02may06-e.htm.

10. Greenberg PE, Kessler RC, Birnbaum HG, Leong SA, Lowe SW, Berglund PA, and Corey-Lisle PK. The economic burden of depression in the United States: how did it change between 1990 and 2000? J Clin Psychiatr. 2003; 64(12):1465-75. PM:14728109.

11. Stephens T and Joubert N. The economic burden of mental health problems in Canada. Chronic Dis Can. 2001; 22(1):18-23. PM:11397346.

12. Tompson MC, Pierre CB, Boger KD, McKowen JW, Chan PT, and Freed RD. Maternal depression, maternal expressed emotion, and youth psychopathology. J Abnorm Child Psychol. 2010; 38(1):105-17. PM:19693663.

13. Eckshtain D, Ellis DA, Kolmodin K, and Naar-King S. The effects of parental depression and parenting practices on depressive symptoms and metabolic control in urban youth with insulin dependent diabetes. J Pediatr Psychol. 2010; 35(4):426-35. PM:19710249.

14. Ramchandani P and Stein A. The impact of parental psychiatric disorder on children. BMJ. 2003; 327(7409):242-3. PM:12896914.

15. Wells KB, Burnam MA, Rogers W, Hays R, and Camp P. The course of depression in adult outpatients: results from the Medical Outcomes Study. Arch Gen Psychiatry. 1992; 49(10):788-94. PM:1417431.

16. Eaton WW, Anthony JC, Gallo J, Cai G, Tien A, Romanoski A, Lyketsos C, and Chen LS. Natural history of Diagnostic Interview Schedule/DSM-IV major depression: the Baltimore Epidemiologic Catchment Area follow-up. Arch Gen Psychiatry. 1997; 54(11):993-9. PM:9366655.

17. Shea MT, Elkin I, Imber SD, Sotsky SM, Watkins JT, Collins JF, Pilkonis PA, Beckham E, Glass DR, and Dolan RT. Course of depressive symptoms over follow-up: findings from the National Institute of Mental Health Treatment of Depression Collaborative Research Program. Arch Gen Psychiatry. 1992; 49(10):782-7. PM:1417430.

18. Heisel MJ. Suicide and its prevention among older adults. Can J Psychiatry. 2006; 51(3):143-54. PM:16618005.

19. Conference Board of Canada. Suicides. How CanadA Performs Website. 2009. Avalable at: http://www.conferenceboard.ca/hcp/Details/society/suicides.aspx.

20. Patten SB, Williams JV, Lavorato DH, Modgill G, Jette N, and Eliasziw M. Major depression as a risk factor for chronic disease incidence: longitudinal analyses in a general population cohort. Gen Hosp Psychiatry. 2008; 30(5):407-13. PM:18774423.

21. Patten SB, Williams JV, Lavorato DH, Campbell NR, Eliasziw M, and Campbell TS. Major depression as a risk factor for high blood pressure: epidemiologic evidence from a national longitudinal study. Psychosom Med. 2009; 71(3):273-9. PM:19196807.

22. Brown LC, Majumdar SR, Newman SC, and Johnson JA. History of depression increases risk of type 2 diabetes in younger adults. Diabet Care. 2005; 28(5):1063-7. PM:15855568.

23. Gilmour H. Depression and risk of heart disease. Health Rep. 2008; 19(3):7-17. PM:18847141.

24. Gariepy G, Wang J, Lesage AD, and Schmitz N. The longitudinal association from obesity to depression: Results from the 12-year National Population Health Survey. Obesity. 2010; 18(5):1033-8. PM:19816409.

25. Patten SB, Williams JV, Lavorato DH, Brown L, McLaren L, and Eliasziw M. Major depression, antidepressant medication and the risk of obesity. Psychother Psychosom. 2009; 78(3):182-6. PM:19321971.

26. Breslau N, Peterson EL, Schultz LR, Chilcoat HD, and Andreski P. Major depression and stages of smoking: a longitudinal investigation. Arch Gen Psychiatry. 1998; 55(2):161-6. PM:9477930.

27. Khaled SM, Bulloch A, Exner DV, and Patten SB. Cigarette smoking, stages of change, and major depression in the Canadian population. Can J Psychiatry. 2009; 54(3):204-8. PM:19321025.

28. Patten SB, Williams JV, Lavorato DH, and Eliasziw M. Major depression and injury risk. Can J Psychiatry. 2010; 55(5):313-8. PM:20482958.

29. Patten SB, Williams JV, Lavorato DH, and Eliasziw M. The effect of major depression on participation in preventive health care activities. BMC Pub Health. 2009; 9:87. PM:19320983.

30. Patten SB, Wang JL, Williams JV, Lavorato DH, Bulloch A, and Eliasziw M. Prospective evaluation of the effect of major depression on working status in a population sample. Can J Psychiatry. 2009; 54(12):841-5. PM:20047723.

31. Beck CA and Patten SB. Adjustment to antidepressant utilization rates to account for depression in remission. Compr Psychiatry. 2004; 45(4):268-74. PM:15224269.

32. Brugha T, Jenkins R, Bebbington P, Meltzer H, Lewis G, and Farrell M. Risk factors and the prevalence of neurosis and psychosis in ethnic groups in Great Britain. Soc Psychiatry Psychiatr Epidemiol. 2004; 39(12):939-46. PM:15583900.

33. Kessler RC, Demler O, Frank RG, Olfson M, Pincus HA, Walters EE, Wang P, Wells KB, and Zaslavsky AM. Prevalence and treatment of mental disorders, 1990 to 2003. N Engl J Med. 2005; 352(24):2515-23. PM:15958807.

34. Gonzalez HM, Vega WA, Williams DR, Tarraf W, West BT, and Neighbors HW. Depression care in the United States: too little for too few. Arch Gen Psychiatry. 2010; 67(1):37-46. PM:20048221.

35. Herman S, Blumenthal JA, Babyak M, Khatri P, Craighead WE, Krishnan KR, and Doraiswamy PM. Exercise therapy for depression in middle-aged and older adults: predictors of early dropout and treatment failure. Health Psychol. 2002; 21(6):553-63. PM:12433007.

36. Cassano P and Fava M. Depression and public health: an overview. J Psychosom Res. 2002; 53(4):849-57. PM:12377293.

37. Drevets WC, Price JL, and Furey ML. Brain structural and functional abnormalities in mood disorders: implications for neurocircuitry models of depression. Brain Struct Funct. 2008; 213(1-2):93-118. PM:18704495.

38. Beck AT. The evolution of the cognitive model of depression and its neurobiological correlates. Am J Psychiatry. 2008; 165(8):969-77. PM:18628348.

39. Spijker J, De Graaf R, Bijl RV, Beekman AT, Ormel J, and Nolen WA. Duration of major depressive episodes in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). Br J Psychiatry. 2002; 181:208-13. PM:12204924.

40. Kupfer DJ. Long-term treatment of depression. J Clin Psychiatr. 1991; 52 Suppl:28-34. PM:1903134.

41. Satyanarayana S, Enns MW, Cox BJ, and Sareen J. Prevalence and correlates of chronic depression in the Canadian community health survey: mental health and well-being. Can J Psychiatry. 2009; 54(6):389-98. PM:19527559.

42. Ciechanowski PS, Katon WJ, and Russo JE. Depression and diabetes: impact of depressive symptoms on adherence, function, and costs. Arch Intern Med. 2000; 160(21):3278-85. PM:11088090.

43. Aro AR, de Koning HJ, Absetz P, and Schreck M. Psychosocial predictors of first attendance for organised mammography screening. J Med Screen. 1999; 6(2):82-8. PM:10444726.

44. McIntyre RS, Soczynska JK, Konarski JZ, and Kennedy SH. The effect of antidepressants on glucose homeostasis and insulin sensitivity: synthesis and mechanisms. Expert Opin Drug Saf. 2006; 5(1):157-68. PM:16370964.

45. Murphy JM, Horton NJ, Monson RR, Laird NM, Sobol AM, and Leighton AH. Cigarette smoking in relation to depression: historical trends from the Stirling County Study. Am J Psychiatry. 2003; 160(9):1663-9. PM:12944343.

46. Van Gool CH, Kempen GI, Penninx BW, Deeg DJ, Beekman AT, and van Eijk JT. Relationship between changes in depressive symptoms and unhealthy lifestyles in late middle aged and older persons: results from the Longitudinal Aging Study Amsterdam. Age Ageing. 2003; 32(1):81-7. PM:12540353.

47. Corcos M, Guilbaud O, Hjalmarsson L, Chambry J, and Jeammet P. Cytokines and depression: an analogic approach. Biomed Pharmacother. 2002; 56(2):105-10. PM:12000135.

48. Kop WJ, Gottdiener JS, Tangen CM, Fried LP, McBurnie MA, Walston J, Newman A, Hirsch C, and Tracy RP. Inflammation and coagulation factors in persons > 65 years of age with symptoms of depression but without evidence of myocardial ischemia. Am J Cardiol. 2002; 89(4):419-24. PM:11835923.

49. Musselman DL, Miller AH, Porter MR, Manatunga A, Gao F, Penna S, Pearce BD, Landry J, Glover S, McDaniel JS, and Nemeroff CB. Higher than normal plasma interleukin-6 concentrations in cancer patients with depression: preliminary findings. Am J Psychiatry. 2001; 158(8):1252-7. PM:11481159.

50. Penninx BW, Kritchevsky SB, Yaffe K, Newman AB, Simonsick EM, Rubin S, Ferrucci L, Harris T, and Pahor M. Inflammatory markers and depressed mood in older persons: results from the Health, Aging and Body Composition study. Biol Psychiatr. 2003; 54(5):566-72. PM:12946885.

51. Taylor CB, Youngblood ME, Catellier D, Veith RC, Carney RM, Burg MM, Kaufmann PG, Shuster J, Mellman T, Blumenthal JA, Krishnan R, and Jaffe AS. Effects of antidepressant medication on morbidity and mortality in depressed patients after myocardial infarction. Arch Gen Psychiatry. 2005; 62(7):792-8. PM:15997021.

52. Wassertheil-Smoller S, Shumaker S, Ockene J, Talavera GA, Greenland P, Cochrane B, Robbins J, Aragaki A, and Dunbar-Jacob J. Depression and cardiovascular sequelae in postmenopausal women: the Women's Health Initiative (WHI). Arch Intern Med. 2004; 164(3):289-98. PM:14769624.

53. Mood Disorders Society of Canada. Quick facts: mental illness & addiction in Canada (Third Edition). 2009. Available at: http://www.mooddisorderscanada.ca/page/quick-facts.

54. Adler DA, McLaughlin TJ, Rogers WH, Chang H, Lapitsky L, and Lerner D. Job performance deficits due to depression. Am J Psychiatry. 2006; 163(9):1569-76. PM:16946182.

55. Lerner D, Adler DA, Chang H, Berndt ER, Irish JT, Lapitsky L, Hood MY, Reed J, and Rogers WH. The clinical and occupational correlates of work productivity loss among employed patients with depression. J Occup Environ Med. 2004; 46(6 Suppl):S46-S55. PM:15194895.

56. Hoge CW, Lesikar SE, Guevara R, Lange J, Brundage JF, Engel CC, Jr., Messer SC, and Orman DT. Mental disorders among U.S. military personnel in the 1990s: association with high levels of health care utilization and early military attrition. Am J Psychiatry. 2002; 159(9):1576-83. PM:12202280.

57. Rapaport MH, Clary C, Fayyad R, and Endicott J. Quality-of-life impairment in depressive and anxiety disorders. Am J Psychiatry. 2005; 162(6):1171-8. PM:15930066.

58. O'Connor EA, Whitlock EP, Beil TL, and Gaynes BN. Screening for depression in adult patients in primary care settings: a systematic evidence review. Ann Intern Med. 2009; 151(11):793-803. PM:19949145.

59. National Institute for Health and Clinical Excellence. Depression: the treatment and management of depression in adults. NICE Clinical Guideline 90. London, UK: National Institute for Health and Clinical Excellence; 2009. Available at: http://www.nice.org.uk/nicemedia/pdf/CG90NICEguideline.pdf.

60. O'Connor EA, Whitlock EP, Gaynes BN, and Beil TL. Screening for depression in adults and older adults in primary care: an updated systematic review. Evidence Syntheses, No. 75. Rockville, MD: Agency for Healthcare and Research Quality; 2009. Available at: PM:20722174.

61. National Collaborating Centre for Mental Health. Depression (amended): management of depression in primary and secondary care. National Clinical Practice Guideline Number 23. London, UK: National Institute for Clinical Excellence (NICE); 2004. Available at: http://www.nccmh.org.uk/downloads/Depression/cg023fullguideline.pdf.

62. Hirschfeld RM and Weissman MM. Risk factors for major depression and bipolar disorder. Davis KL, Charney D, Coyle JT, Nemeroff C, editors, In: Neuropsychopharmacology: The Fifth Generation of Progress. American College of Neuropsychopharmacology; 2002. Chapter 70, p. 1017-25.

63. Patten SB and Juby H. A profile of clinical depression in Canada. Research Data Centres of Canada; 2008. Available at: http://hdl.handle.net/1880/46327.

64. Wang J, Williams J, Lavorato D, Schmitz N, Dewa C, and Patten SB. The incidence of major depression in Canada: the National Population Health Survey. J Affect Disord. 2010; 123:158-63. http://www.sciencedirect.com/science/article/pii/S0165032709003437.

65. Smith KLW, Matheson FI, Moineddin R, and Glazier RH. Gender, income and immigration differences in depression in Canadian urban centres. Can J Public Health. 2007; 98(2):149-53. PM:17441541.

66. Sartorius N, Ustun TB, Lecrubier Y, and Wittchen HU. Depression comorbid with anxiety: results from the WHO study on psychological disorders in primary health care. Br J Psychiatry. 1996; (30):38-43. PM:8864147.

67. MacMillan HL, Patterson CJS, and Wathen CN. Screening for depression in primary care: updated recommendations from the Canadian Task Force on Preventative Health Care. London, ON: Canadian Task Force on Preventative Health Care; 2004. Available at: http://www.canadiantaskforce.ca/docs/CTF_Depression_TR_2004_final.pdf.

68. Pignone MP, Gaynes BN, Rushton JL, Burchell CM, Orleans CT, Mulrow CD, and Lohr KN. Screening for depression in adults: a summary of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med. 2002; 136(10):765-76. PM:12020146.

69. National Collaborating Centre for Mental Health. Depression. The treatment and management of depression in adults (updated edition). NICE Clinical Guideline 90. London, UK: National Institute for Health and Clinical Excellence (NICE); 2009. Available at: http://publications.nice.org.uk/depression-in-adults-cg90.

70. Palmer SC and Coyne JC. Screening for depression in medical care: pitfalls, alternatives, and revised priorities. J Psychosom Res. 2003; 54(4):279-87. PM:12670603.

71. Bower P and Gilbody S. Stepped care in psychological therapies: access, effectiveness and efficiency. Narrative literature review. Br J Psychiatry. 2005; 186:11-7. PM:15630118.

72. Bilsker D, Goldner EM, and Jones W. Health service patterns indicate potential benefit of supported self-management for depression in primary care. Can J Psychiatry. 2007; 52(2):86-95. PM:17375863.

73. McIntyre RS, Fallu A, and Konarski JZ. Measurable outcomes in psychiatric disorders: remission as a marker of wellness. Clin Ther. 2006; 28(11):1882-91. PM:17213009.

74. Paykel ES, Ramana R, Cooper Z, Hayhurst H, Kerr J, and Barocka A. Residual symptoms after partial remission: an important outcome in depression. Psychol Med. 1995; 25(6):1171-80. PM:8637947.

75. Feightner JW. Early detection of depression. Chapter 39. Ottawa: Health Canada; 1994. Available at: http://www.phac-aspc.gc.ca/publicat/clinic-clinique/pdf/s6c39e.pdf.

76. Berg AO, Allan JD, Frame PS, Homer CJ, Johnson MS, Klein JD, Lieu TA, Mulrow CD, Orleans CT, Peipert JF, Pender NJ, Siu AL, Teutsch SM, Westhoff C, and Woolf SH. Screening for depression: recommendations and rationale. Ann Intern Med. 2002; 136(10):760-4. PM:12020145.

77. Pignone M, Gaynes BN, Rushton JL, Mulrow CD, Orleans CT, Whitener BL, Milles C, and Lohr KN. Screening for depression: systematic evidence reviews, No. 6. 02-S002. Rockville, MD: Agency for Healthcare Research and Quality; 2002. Available at: PM:20722156.

78. MacMillan HL, Patterson CJ, Wathen CN, Feightner JW, Bessette P, Elford RW, Feig DS, Langley J, Palda VA, Patterson C, Reeder BA, Walton R, and Canadian Task Force on Preventive Health Care. Screening for depression in primary care: recommendation statement from the Canadian Task Force on Preventive Health Care. Can Med Assoc J. 2005; 172(1):33-5. PM:15632399.

79. National Collaborating Centre for Mental Health. Depression in adults with a chronic physical health problem: treatment and management. NICE Clinical Guideline 91. London, UK: National Institute for Health and Clinical Excellence (NICE); 2009. Available at: http://www.nice.org.uk/CG91.

80. Effective Public Health Practice Project (EPHPP). Quality assessment tool for quantitative studies. 2007. Available at: http://www.ephpp.ca/tools.html.

81. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ, and Grade Working Group. GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008; 336(7650336):924-6. PM:18436948.

82. GRADEpro. [Computer program]. Version 3.2 for Windows [computer program]. 2008.

83. GRADE working group. 2000. Available at: http://www.gradeworkinggroup.org/.

84. DerSimonian R and Laird N. Meta-analysis in clinical trials. Contr Clin Trials. 1986; 7(3):177-88. PM:3802833.

85. Oyama H, Fujita M, Goto M, Shibuya H, and Sakashita T. Outcomes of community-based screening for depression and suicide prevention among Japanese elders. Gerontologist. 2006; 46(6):821-6. PM:17169937.

86. Oyama H, Koida J, Sakashita T, and Kudo K. Community-based prevention for suicide in elderly by depression screening and follow-up. Community Ment Health J. 2004; 40(3):249-63. PM:15259630.

87. Oyama H, Goto M, Fujita M, Shibuya H, and Sakashita T. Preventing elderly suicide through primary care by community-based screening for depression in rural Japan. Crisis. 2006; 27(2):58-65. PM:16913326.

88. Oyama H, Ono Y, Watanabe N, Tanaka E, Kudoh S, Sakashita T, Sakamoto S, Neichi K, Satoh K, Nakamura K, and Yoshimura K. Local community intervention through depression screening and group activity for elderly suicide prevention. Psychiatry Clin Neurosci. 2006; 60(1):110-4. PM:16472368.

89. Oyama H, Sakashita T, Hojo K, Ono Y, Watanabe N, Takizawa T, Sakamoto S, Takizawa S, Tasaki H, and Tanaka E. A community-based survey and screening for depression in the elderly: the short-term effect on suicide risk in Japan. Crisis. 2010; 31(2):100-8. PM:20418216.

90. Kessler RC and 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; 13(2):93-121. PM:15297906.

91. Zung WW. A self-rating depression scale. Arch Gen Psychiatry. 1965; 12:63-70. PM:14221692.

92. Hoyl MT, Alessi CA, Harker JO, Josephson KR, Pietruszka FM, Koelfgen M, Mervis JR, Fitten LJ, and Rubenstein LZ. Development and testing of a five-item version of the geriatric depression scale. J Am Geriatr Soc. 1999; 47(7):873-8. PM:10404935.

93. Takahashi R and Sakurai M. International standardization of clinical assessment of psychiatric symptoms. Jpn J Clin Psychiatr. 1977; (6):385-8.

94. Marcus G. The ICD-9 and the SADD-criteria for depression. Acta Psychiatr Scand Suppl. 1983; 310:31-41. PM:6582769.

95. Cochrane handbook for systematic reviews of interventions. Version 5.1.0 ed. New York, NY: John Wiley & Sons, Ltd. Publications; 2011.

96. Wells, G. A., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., and Tugwell, P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. 2008. Available at: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.

97. Kirpinar I, Gozum S, and Pasinlioglu T. Prospective study of postpartum depression in eastern Turkey prevalence, socio-demographic and obstetric correlates, prenatal anxiety and early awareness. J Clin Nurs. 2010; 19(3-4):422-31. PM:20500282.

98. Dennis CL. Can we identify mothers at risk for postpartum depression in the immediate postpartum period using the Edinburgh Postnatal Depression Scale? J Affect Disord. 2004; 78(2):163-9. PM:14706728.

99. Dennis CL and Ross LE. Depressive symptomatology in the immediate postnatal period: Identifying maternal characteristics related to true- and false-positive screening scores. Can J Psychiatry. 2006; 51(5):265-73. PM:16986815.

100. Kim JJ, Gordon TE, La Porte LM, Adams M, Kuendig JM, and Silver RK. The utility of maternal depression screening in the third trimester. Am J Obstet Gynecol. 2008; 199(5):509-e1-5. PM:18533122.

101. Lee DT, Yip AS, Chan SS, Tsui MH, Wong WS, and Chung TK. Postdelivery screening for postpartum depression. Psychosom Med. 2003; 65(3):357-61. PM:12764207.

102. Sheeder J, Kabir K, and Stafford B. Screening for postpartum depression at well-child visits: is once enough during the first 6 months of life? Pediatrics. 2009; 123(6):e982-e988. PM:19482749.

103. Davies BR, Howells S, and Jenkins M. Early detection and treatment of postnatal depression in primary care. J Adv Nurs. 2003; 44(3):248-55. PM:14641394.

104. Rychnovsky J and Beck CT. Screening for postpartum depression in military women with the Postpartum Depression Screening Scale. Mil Med. 2006; 171(11):1100-4. PM:17153549.

105. Kaaya SF, Mbwambo JK, Kilonzo GP, Van Den Borne H, Leshabari MT, Fawzi MC, and Schaalma H. Socio-economic and partner relationship factors associated with antenatal depressive morbidity among pregnant women in Dar es Salaam, Tanzania. Tanzan J Health Res. 2010; 12(1):23-35. PM:20737826.

106. Turvey CL, Willyard D, Hickman DH, Klein DM, and Kukoyi O. Telehealth screen for depression in a chronic illness care management program. Telemed J E Health. 2007; 13(1):51-6. PM:17309355.

107. U.S.Preventive Services Task Force. Screening for depression in adults: U.S. preventive services task force recommendation statement. Ann Intern Med. 2009; 151(11):784-92. PM:19949144.

108. Valenstein M, Vijan S, Zeber JE, Boehm K, and Buttar A. The cost-utility of screening for depression in primary care. Ann Intern Med. 2001; 134(5):345-60. PM:11242495.

109. Leigh B and Milgrom J. Acceptability of antenatal screening for depression in routine antenatal care. Aust J Adv Nurs. 2007; 24(3):14-8. PM:17518160.

110. Buist A, Condon J, Brooks J, Speelman C, Milgrom J, Hayes B, Ellwood D, Barnett B, Kowalenko N, Matthey S, Austin MP, and Bilszta J. Acceptability of routine screening for perinatal depression. J Affect Disord. 2006; 93(1-3):233-7. PM:16647761.

111. Gemmill AW, Leigh B, Ericksen J, and Milgrom J. A survey of the clinical acceptability of screening for postnatal depression in depressed and non-depressed women. BMC Pub Health. 2006; 6:211. PM:16914061.

112. Allen K, Cull A, and Sharpe M. Diagnosing major depression in medical outpatients: acceptability of telephone interviews. J Psychosom Res. 2003; 55(4):385-7. PM:14507551.

113. Esler D, Johnston F, Thomas D, and Davis B. The validity of a depression screening tool modified for use with Aboriginal and Torres Strait Islander people. Aust N Z J Public Health. 2008; 32(4):317-21. PM:18782392.

114. Esler DM, Johnston F, and Thomas D. The acceptability of a depression screening tool in an urban, Aboriginal community-controlled health service. Aust N Z J Public Health. 2007; 31(3):259-63. PM:17679245.

115. First Nations & Inuit Health: mental health and wellness. 2011. Available at: http://www.hc-sc.gc.ca/fniah-spnia/promotion/mental/index-eng.php.

116. First Nations, Métis, and nuit GBA. Depression: A culturally relevant perspective on the issue. 2009. Available at: http://www.aboriginalgba.ca/category.aspx?catid=117&rt=1.

117. Clarke PJ. Validation of two postpartum depression screening scales with a sample of First Nations and Metis women. Can J Nurs Res. 2008; 40(1):113-25. PM:18459275.

118. Lin E, Diaz-Granados N, Sterwart D, Rhodes A, Yeritsyan N, Johns A, Duong-Hua M, and Bierman AS. Depression: project for an Ontario women's health evidence-based report. Bierman AS, editor, In: Ontario Women's Health Evidence-Based Report (Power Study). Volume 1 ed. Toronto, ON: Keenan Research Centre in the Li Ka Shing Knowledge Institute, St. Michael's Hospital; 2009. Chapter 5.

119. Starkes JM, Poulin CC, and Kisely SR. Unmet need for the treatment of depression in Atlantic Canada. Can J Psychiatry. 2005; 50(10):580-90. PM:16276848.

120. Fisher KM and Copenhaver V. Assessing the mental health of rural older adults in public housing facilities: a comparison of screening tools. J Gerontol Nurs. 2006; 32(9):26-33. PM:16972606.

121. Ali JS, McDermott S, and Gravel RG. Recent research on immigrant health from statistics Canada's population surveys. Can J Public Health. 2004; 95(3):I9-13. PM:15191126.

122. Stewart DE, Gagnon A, Saucier JF, Wahoush O, and Dougherty G. Postpartum depression symptoms in newcomers. Can J Psychiatry. 2008; 53(2):121-4. PM:18357931.

123. Huang FY, Chung H, Kroenke K, Delucchi KL, and Spitzer RL. Using the Patient Health Questionnaire-9 to measure depression among racially and ethnically diverse primary care patients. J Gen Intern Med. 2006; 21(6):547-52. PM:16808734.

124. Lowe B, Schulz U, Grafe K, and Wilke S. Medical patients' attitudes toward emotional problems and their treatment. What do they really want? J Gen Intern Med. 2006; 21(1):39-45. PM:16423121.

125. Dobscha SK, Corson K, and Gerrity MS. Depression treatment preferences of VA primary care patients. Psychosom. 2007; 48(6):482-8. PM:18071094.

126. Givens JL, Houston TK, Van Voorhees BW, Ford DE, and Cooper LA. Ethnicity and preferences for depression treatment. Gen Hosp Psychiatry. 2007; 29(3):182-91. PM:17484934.

127. Johnson MD, Meredith LS, Hickey SC, and Wells KB. Influence of patient preference and primary care clinician proclivity for watchful waiting on receipt of depression treatment. Gen Hosp Psychiatry. 2006; 28(5):379-86. PM:16950372.

128. Hodges L, Butcher I, Kleiboer A, McHugh G, Murray G, Walker J, Wilson R, and Sharpe M. Patient and general practitioner preferences for the treatment of depression in patients with cancer: How, who, and where? J Psychosom Res. 2009; 67(5):399-402. PM:19837202.

129. Kocsis JH, Leon AC, Markowitz JC, Manber R, Arnow B, Klein DN, and Thase ME. Patient preference as a moderator of outcome for chronic forms of major depressive disorder treated with nefazodone, cognitive behavioral analysis system of psychotherapy, or their combination. J Clin Psychiatr. 2009; 70(3):354-61. PM:19192474.

130. Hanson AE and Scogin F. Older adults' acceptance of psychological, pharmacological, and combination treatments for geriatric depression. J Gerontol B Psychol Sci Soc Sci. 2008; 63(4):245-8. PM:18689767.

131. Lin P, Campbell DG, Chaney EF, Liu CF, Heagerty P, Felker BL, and Hedrick SC. The influence of patient preference on depression treatment in primary care. Ann Behav Med. 2005; 30(2):164-73. PM:16173913.

132. Raue PJ, Schulberg HC, Heo M, Klimstra S, and Bruce ML. Patients' depression treatment preferences and initiation, adherence, and outcome: a randomized primary care study. Psychiatr Serv. 2009; 60(3):337-43. PM:19252046.

133. Sharp DJ, Chew-Graham C, Tylee A, Lewis G, Howard L, Anderson I, Abel K, Turner KM, Hollinghurst SP, Tallon D, McCarthy A, and Peters TJ. A pragmatic randomised controlled trial to compare antidepressants with a community-based psychosocial intervention for the treatment of women with postnatal depression: the RESPOND trial. Health Technol Assess. 2010; 14(43):iii-xi,1. PM:20860888.

134. Kwan BM, Dimidjian S, and Rizvi SL. Treatment preference, engagement, and clinical improvement in pharmacotherapy versus psychotherapy for depression. Behav Res Ther. 2010; 48(8):799-804. PM:20462569.

135. Sleath B, West S, Tudor G, Perreira K, King V, and Morrissey J. Ethnicity and depression treatment preferences of pregnant women. J Psychosom Obstet Gynaecol. 2005; 26(2):135-40. PM:16050539.

136. Mergl R, Henkel V, Allgaier A-K, Kramer D, Hautzinger M, Kohnen R, Coyne J, and Hegerl U. Are treatment preferences relevant in response to serotonergic antidepressants and cognitive-behavioral therapy in depressed primary care patients? Results from a randomized controlled trial including a patients' choice arm. Psychother Psychosom. 2011; 80(1):39-47. PM:20975325.

137. Gum AM, Arean PA, Hunkeler E, Tang L, Katon W, Hitchcock P, Steffens DC, Dickens J, and Unutzer J. Depression treatment preferences in older primary care patients. Gerontologist. 2006; 46(1):14-22. PM:16452280.

138. Fann JR, Jones AL, Dikmen SS, Temkin NR, Esselman PC, and Bombardier CH. Depression treatment preferences after traumatic brain injury. J Head Trauma Rehabil. 2009; 24(4):272-8. PM:19625866.

139. Van Voorhees BW, Fogel J, Houston TK, Cooper LA, Wang NY, and Ford DE. Beliefs and attitudes associated with the intention to not accept the diagnosis of depression among young adults. Ann Fam Med. 2005; 3(1):38-46. PM:15671189.

140. Leykin Y, Derubeis RJ, Gallop R, Amsterdam JD, Shelton RC, and Hollon SD. The relation of patients' treatment preferences to outcome in a randomized clinical trial. Behav Ther. 2007; 38(3):209-17. PM:17697846.

141. Givens JL, Datto CJ, Ruckdeschel K, Knott K, Zubritsky C, Oslin DW, Nyshadham S, Vanguri P, and Barg FK. Older patients' aversion to antidepressants. a qualitative study. J Gen Intern Med. 2006; 21(2):146-51. PM:16336620.

142. Brown C, Battista DR, Bruehlman R, Sereika SS, Thase ME, and Dunbar-Jacob J. Beliefs about antidepressant medications in primary care patients: relationship to self-reported adherence. Med Care. 2005; 43(12):1203-7. PM:16299431.

143. Stecker T and Alvidrez J. Patient decision-making regarding entry into psychotherapy to treat depression. Issues Ment Health Nurs. 2007; 28(7):811-20. PM:17654113.

144. Gilbody S, Bower P, Fletcher J, Richards D, and Sutton AJ. Collaborative care for depression: a cumulative meta-analysis and review of longer-term outcomes. Arch Intern Med. 2006; 166(21):2314-21. PM:17130383.

145. Arroll B, Elley CR, Fishman T, Goodyear-Smith FA, Kenealy T, Blashki G, Kerse N, and MacGillivray S. Antidepressants versus placebo for depression in primary care. Cochrane Database Syst Rev. 2009; (3):CD007954. PM:19588448.

146. Cuijpers P, van Straten A, and Smit F. Psychological treatment of late-life depression: a meta-analysis of randomized controlled trials. Int J Geriatr Psychiatry. 2006; 21(12):1139-49. PM:16955421.

147. Gartlehner G, Hansen RN, Thieda P, DeVeaugh-Geiss A, Gaynes BN, Krebs EE, Morgan LC, Lux LJ, Shumate JA, Monroe LG, and Lohr KN. Comparative effectiveness of second-generation antidepressants in the pharmacologic treatment of adult depression. 07-EHC007-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2007. Available at: PM:20704050.

148. Baumeister H, Hutter N, and Bengel J. Psychological and pharmacological interventions for depression in patients with coronary artery disease. Cochrane Database Syst Rev. 2011; 9:CD008012. PM:21901717.

149. Dowlati Y, Herrmann N, Swardfager WL, Reim EK, and Lanctot KL. Efficacy and tolerability of antidepressants for treatment of depression in coronary artery disease: a meta-analysis. Can J Psychiatry. 2010; 55(2):91-9. PM:20181304.

150. Kang-Yi CD and Gellis ZD. A systematic review of community-based health interventions on depression for older adults with heart disease. Aging Ment Health. 2010; 14(1):1-19. PM:20155517.

151. Mazza M, Lotrionte M, Biondi-Zoccai G, Abbate A, Sheiban I, and Romagnoli E. Selective serotonin reuptake inhibitors provide significant lower re-hospitalization rates in patients recovering from acute coronary syndromes: evidence from a meta-analysis. J Psychopharmacol. 2010; 24(12):1785-92. PM:19965939.

152. O'Neil A, Sanderson K, Oldenburg B, and Taylor CB. Impact of depression treatment on mental and physical health-related quality of life of cardiac patients: a meta-analysis. J Cardpulm Rehabil Prev. 2011; 31(3):146-56. PM:21150634.

153. Whalley B, Rees K, Davies P, Bennett P, Ebrahim S, Liu Z, West R, Moxham T, Thompson DR, and Taylor RS. Psychological interventions for coronary heart disease. Cochrane Database Syst Rev. 2011; (8):CD002902. PM:21833943.

154. Rayner L, Price A, Evans A, Valsraj K, Higginson IJ, and Hotopf M. Antidepressants for depression in physically ill people. Cochrane Database Syst Rev. 2010; (3):CD007503. PM:20238354.

155. Rizzo M, Creed F, Goldberg D, Meader N, and Pilling S. A systematic review of non-pharmacological treatments for depression in people with chronic physical health problems. J Psychosom Res. 2011; 71(1):18-27. PM:21665008.

156. Taylor D, Meader N, Bird V, Pilling S, Creed F, Goldberg D, and pharmacology subgroup of the National Institute for Health and Clinical Excellence Guideline Development Group for Depression in Chronic Physical Health Problems. Pharmacological interventions for people with depression and chronic physical health problems: systematic review and meta-analyses of safety and efficacy. Br J Psychiatry. 2011; 198(3):179-88. PM:21357876.

157. Dai B, Li J, and Cuijpers P. Psychological treatment of depressive symptoms in Chinese elderly inpatients with significant medical comorbidity: a meta-analysis. BMC Psychiatr. 2011; 11(92) http://www.biomedcentral.com/content/pdf/1471-244x-11-92.pdf.

158. Iovieno N, Tedeschini E, Ameral VE, Rigatelli M, and Papakostas GI. Antidepressants for major depressive disorder in patients with a co-morbid axis-III disorder: a meta-analysis of patient characteristics and placebo response rates in randomized controlled trials. Int Clin Psychopharmacol. 2011; 26(2):69-74. PM:20962663.

159. Meijer A, Roseman M, Milette K, Coyne JC, Stefanek ME, Ziegelstein RC, Arthurs E, Leavens A, Palmer SC, Stewart DE, de JP, and Thombs BD. Depression screening and patient outcomes in cancer: a systematic review. PLoS ONE. 2011; 6(11):e27181. http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0027181.

160. van der Feltz-Cornelis CM, Nuyen J, Stoop C, Chan J, Jacobson AM, Katon W, Snoek F, and Sartorius N. Effect of interventions for major depressive disorder and significant depressive symptoms in patients with diabetes mellitus: a systematic review and meta-analysis. Gen Hosp Psychiatry. 2010; 32(4):380-95. PM:20633742.

161. van Straten A, Geraedts A, Verdonck-de Leeuw I, Andersson G, and Cuijpers P. Psychological treatment of depressive symptoms in patients with medical disorders: a meta-analysis. J Psychosom Res. 2010; 69(1):23-32. PM:20630260.

162. Beltman MW, Voshaar RC, and Speckens AE. Cognitive-behavioural therapy for depression in people with a somatic disease: Meta-analysis of randomised controlled trials. Br J Psychiatry. 2010; 197(1):11-9. PM:20592427.

163. Henschke N, Ostelo RW, van Tulder MW, Vlaeyen JW, Morley S, Assendelft WJ, and Main CJ. Behavioural treatment for chronic low-back pain. Cochrane Database Syst Rev. 2010; (7):CD002014. PM:20614428.

164. Urquhart DM, Hoving JL, Assendelft-Willem JJ, Roland M, and van-Tulder MW. Antidepressants for non-specific low back pain. Cochrane Database Syst Rev. 2008; (1):CD001703. http://ww.cfah.org/hbns/archives/viewSupportDoc.cfm?supportingDocID=538.

165. Pani PP, Vacca R, Trogu E, Amato L, and Davoli M. Pharmacological treatment for depression during opioid agonist treatment for opioid dependence. Cochrane Database Syst Rev. 2010; (9):CD008373. PM:20824876.

166. Pedrelli P, Iovieno N, Vitali M, Tedeschini E, Bentley KH, and Papakostas GI. Treatment of major depressive disorder and dysthymic disorder with antidepressants in patients with comorbid opiate use disorders enrolled in methadone maintenance therapy: a meta-analysis. J Clin Psychopharmacol. 2011; 31(5):582-6. PM:21869696.

167. Cuijpers P, Brannmark JG, and van Straten A. Psychological treatment of postpartum depression: a meta-analysis. J Clin Psychol. 2008; 64(1):103-18. PM:18161036.

168. Dennis CL and Hodnett E. Psychosocial and psychological interventions for treating postpartum depression. Cochrane Database Syst Rev. 2007; (4):CD006116. PM:17943888.

169. Williams JW, Jr., Mulrow CD, Kroenke K, Dhanda R, Badgett RG, Omori D, and Lee S. Case-finding for depression in primary care: a randomized trial. Am J Med. 1999; 106(1):36-43. PM:10320115.

170. National Collaborating Centre for Mental Health. The NICE guideline on the management and treatment of depression in adults (updated edition). National Clinical Practice Guideline 90. London (UK): National Institute for Health and Clinical Excellence (NICE); 2010. Available at: http://www.nice.org.uk/nicemedia/live/12329/45896/45896.pdf.

171. Scottish Intercollegiate Guidelines Network (SIGN). Management of perinatal mood disorders. SIGN publication no. 127. Edinburgh: SIGN; 2012. Available at: http://www.sign.ac.uk.

172. National Collaborating Centre for Mental Health. Antenatal and postnatal mental health: clinical management and service guidance. NICE Clinical Guideline 45. London, UK: National Institute for Health and Clinical Excellence (NICE); 2007. Available at: http://www.nice.org.uk/nicemedia/live/11004/30433/30433.pdf.

173. 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(1):19-24. PM:21483194.

174. Szigethy E, Levy-Warren A, Whitton S, Bousvaros A, Gauvreau K, Leichtner AM, and Beardslee WR. Depressive symptoms and inflammatory bowel disease in children and adolescents: a cross-sectional study. J Pediatr Gastroenterol Nutr. 2004; 39(4):395-403. PM:15448431.

Figure 2. Flow of Studies to Final Number of Eligible Studies in Stage 1

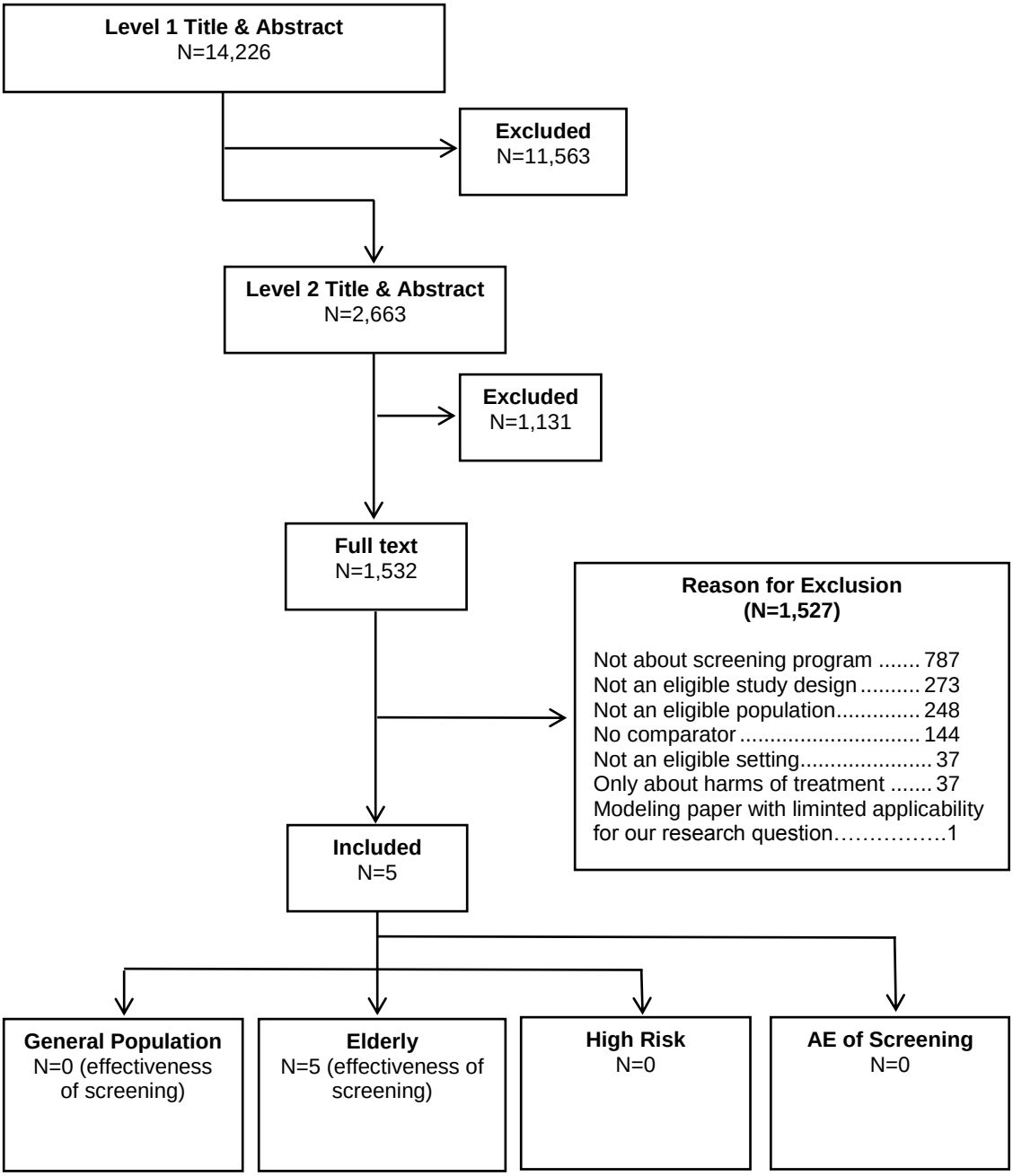
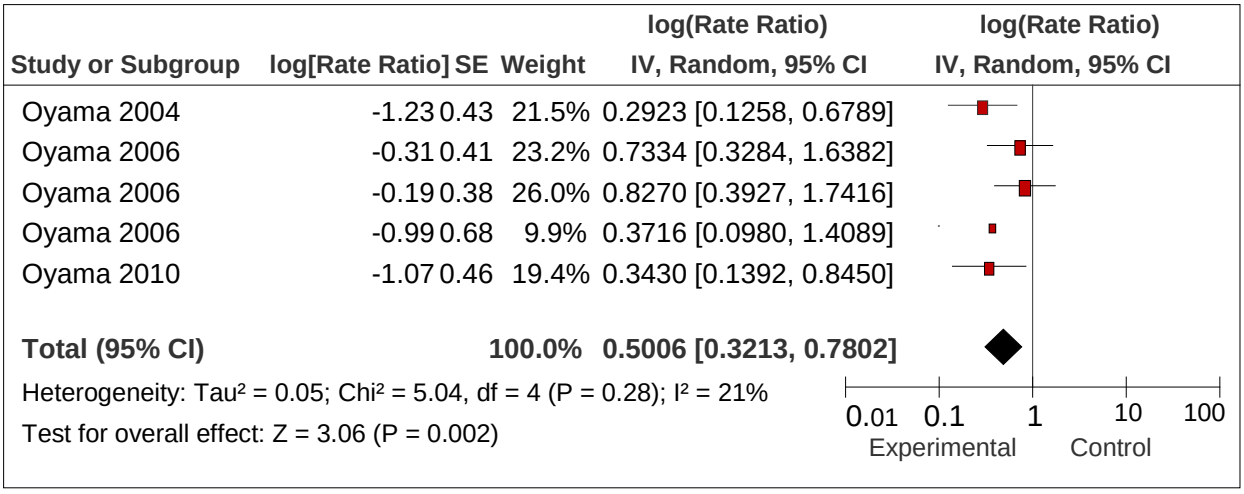


Figure 3. Forest Plot: Effect of Community-based Suicide Prevention Program (including screening for depression) Overall analysis



NOTE 1: the number of studies (n=5) is too few to assess publication bias confidently using a funnel plot (threshold rule of thumb value is >=10) – may not want to include the plot in the report when circulated for peer review
NOTE 2: The studies that were reviewed here evaluated the effectiveness of the community-based depression screening (CDS) programs which incorporated screening for depression, followup with mental health care or psychiatric treatment, and health education in the community setting. As such, the observed reduction in suicide rates cannot be attributed solely to the screening component of these programs.

Figure 4. Forest Plot: Effect of Community-based Suicide Prevention Program (including screening for depression) on Completed Suicide by Gender

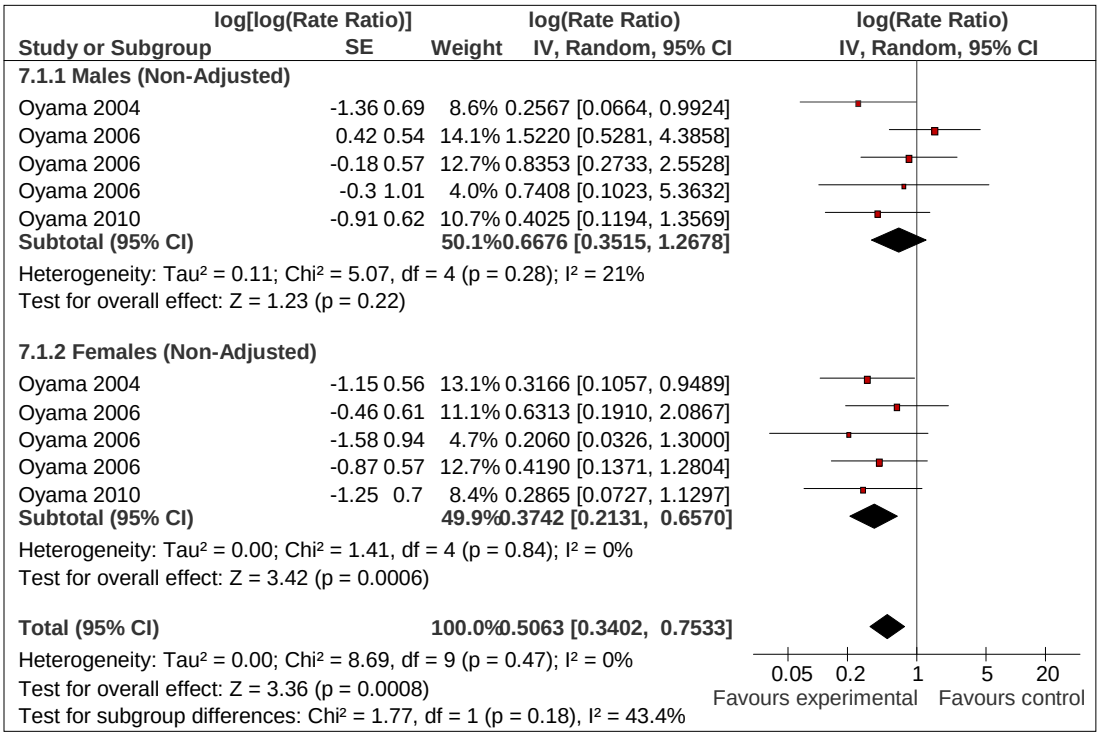


Figure 5. Forest Plot: Effect of Community-based Suicide Prevention Program (including screening for depression) By Age

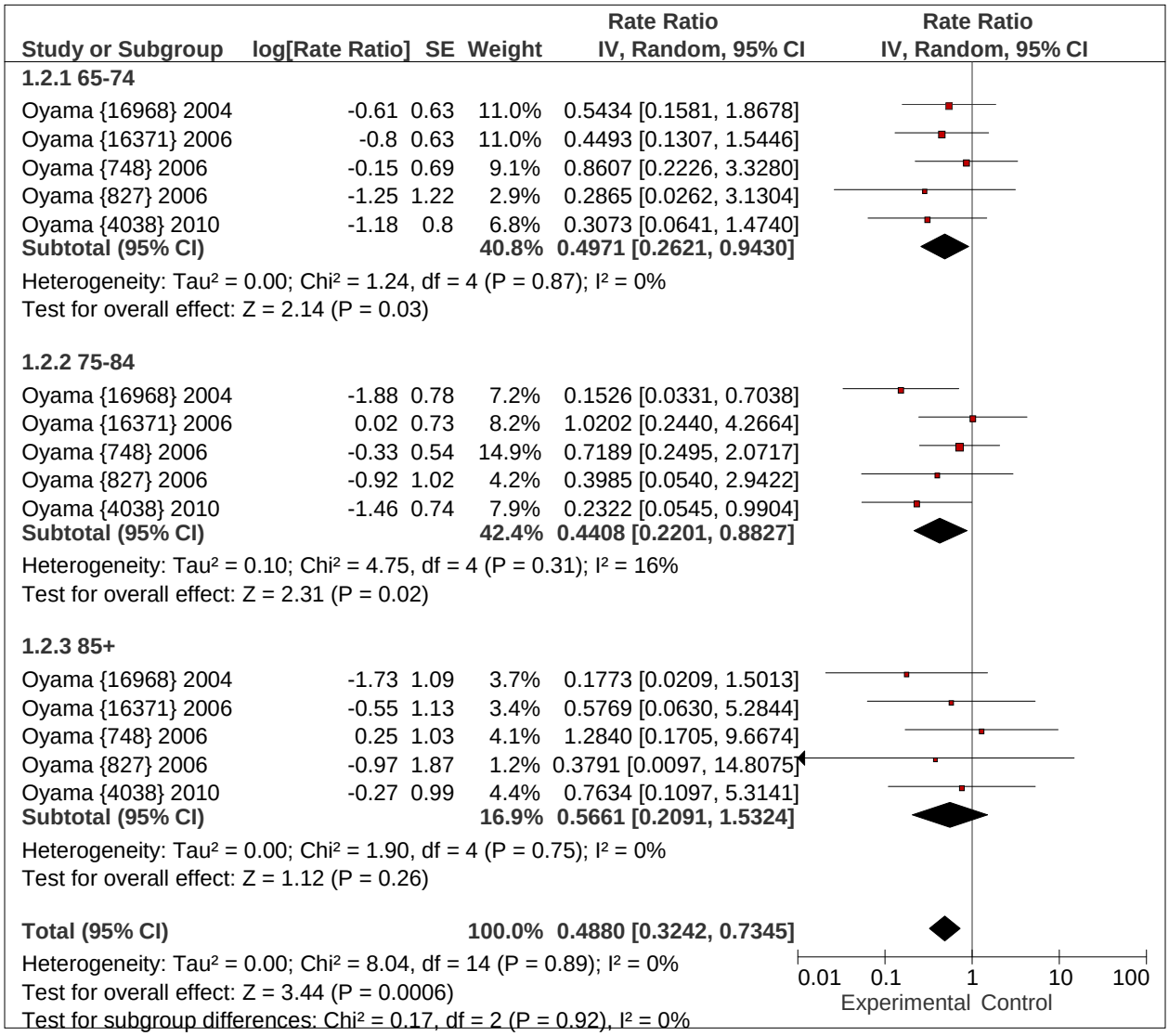
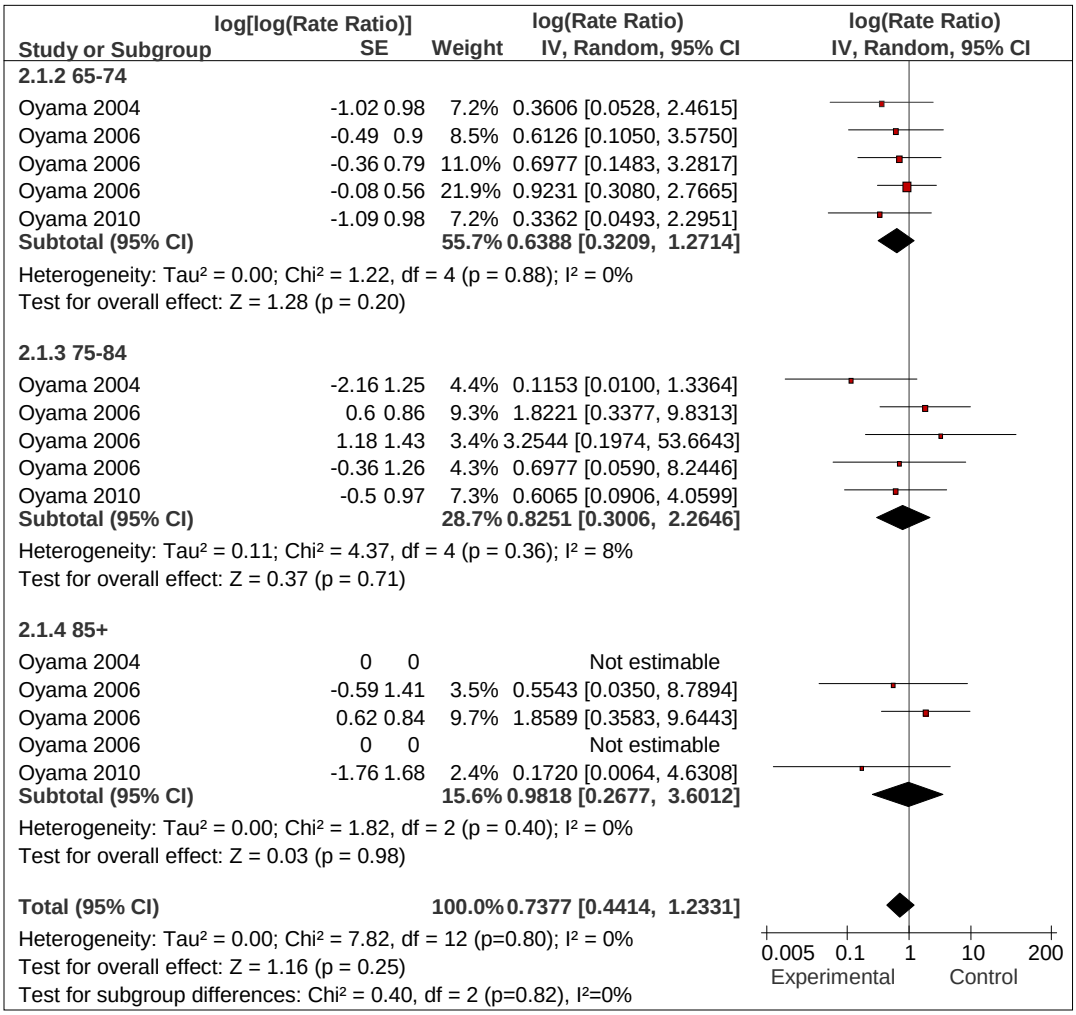
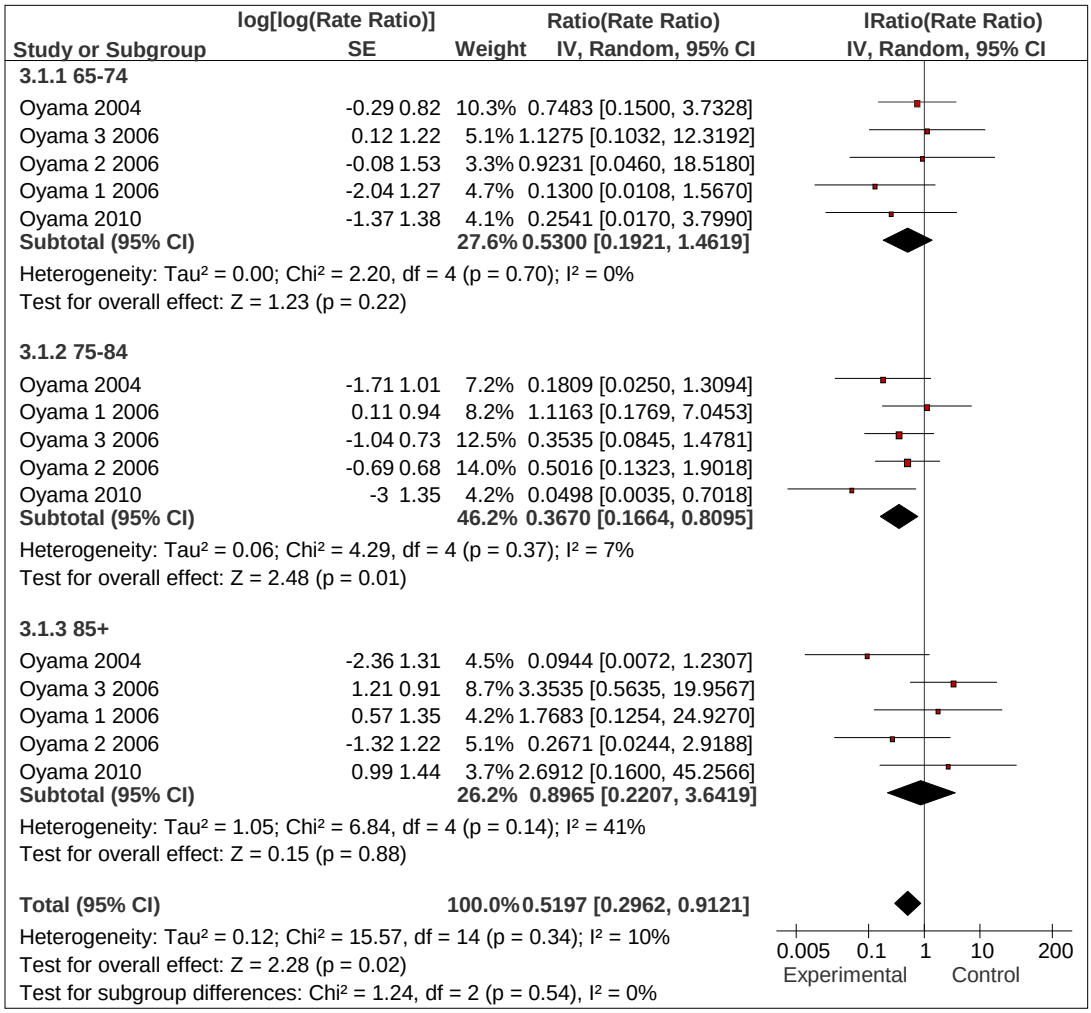


Figure 6. Forest Plot: Effect of Community-based Suicide Prevention Program (including screening for depression) by Age Group - Male



NOTE 1: the number of studies (n=5) is too few to assess publication bias confidently using a funnel plot (threshold rule of thumb value is >=10) – may not want to include the plot in the report when circulated for peer review
NOTE 2: The studies that were reviewed here evaluated the effectiveness of the community-based depression screening (CDS) programs which incorporated screening for depression, followup with mental health care or psychiatric treatment, and health education in the community setting. As such, the observed reduction in suicide rates cannot be attributed solely to the screening component of these programs.

Figure 7. Forest Plot: Effect of Community-based Suicide Prevention Program (including screening for depression) by Age Group - Female



NOTE 1: the number of studies (n=5) is too few to assess publication bias confidently using a funnel plot (threshold rule of thumb value is >=10) – may not want to include the plot in the report when circulated for peer review
NOTE 2: The studies that were reviewed here evaluated the effectiveness of the community-based depression screening (CDS) programs which incorporated screening for depression, followup with mental health care or psychiatric treatment, and health education in the community setting. As such, the observed reduction in suicide rates cannot be attributed solely to the screening component of these programs.

Table 4. Summary of Quality Assessment for Quasi-experimental papers

Study	A: Selection				B: Comparability		C: Outcome			Total Stars
	A1: Representative-ness of exposed cohort	A2: Selection of non-exposed cohort	A3: Ascertainment of exposure	A4: Outcome not present at outset	B1a: Comparability of cohorts: age	B1b: Comparability of cohorts on additional factor	C1: Assessment of outcome	C2: Length of follow-up	C3: Adequacy of cohort follow-up	
Oyama [87]	Truly representative ★	Different source	Written self-report	Yes★	Yes★	Yes★	Record linkage★	Yes★	Complete follow-up – all subjects accounted for★	7
Oyama [88]	Truly representative ★	Different source	Written self-report	Yes★	Yes★	Yes★	Record linkage★	Yes★	Complete follow-up – all subjects accounted for★	7
Oyama [85]	Truly representative ★	Different source	Written self-report	Yes★	Yes★	Yes★	Record linkage★	Yes★	Complete follow-up – all subjects accounted for★	7
Oyama [89]	Truly representative ★	Different source	Written self-report	Yes★	Yes★	Yes★	Record linkage★	Yes★	Complete follow-up – all subjects accounted for★	7
Oyama [174]	Truly representative ★	Different source	Written self-report	Yes★	Yes★	Yes★	Record linkage★	Yes★	Complete follow-up – all subjects accounted for★	7

Assessed using Newcastle-Ottawa Scale [96]

Table 5. Summary of Findings KQ1 – Effect of Community-based Suicide Prevention Program (including screening for depression) - Incidence of Suicide (overall)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CDS screening - Elderly	Control	Relative (95% CI)	Absolute		
Overall (follow-up 5-10 years; assessed with: CDS (community depression screening))												
5[1]	observational studies	no serious risk of bias[2]	no serious inconsistency[3]	very serious[4,5]	no serious imprecision[6]	none[7]	65/70053 (0.09%)	145/113324 (0.13%)	RR 0.5006 (0.3213 to 0.7802)	1 fewer per 1000 (from 0 fewer to 1 fewer)	⊕OOO VERY LOW	CRITICAL

[1] Oyama, Koida et al 2004; Oyama Fujita et al 2006; Oyama, Goto et al 2006; Oyama, Ono et al 2006; Oyama, Sakashita et al 2010
[2] The quality assessment tools identified a few issues with the papers (e.g. selection of non-exposed cohort, blinding and reporting of withdrawals/drop-outs) however the evidence was not downgraded for these reasons
[3] Heterogeneity statistics not significant Heterogeneity: $Tau^2 = 0.05$; $Chi^2 = 5.04$, df = 4 (P = 0.28); $I^2 = 21\%$
[4] Directness was downgraded due to concerns regarding population characteristics. The included papers all looked at elderly, rural Japanese populations which are unlikely to be representative of Canadians at average or high risk for depression
[5] Directness was downgraded for the second time due to concerns regarding the Community Depression Screening: The studies that were reviewed here evaluated the effectiveness of the community-based depression screening (CDS) programs which incorporated screening for depression, followup with mental health care or psychiatric treatment, and health education in the community setting. As such, the observed reduction in suicide rates cannot be attributed solely to the screening component of these programs.
[6] The number of events is small (<300, a threshold rule of thumb value for dichotomous outcomes), however considering the specific outcome the evidence is not downgraded.
[7] Funnel plot of comparison indicates potential asymmetry and thus potential publication bias. However, the number of papers (n=5) was too small to assess this with confidence (>=10 papers is the threshold rule of thumb value).

Table 6. Summary of Findings KQ1 – Effect of Community-based Suicide Prevention Program (including screening for depression) - Incidence of Suicide (by age groups)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CDS screening - Elderly by age groups	Control	Relative (95% CI)	Absolute		
Overall - By Age Groups - 65-74 (follow-up 5-10 years; assessed with: CDS (community depression screening))												
5[1]	observational studies	no serious risk of bias[2]	no serious inconsistency[3]	very serious[4,5]	no serious imprecision[6]	none[7]	26/35843 (0.07%)	63/61086 (0.1%)	RR 0.4971 (0.2621 to 0.9430)	1 fewer per 1000 (from 0 fewer to 1 fewer)	⊕OOO VERY LOW	CRITICAL
Overall - By Age Groups - 75-84 (follow-up 5-10 years; assessed with: CDS (community depression screening))												
5[1]	observational studies	no serious risk of bias[2]	no serious inconsistency[8]	very serious[4,5]	no serious imprecision[6]	none[7]	24/24441 (0.1%)	59/38644 (0.15%)	RR 0.4408 (0.2201 to 0.8827)	1 fewer per 1000 (from 0 fewer to 1 fewer)	⊕OOO VERY LOW	CRITICAL
Overall - By Age Groups - 85+ (follow-up 5-10 years; assessed with: CDS (community depression screening))												
5[1]	observational studies	no serious risk of bias[2]	no serious inconsistency[9]	very serious[4,5]	no serious imprecision[6]	none[7]	15/9769 (0.15%)	23/13594 (0.17%)	RR 0.5661 (0.2091 to 1.5324)	1 fewer per 1000 (from 1 fewer to 1 more)	⊕OOO VERY LOW	CRITICAL

[1] Oyama, Koida et al 2004; Oyama Fujita et al 2006; Oyama, Goto et al 2006; Oyama, Ono et al 2006; Oyama, Sakashita et al 2010
[2] The quality assessment tools identified a few issues with the papers (e.g. selection of non-exposed cohort, blinding and reporting of withdrawals/drop-outs) however the evidence was not downgraded for these reasons
[3] Heterogeneity statistics not significant Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 1.24$, df = 4 (P = 0.87); $I^2 = 0\%$
[4] Directness was downgraded due to concerns regarding population characteristics. The included papers all looked at elderly, rural Japanese populations which are unlikely to be representative of Canadians at average or high risk for depression
[5] Directness was downgraded for the second time due to concerns regarding the Community Depression Screening: The studies that were reviewed here evaluated the effectiveness of the community-based depression screening (CDS) programs which incorporated screening for depression, follow-up with mental health care or psychiatric treatment, and health education in the community setting. As such, the observed reduction in suicide rates cannot be attributed solely to the screening component of these programs.
[6] The number of events is small (<300, a threshold rule of thumb value for dichotomous outcomes), however considering the specific outcome the evidence is not downgraded.
[7] Funnel plot of comparison indicates potential asymmetry and thus potential publication bias. However, the number of papers (n=5) was too small to assess this with confidence (>=10 papers is the threshold rule of thumb value).

[8] Heterogeneity statistics not significant Heterogeneity: $Tau^2 = 0.11$; $Chi^2 = 4.75$, df = 4 (P = 0.31); $I^2 = 16\%$
[9] Heterogeneity statistics not significant Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 1.90$, df = 4 (P = 0.75); $I^2 = 0\%$

Table 7. Summary of Findings KQ1a Effect of Community-based Suicide Prevention Program (including screening for depression) - Incidence of Suicide by Gender

Quality assessment							No. of patients		Effect	Quality	Importance
No. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CDS	Control	Ratio of Rate Ratios (RRR) (95% CI)		
Total Males (followup 5-10 patient-years; assessed with: CDS (community depression screening))											
5[a]	observational studies	no serious risk of bias[b]	no serious inconsistency[c]	very serious[d]	no serious imprecision[e]	none[f]	37/29182 [g]	66/46905[h]	0.6676 (0.3515 to 1.2678)	⊕OOO VERY LOW	CRITICAL
Total Females (followup 5-10 patient-years; assessed with: CDS)											
5	observational studies	no serious risk of bias[b]	no serious inconsistency[i]	very serious[d]	no serious imprecision[e]	none[f]	28/40871 [g]	79/66419[h]	0.3742 (0.2131 to 0.6570)	⊕OOO VERY LOW	CRITICAL

[a] Oyama, Koida et al 2004; Oyama Fujita et al 2006; Oyama, Goto et al 2006; Oyama, Ono et al 2006; Oyama, Sakashita et al 2010

[b] The quality assessment tools identified a few issues with the papers (e.g. selection of non-exposed cohort, blinding and reporting of withdrawals/drop-outs) however the evidence was not downgraded for these reasons

[c] Heterogeneity statistics were not significant Heterogeneity: $Tau^2 = 0.11$; $Chi^2 = 5.08$, df = 4 (p = 0.28); $I^2 = 21\%$

[d] Directness was downgraded due to concerns regarding population characteristics. The included papers all looked at elderly, rural Japanese populations which are unlikely to be representative of Canadians at average or high risk for depression.
Directness was downgraded for the second time due to concerns regarding the Community Depression Screening: The studies that were reviewed here evaluated the effectiveness of the community-based depression screening (CDS) programs which incorporated education on suicide prevention, screening for depression, and , if needed, followup with mental health care or psychiatric treatment. The education component occurred first and provided information on suicide prevention, screening for depression followed. As such, the observed reduction in suicide rates cannot be attributed solely to the screening component of these programs.

[e] The number of events is small (<300, a threshold rule of thumb value for dichotomous outcomes), however considering the specific outcome the evidence is not downgraded.

[f] Funnel plot of comparison indicates potential asymmetry and thus potential publication bias. However, the number of papers (n=5) was too small to assess this with confidence (>=10 papers is the threshold rule of thumb value).

[g] Intervention data is based on post implementation

[h] Control data is based on post implementation.

[i] Heterogeneity statistics were not significant Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 1.41$, df = 4 (p = 0.84); $I^2 = 0\%$

Table 8. Summary of Findings KQ1 – Effect of Community-based Suicide Prevention Program (including screening for depression) - Incidence of Suicide in Males (by age)

Quality assessment							No of patients		Effect	Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CDS (Males)	Control	Ratio of Rate Ratios (RRR) (95% CI)		
Males - 65-74 (followup 5-10 years; assessed with: CDS (community depression scale))											
5[a]	observational studies	no serious risk of bias[b]	no serious inconsistency[c]	very serious[d]	no serious imprecision[e]	none[f]	18/16280	36/27400[h]	0.6388 (0.3209 to 1.2714)	⊕OOO VERY LOW	CRITICAL
Males - 75-84 (followup 5-10 years; assessed with: CDS (community depression scale))											
5[a]	observational studies	no serious risk of bias[b]	no serious inconsistency[i]	very serious[d]	no serious imprecision[e]	none[f]	11/9864 [g]	23/15249[h]	0.8251(0.3006 to 2.2646)	⊕OOO VERY LOW	CRITICAL
Males - 85+ (followup 5-10 years; assessed with: CDS (community depression scale))											
5[a]	observational studies	no serious risk of bias[b]	no serious inconsistency[j]	very serious[d]	no serious imprecision[e]	none[f]	8/3038 [g]	7/4256 [h]	0.9818 (0.2677 to 3.6012)	⊕OOO VERY LOW	CRITICAL

[a] Oyama, Koida et al 2004; Oyama Fujita et al 2006; Oyama, Goto et al 2006; Oyama, Ono et al 2006; Oyama, Sakashita et al 2010

[b] The quality assessment tools identified a few issues with the papers (e.g. selection of non-exposed cohort, blinding and reporting of withdrawals/drop-outs) however the evidence was not downgraded for these reasons

[c] Heterogeneity statistics not significant Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 1.22$, df = 4 (p = 0.88); $I^2 = 0\%$

[d] Directness was downgraded due to concerns regarding population characteristics. The included papers all looked at elderly, rural Japanese populations which are unlikely to be representative of Canadians at average or high risk for depression.
Directness was downgraded for the second time due to concerns regarding the Community Depression Screening: The studies that were reviewed here evaluated the effectiveness of the community-based depression screening (CDS) programs which incorporated education on suicide prevention, screening for depression, and , if needed, followup with mental health care or psychiatric treatment. The education component occurred first and provided information on suicide prevention, screening for depression followed. As such, the observed reduction in suicide rates cannot be attributed solely to the screening component of these programs.

[e] The number of events is small (<300, a threshold rule of thumb value for dichotomous outcomes), however considering the specific outcome the evidence is not downgraded.

[f] Funnel plot of comparison indicates potential asymmetry and thus potential publication bias. However, the number of papers (n=5) was too small to assess this with confidence (>=10 papers is the threshold rule of thumb value).

[g] Intervention data is based on post implementation.

[h] Control data is based on post implementation data.

[i] Heterogeneity statistic not significant Heterogeneity: $Tau^2 = 0.11$; $Chi^2 = 4.36$, df = 4 (p = 0.36); $I^2 = 8\%$

[j] Heterogeneity statistic is not significant Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 1.83$, df = 2 (p = 0.40); $I^2 = 0\%$

Table 9. Summary of Findings KQ1 – Effect of Community-based Suicide Prevention Program (including screening for depression) - Incidence of Suicide in Females (by age)

Quality assessment							No of patients		Effect	Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	CDS/screening (Females)	Control	Ratio of Rate Ratios (RRR) (95% CI)		
Females - 65-74 (followup 5-10 years; assessed with: CDS (community depression scale))											
5[a]	observational studies	no serious risk of bias[b]	no serious inconsistency[c]	very serious[d]	no serious imprecision[e]	none[f]	8/19563 [g]	27/33686[h]	0.5300 (0.1921 to 1.4619)	⊕OOO VERY LOW	CRITICAL
Females - 75-84 (followup 5-10 years; assessed with: CDS (community depression scale))											
5[a]	observational studies	no serious risk of bias[b]	no serious inconsistency[i]	very serious[d]	no serious imprecision[e]	none[f]	13/14577 [g]	36/23395[h]	0.3670 (0.1664 to 0.8095)	⊕OOO VERY LOW	CRITICAL
Females - 85+ (followup 5-10 years; assessed with: CDS (community depression scale))											
5[a]	observational studies	no serious risk of bias[b]	no serious inconsistency[j]	very serious[d]	no serious imprecision[e]	none[f]	7/6731[g]	16/9338 [h]	0.8965 (0.2207 to 3.6419)	⊕OOO VERY LOW	CRITICAL

[a] Oyama, Koida et al 2004; Oyama Fujita et al 2006; Oyama, Goto et al 2006; Oyama, Ono et al 2006; Oyama, Sakashita et al 2010

[b] The quality assessment tools identified a few issues with the papers (e.g. selection of non-exposed cohort, blinding and reporting of withdrawals/drop-outs) however the evidence was not downgraded for these reasons

[c] Heterogeneity statistic not significant Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 2.20$, df = 4 (p = 0.70); $I^2 = 0\%$

[d] Directness was downgraded due to concerns regarding population characteristics. The included papers all looked at elderly, rural Japanese populations which are unlikely to be representative of Canadians at average or high risk for depression.
Directness was downgraded for the second time due to concerns regarding the Community Depression Screening: The studies that were reviewed here evaluated the effectiveness of the community-based depression screening (CDS) programs which incorporated education on suicide prevention, screening for depression, and , if needed, followup with mental health care or psychiatric treatment. The education component occurred first and provided information on suicide prevention, screening for depression followed. As such, the observed reduction in suicide rates cannot be attributed solely to the screening component of these programs.

[e] The number of events is small (<300, a threshold rule of thumb value for dichotomous outcomes), however considering the specific outcome the evidence is not downgraded.

[f] Funnel plot of comparison indicates potential asymmetry and thus potential publication bias. However, the number of papers (n=5) was too small to assess this with confidence (>=10 papers is the threshold rule of thumb value).

[g] Intervention data is based on post implementation.

[h] Control data is based on post implementation data.

[i] Heterogeneity statistic is not significant Heterogeneity: $Tau^2 = 0.06$; $Chi^2 = 4.29$, df = 4 (p = 0.37); $I^2 = 7\%$

[j] Heterogeneity statistic is not significant Heterogeneity: $Tau^2 = 1.05$; $Chi^2 = 6.84$, df = 4 (p = 0.14); $I^2 = 41\%$

Table 10. Summary of Findings KQ1 – Effect of Community-based Suicide Prevention Program (including screening for depression) - Incidence of Suicide (overall)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with CDS screening - Elderly (95% CI)
Overall CDS (community depression screening)	183377 (5 studies[1]) 5-10 years	⊕⊖⊖⊖ **VERY LOW**[2,3,4,5,6,7] due to indirectness	**RR 0.5006** (0.3213 to 0.7802)	**1 per 1000**	**1 fewer per 1000** (from 0 fewer to 1 fewer)

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; **RR:** Risk ratio;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Oyama, Koida et al 2004; Oyama Fujita et al 2006; Oyama, Goto et al 2006; Oyama, Ono et al 2006; Oyama, Sakashita et al 2010
[2] The quality assessment tools identified a few issues with the papers (e.g. selection of non-exposed cohort, blinding and reporting of withdrawals/drop-outs) however the evidence was not downgraded for these reasons
[3] Heterogeneity statistics not significant Heterogeneity: $Tau^2 = 0.05$; $Chi^2 = 5.04$, df = 4 (P = 0.28); $I^2 = 21\%$
[4] Directness was downgraded due to concerns regarding population characteristics. The included papers all looked at elderly, rural Japanese populations which are unlikely to be representative of Canadians at average or high risk for depression
[5] Directness was downgraded for the second time due to concerns regarding the Community Depression Screening: The studies that were reviewed here evaluated the effectiveness of the community-based depression screening (CDS) programs which incorporated screening for depression, follow-up with mental health care or psychiatric treatment, and health education in the community setting. As such, the observed reduction in suicide rates cannot be attributed solely to the screening component of these programs.
[6] The number of events is small (<300, a threshold rule of thumb value for dichotomous outcomes), however considering the specific outcome the evidence is not downgraded.
[7] Funnel plot of comparison indicates potential asymmetry and thus potential publication bias. However, the number of papers (n=5) was too small to assess this with confidence (>=10 papers is the threshold rule of thumb value).

Table 11. Summary of Findings KQ1 – Effect of Community-based Suicide Prevention Program (including screening for depression) - Incidence of Suicide (by age)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with CDS screening - Elderly by age groups (95% CI)
Overall - By Age Groups - 65-74 CDS (community depression screening)	96929 (5 studies[1]) 5-10 years	⨁⊖⊖⊖ **VERY LOW**[2,3,4,5,6,7] due to indirectness	**RR 0.4971** (0.2621 to 0.9430)	**1 per 1000**	**1 fewer per 1000** (from 0 fewer to 1 fewer)
Overall - By Age Groups - 75-84 CDS (community depression screening)	63085 (5 studies[1]) 5-10 years	⨁⊖⊖⊖ **VERY LOW**[2,4,5,6,7,8] due to indirectness	**RR 0.4408** (0.2201 to 0.8827)	**2 per 1000**	**1 fewer per 1000** (from 0 fewer to 1 fewer)
Overall - By Age Groups - 85+ CDS (community depression screening)	23363 (5 studies[1]) 5-10 years	⨁⊖⊖⊖ **VERY LOW**[2,4,5,6,7,9] due to indirectness	**RR 0.5661** (0.2091 to 1.5324)	**2 per 1000**	**1 fewer per 1000** (from 1 fewer to 1 more)

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval; **RR:** Risk ratio;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Oyama, Koida et al 2004; Oyama Fujita et al 2006; Oyama, Goto et al 2006; Oyama, Ono et al 2006; Oyama, Sakashita et al 2010
[2] The quality assessment tools identified a few issues with the papers (e.g. selection of non-exposed cohort, blinding and reporting of withdrawals/drop-outs) however the evidence was not downgraded for these reasons
[3] Heterogeneity statistics not significant Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 1.24$, df = 4 (P = 0.87); $I^2 = 0\%$
[4] Directness was downgraded due to concerns regarding population characteristics. The included papers all looked at elderly, rural Japanese populations which are unlikely to be representative of Canadians at average or high risk for depression
[5] Directness was downgraded for the second time due to concerns regarding the Community Depression Screening: The studies that were reviewed here evaluated the effectiveness of the community-based depression screening (CDS) programs which incorporated screening for depression, follow-up with mental health care or psychiatric treatment, and health education in the community setting. As such, the observed reduction in suicide rates cannot be attributed solely to the screening component of these programs.
[6] The number of events is small (<300, a threshold rule of thumb value for dichotomous outcomes), however considering the specific outcome the evidence is not downgraded.
[7] Funnel plot of comparison indicates potential asymmetry and thus potential publication bias. However, the number of papers (n=5) was too small to assess this with confidence (>=10 papers is the threshold rule of thumb value).

[8] Heterogeneity statistics not significant Heterogeneity: $Tau^2 = 0.11$; $Chi^2 = 4.75$, df = 4 (P = 0.31); $I^2 = 16\%$
[9] Heterogeneity statistics not significant Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 1.90$, df = 4 (P = 0.75); $I^2 = 0\%$

Table 12 Summary of Findings: Effect of Community-based Suicide Prevention Program (including screening for depression) by gender, Oyama[85-89]

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	**CDS/screening**				
Total Males and Females (Non-Adjusted) - Males (Non-Adjusted)Table 4 CDS (community depression scale) Follow-up: 5-10 patient-years	**1 per 1000**[1]	**1 per 1000** (0 to 2)[2]	**0.6676** (0.3515 to 1.2678)	76087 (5 studies[3])	⊕⊖⊖⊖ **very low**[4,5,6,7,8,9]	
Total Males and Females (Non-Adjusted) - Females (Non-Adjusted) (table 4) CDS Follow-up: 5-10 patient-years	**1 per 1000**[1]	**0 per 1000** (0 to 1)[2]	**0.3742** (0.2131 to 0.6570)	107290 (5 studies)	⊕⊖⊖⊖ **very low**[4,6,7,8,9,10]	

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Control data is based on post implementation data.
[2] Intervention data is based on post implementation.
[3] Oyama, Koida et al 2004; Oyama Fujita et al 2006; Oyama, Goto et al 2006; Oyama, Ono et al 2006; Oyama, Sakashita et al 2010
[4] The quality assessment tools identified a few issues with the papers (e.g. selection of non-exposed cohort, blinding and reporting of withdrawals/drop-outs) however the evidence was not downgraded for these reasons
[5] Heterogeneity statistics were not significant Heterogeneity: TauÂ² = 0.11; ChiÂ² = 5.08, df = 4 (p = 0.28); IÂ² = 21%
[6] Directness was downgraded due to concerns regarding population characteristics. The included papers all looked at elderly, rural Japanese populations which are unlikely to be representative of Canadians at average or high risk for depression.The studies that were reviewed here evaluated the effectiveness of the community-based depression screening (CDS) programs which incorporated screening for depression, followup with mental health care or psychiatric treatment, and health education in the community setting. As such, the observed reduction in suicide rates cannot be attributed solely to the screening component of these programs.
[7] Directness was downgraded for the second time due to concerns regarding the Community Depression Screening: The studies that were reviewed here evaluated the effectiveness of the community-based depression screening (CDS) programs which incorporated education on suicide prevention, screening for depression, and , if needed, followup with mental health care or psychiatric treatment. The education component occurred first and provided information on suicide prevention, screening for depression followed. As such, the observed reduction in suicide rates cannot be attributed solely to the screening component of these programs.
[8] The number of events is small (<300, a threshold rule of thumb value for dichotomous outcomes), however considering the specific outcome the evidence is not downgraded.
[9] Funnel plot of comparison indicates potential asymmetry and thus potential publication bias. However, the number of papers (n=5) was too small to assess this with confidence (>=10 papers is the threshold rule of thumb value).
[10] Heterogeneity statistics were not significant Heterogeneity: TauÂ² = 0.00; ChiÂ² = 1.41, df = 4 (p = 0.84); IÂ² = 0%

Table 13. Summary of Findings Effect of Community-based Suicide Prevention Program (including screening for depression) male and age, Oyama[85-89]

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk **Control**	Corresponding risk **CDS/screening (Males)**				
Males - 65-74 (table 5) CDS (community depression scale) Follow-up: 5-10 years	**1 per 1000**[1]	**1 per 1000** (0 to 2)[2]	**0.6388** (0.3209 to 1.2714)	43680 (5 studies[3])	⊕⊖⊖⊖ **very low**[4,5,6,7,8,9]	
Males - 75-84 (table 5) CDS (community depression scale) Follow-up: 5-10 years	**2 per 1000**[1]	**1 per 1000** (0 to 3)[2]	**0.8251** (0.3006 to 2.2646)	25113 (5 studies[3])	⊕⊖⊖⊖ **very low**[4,6,7,8,9,10]	
Males - 85+(table 5) CDS (community depression scale) Follow-up: 5-10 years	**2 per 1000**[1]	**2 per 1000** (0 to 6)[2]	**0.9818** (0.2677 to 3.6012)	7294 (5 studies[3])	⊕⊖⊖⊖ **very low**[4,6,7,8,9,11]	

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Control data is based on post implementation data.
[2] Intervention data is based on post implementation.
[3] Oyama, Koida et al 2004; Oyama Fujita et al 2006; Oyama, Goto et al 2006; Oyama, Ono et al 2006; Oyama, Sakashita et al 2010
[4] The quality assessment tools identified a few issues with the papers (e.g. selection of non-exposed cohort, blinding and reporting of withdrawals/drop-outs) however the evidence was not downgraded for these reasons
[5] Heterogeneity statistics not significant Heterogeneity: TauÂ² = 0.00; ChiÂ² = 1.22, df = 4 (p = 0.88); IÂ² = 0%
[6] Directness was downgraded due to concerns regarding population characteristics. The included papers all looked at elderly, rural Japanese populations which are unlikely to be representative of Canadians at average or high risk for depression. The studies that were reviewed here evaluated the effectiveness of the community-based depression screening (CDS) programs which incorporated screening for depression, followup with mental health care or psychiatric treatment, and health education in the community setting. As such, the observed reduction in suicide rates cannot be attributed solely to the screening component of these programs.
[7] Directness was downgraded for the second time due to concerns regarding the Community Depression Screening: The studies that were reviewed here evaluated the effectiveness of the community-based depression screening (CDS) programs which incorporated education on suicide prevention, screening for depression, and , if needed, followup with mental health care or psychiatric treatment. The education component occurred first and provided information on suicide prevention, screening for depression followed. As such, the observed reduction in suicide rates cannot be attributed solely to the screening component of these programs.
[8] The number of events is small (<300, a threshold rule of thumb value for dichotomous outcomes), however considering the specific outcome the evidence is not downgraded.
[9] Funnel plot of comparison indicates potential asymmetry and thus potential publication bias. However, the number of papers (n=5) was too small to assess this with confidence (>=10 papers is the threshold rule of thumb value).
[10] Heterogeneity statistic not significant Heterogeneity: TauÂ² = 0.11; ChiÂ² = 4.36, df = 4 (p = 0.36); IÂ² = 8%
[11] Heterogeneity statistic is not significant Heterogeneity: TauÂ² = 0.00; ChiÂ² = 1.83, df = 2 (p = 0.40); IÂ² = 0%

Table 14. Summary of Findings: Effect of Community-based Suicide Prevention Program (including screening for depression) female and age group, Oyama[85-89]

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk **Control**	Corresponding risk **CDS/screening (Females)**				
Females - 65-74 (table 6) CDS (community depression scale) Follow-up: 5-10 years	**1 per 1000**[1]	**0 per 1000** (0 to 1)[2]	**0.5300** (0.1921 to 1.4619)	53249 (5 studies[3])	⊕⊖⊖⊖ **very low**[4,5,6,7,8,9]	
Females - 75-84 (table 6) CDS (community depression scale) Follow-up: 5-10 years	**2 per 1000**[1]	**1 per 1000** (0 to 1)[2]	**0.3670** (0.1664 to 0.8095)	37972 (5 studies[3])	⊕⊖⊖⊖ **very low**[4,6,7,8,9,10]	
Females - 85+ (table 6) CDS (community depression scale) Follow-up: 5-10 years	**2 per 1000**[1]	**2 per 1000** (0 to 6)[2]	**0.8965** (0.2207 to 3.6419)	16069 (5 studies[3])	⊕⊖⊖⊖ **very low**[4,6,7,8,9,11]	

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: Confidence interval;

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1] Control data is based on post implementation data.
[2] Intervention data is based on post implementation.
[3] Oyama, Koida et al 2004; Oyama Fujita et al 2006; Oyama, Goto et al 2006; Oyama, Ono et al 2006; Oyama, Sakashita et al 2010
[4] The quality assessment tools identified a few issues with the papers (e.g. selection of non-exposed cohort, blinding and reporting of withdrawals/drop-outs) however the evidence was not downgraded for these reasons
[5] Heterogeneity statistic not significant Heterogeneity: TauÂ² = 0.00; ChiÂ² = 2.20, df = 4 (p = 0.70); IÂ² = 0%
[6] Directness was downgraded due to concerns regarding population characteristics. The included papers all looked at elderly, rural Japanese populations which are unlikely to be representative of Canadians at average or high risk for depression. The studies that were reviewed here evaluated the effectiveness of the community-based depression screening (CDS) programs which incorporated screening for depression, followup with mental health care or psychiatric treatment, and health education in the community setting. As such, the observed reduction in suicide rates cannot be attributed solely to the screening component of these programs.
[7] Directness was downgraded for the second time due to concerns regarding the Community Depression Screening: The studies that were reviewed here evaluated the effectiveness of the community-based depression screening (CDS) programs which incorporated education on suicide prevention, screening for depression, and , if needed, followup with mental health care or psychiatric treatment. The education component occurred first and provided information on suicide prevention, screening for depression followed. As such, the observed reduction in suicide rates cannot be attributed solely to the screening component of these programs.
[8] The number of events is small (<300, a threshold rule of thumb value for dichotomous outcomes), however considering the specific outcome the evidence is not downgraded.
[9] Funnel plot of comparison indicates potential asymmetry and thus potential publication bias. However, the number of papers (n=5) was too small to assess this with confidence (>=10 papers is the threshold rule of thumb value).
[10] Heterogeneity statistic is not significant Heterogeneity: TauÂ² = 0.06; ChiÂ² = 4.29, df = 4 (p = 0.37); IÂ² = 7%
[11] Heterogeneity statistic is not significant Heterogeneity: TauÂ² = 1.05; ChiÂ² = 6.84, df = 4 (p = 0.14); IÂ² = 41%

Table 15. Summary of effect size of included systematic reviews in high risk group: cardiovascular n=6

Study Detail	Intervention versus Placebo
	Outcome: Effect size
Baumeister[148] 2011	**Pharmacological versus Placebo** All-cause mortality: OR: 0.39 (0.07,2.02) n.s. Hospitalization: Non-specified OR: 0.58 (0.39, 0.85) sig p=0.0060 Hospitalization: ER visits OR: 0.58 (0.34, 1.00) n.s Depression score-short term: SMD: -0.24 (-0.38, -0.09) sig Depression remission - short term: OR: 1.80 (1.18, 2.74) sig p=0.0066 **Psychological with usual care** Depression scores – short term: SMD -0.36 (1.27, 0.54) n.s.
Dowlati[149] 2010	**Pharmacological versus Placebo** Response: OR: 1.72 (1.17, 2.54) sig p=0.006 Remission: OR: 1.80 (1.18, 2.74) sig p=0.007 Response (HDRS): MD: 1.41 (0.53, 2.29) sig p=0.002 Remission (BDI): MD: 2.27 (0.60, 3.94) sig p=0.008
Kang-Yi[150] 2010	**Psychological versus usual care** Depression severity (GDS): MWD: 0.65 (-0.11, 1.41) sig Depression severity (**SCL-90R**): MWD: 0.50 (0.01, 0.99) sig Depression severity (PHQ): MWD: 0.24 (-0.10, 0.58) sig Depression severity (HADS): MWD: 0.20 (-0.23, 0.64) sig Depression severity (**SDS**): MWD: 0.11 (-0.53, 0.75) n.s
Mazza[151] 2010	**Pharmacological versus Placebo** Change in score (baseline to followup): RD: -2.38 (- 6.11,1.35) n.s. p=0.21
O'Neil[152] 2011	**Psychological versus usual care** Variations of the DSM-IV -Mental HRQOL : SMD: -0.29 (- 0.38,-0.20) (p<0.00001)
Whalley[153] 2011	**Psychological versus usual care** Total Mortality: RR: 0.89 (0.75, 1.05) n.s. p=0.17 Depression: SMD: -0.21 (-0.35, -0.08) sig p=0.0019

Abbreviations: BDI = Beck Depression Index; DSM = Diagnostic Statistics Manual; ER = ermergency room; GDS = Geriatric Depression Scale; HADS = Hospital Anxiety and Depression Scale; HDRS = Hamilton Depression Rating Scale; HRQOL = health-related quality of life; MD = mean difference; MWD = mean weighted difference; n.s. = not significant; OR = Odds Ratio; PHQ = Patient Health Questionnaire; RD = Risk Differences; SCL-90R = the Symptom Checklist (90 item, revised); SDS = Self-rating Depression Scale; sig = statistically significant; SMD = standard mean difference;

Table 16. Summary of outcome and effect size of included systematic reviews– chronic disease n=9

Study Detail	Intervention versus Placebo
	Outcome: Effect size
Rizzo[155] 2011 (chronic physical health issues)	**Psychological versus usual care** Non-response: (Individual-based C&B vs. Usual Care) SMD: -0.55 (−0.97, −0.13) sig; RR: 0.63 (0.23, 1.71) Non-remission: (Group based C&B vs. Usual Care) SMD: -0.58 (−0.99, −0.17) sig; RR: 0.41 (0.22, 0.75) Non-remission: (Group based C&B vs. Usual Care): Outlier removed SMD: -0.28 (-0.47 to −0.10) sig
Taylor[156] 2011 (chronic physical health issues)	**Pharmacological versus usual care** Remission: RR: 0.81 (0.73, 0.91); double blind only: RR: 0.88 (0.81, 0.95) sig Response: RR: 0.83 (0.71, 0.97); double blind only: RR: 0.89 (0.81, 0.98) sig Quality of life: SMD: -0.27 (-0.44, -0.10) sig Remission: RR: 0.70 (0.40,1.25) n.s. Observer-rated depression scales: SMD: -0.70 (-0.97, -0.43) n.s.
Rayner[154] 2010 (Physically illness)	**Pharmacological versus usual care** Response (6 to 8 weeks): OR: 2.33 (1.80, 3.00) sig p<0.000001 Response (4 to 5 weeks): OR: 2.29 (1.16, 4.54) sig p=0.02 Response (9 to 18 weeks): OR: 2.08 (1.33, 3.24) sig p=0.001 Response (>18 weeks): OR: 2.13 (1.31, 3.47) sig p=0.002
Beltman[162] 2010 (somatic disease)	**Psychological versus usual care** Depressive disorder: SMD: -0.83 (-1.36 to 0.31) sig p=0.002 Depressive symptoms: SMD: -0.16 (-0.27 to -0.06) sig p=0.001
Dai[157] 2011 (Medical comorbidity, somatic symptoms)	**Psychological versus usual care** Depression symptom: Cohen's d: 0.80 (0.60, 0.99) sig p <0.001
Straten[161] 2010 (Medical disorder)	**Psychological versus usual care** Change scores (overall): d: 0.42 (0.21, 0.63) sig Change scores: d: 0.47 (0.29, 0.66) sig
Lovieno[158] 2011 (Axis-III disorder (cancer))	**Pharmacological versus usual care** Depression symptom: RR 1.26 (0.88, 1.80) n.s. p 0.19
Van der Feltz-Cornelis[160] 2010 (Type 1 and Type 2 Diabetes Mellitus)	**Pharmacological versus usual care** Depressive symptom severity: Cohen's d: -0.512 (−0.633, -0.390) sig
Meijer[159] 2011 (cancer)	**Psychological versus usual care** Remission: RR: 2.0 (1.1, 3.5) sig No longer having depression: RR 1.4 (1.1, 1.9) sig Response: Hedges's g 0.37 (0.09, 0.65) sig Response: RR: 1.5 (1.1, 2.1) sig

Abbreviations: C&B = cognitive and behavioural based interventions; d = differences; n.s. = not significant; OR = Odds Ratio; sig = statistically significant; SMD = standard mean difference

Table 17. Summary of effect size of included systematic reviews in high risk group: – chronic pain n=2

Study Detail	Intervention versus usual care
	Outcome: Effect size
Henschke[163] 2011	**Psychological versus usual care** Depression (short term): SMD: -0.11 (-0.67, 0.44) n.s. Depression (short term): SMD: -1.92 (-6.16,2.32) n.s.
Urquhart[164] 2010	**Pharmacological versus Placebo** Depression (short term): SMD: 0.06 (-0.29, 0.40) n.s.

Abbreviations: n.s. = not significant; SMD = standard mean difference

Table 18. Summary of effect size of included systematic reviews in high risk group: alcohol and substance abuse disorder n=2

Study Detail	Intervention versus Placebo
	Outcome: Effect size
Pani[165] 2010	**Pharmacological versus Placebo** Severity of depression (HDRS): SMD: -0.31 (-0.64, 0.01) n.s. p=0.06 Severity of depression (MADRS): SMD: 0.18 (-0.39, 0.74) n.s. p=0.54 Severity of depression (POMS): SMD: -0.50 (-1.12, 0.11) n.s. p=0.11 Severity of depression (AUSSI): SMD: 0.32 (-0.24, 0.89) n.s. p=0.26 Severity of depression (Global Improvement Rating): SMD: -0.52 (-1.26, 0.22) n.s. p=0.17 Severity of depression (BDI score): SMD: -0.58 (-2.30, 1.14) n.s. p=0.51 Severity of depression (“Very much” or “much” improved): RR: 2.03 (1.17, 3.53) sig p=0.01 Severity of depression (>50% reduction, HDRS score): RR: 0.96 (0.54, 1.71) n.s. p=0.88 Severity of depression (HDRS): SMD: -0.23 (-0.69, 0.23) n.s. p=0.33 Severity of depression (Global Improvement Rating): SMD: -0.52 (-1.26, 0.22) n.s. p=0.17 Severity of depression (BDI score): SMD: 0.26 (-0.33, 0.85) n.s. p= 0.39
Pedrelli[166] 2011	**Pharmacological versus Placebo** Response: RR: 1.182 (0.822,1.70) n.s. p= 0.366

Abbreviations: AUSSI = Affect Underpinned by the Severity of Social Impairment; BDI = Beck Depression Index; HDRS = Hamilton Depression Rating Scale; MADRS = Montgomery-Asberg Depression Rating Scale; n.s. = not significant; POMS = Profile of Mood States Scale; RR = relative ratio; sig = statistically significant; SMD = standard mean difference

Table 19. Summary of effect size of included systematic reviews in high risk group - perinatal n=2

Study Detail	Intervention versus usual care
	Outcome: Effect size
Cuijpers[167] 2008	**Psychological versus usual care** Depression symptom: d: 0.61 (0.37, 0.85) sig p <0.001 Depression symptom: d: 0.36 (0.15, 0.58) sig
Dennis[168] 2009	**Psychological versus usual care** Evidence of depression at final assessment <1 yr: RR: 0.72 (0.57, 0.90) sig Evidence of depression at final assessment <1 yr: RR: 0.80 (0.66, 0.98) sig Evidence of depression at final assessment <1 yr: RR: 0.67 (0.33, 1.37) n.s. Evidence of depression at final assessment <1 yr: RR: 0.75 (0.63, 0.88) sig Evidence of depression at final assessment immediately post-Tx: RR: 0.79 (0.62, 1.01) n.s. Evidence of depression at final assessment immediately post-Tx: RR: 0.80 (0.66, 0.98) sig Evidence of depression at final assessment immediately post-Tx: RR: 0.48 (0.29, 0.80) sig

Abbreviations: d = differences; n.s. = not significant; RR = relative risk; sig = statistically significant; Tx = treatment

APPENDIX A

Search Strategies

Appendix A: Detailed Search Strategies

KQ1 and 2
Medline-OVID
May 23, 2012
1. depressive disorder/ or depression, postpartum/ or depressive disorder, major/ or dysthymic disorder/ or seasonal affective disorder/
2. Depression/
3. dysthym*.tw.
4. (subclinical adj2 depressi*).tw.
5. (subsyndromal adj2 depressi*).tw.
6. (subthreshold adj2 depressi*).tw.
7. (subdiagnostic adj2 depressi*).tw.
8. (postnatal adj2 depressi*).tw.
9. (postpartum adj2 depression).mp.
10. (perinatal adj2 depressi*).tw.
11. or/1-10
12. mass screening/
13. screen*.mp.
14. case finding.tw.
15. casefinding.tw.
16. or/12-15
17. 11 and 16
18. depressive disorder/ or depression, postpartum/ or depressive disorder, major/ or dysthymic disorder/ or seasonal affective disorder/
19. Depression/
20. dysthym*.tw.
21. (subclinical adj2 depressi*).tw.
22. (subsyndromal adj2 depressi*).tw.
23. (subthreshold adj2 depressi*).tw.
24. (subdiagnostic adj2 depressi*).tw.
25. (postnatal adj2 depressi*).tw.
26. (postpartum adj2 depression).mp.
27. (perinatal adj2 depressi*).tw.
28. or/18-27
29. mass screening/
30. screen*.mp.
31. case finding.tw.
32. casefinding.tw.
33. or/29-32
34. 28 and 33
35. (adverse adj2 (effects or events)).tw.
36. harm*.tw.
37. label*.tw.
38. adverse effects.fs.
39. ((inappropriat* or unnecess* or unneed*) adj3 (treat* or Surg* or therap* or regimen*)).mp.

40. over treatment.mp.
41. overtreatment.mp.
42. or/35-41
43. 34 and 42
44. 17 or 43
45. limit 44 to yr="1994 -Current"

EMBASE-OVID
May 23, 2012
1. (depressi* or dysthm* or affective or mood).ti.
2. (subclinical adj2 depressi*).tw.
3. (subsyndromal adj2 depressi*).tw.
4. (subthreshold adj2 depressi*).tw.
5. (subdiagnostic adj2 depressi*).tw.
6. exp depression/
7. (perinatal adj2 depressi*).tw.
8. (postnatal adj2 depressi*).tw.
9. (postpartum adj2 depression).mp.
10. or/1-9
11. mass screening/
12. screen*.mp.
13. case finding.tw.
14. casefinding.tw.
15. or/11-14
16. 10 and 15
17. human/
18. nonhuman/
19. animal/
20. animal experiment/
21. or/18-21
22. 32 not (32 and 28)
34. 27 not 33
35. trial.tw.
36. 34 or 35
37. 16 and 36
38. limit 37 to yr="1994 -Current"
39. limit 38 to (english or french)
40. limit 39 to (editorial or letter or note)
41. 39 not 40
42. (adverse adj2 (effects or events)).tw.
43. harm*.tw.
44. label*.tw.
45. unnecessary.tw.
46. overtreatment.tw.
47. over treatment.tw.
48. harm reduction/

49. or/42-48
50. (depressi* or dysthm* or affective or mood).ti.
51. (subclinical adj2 depressi*).tw.
52. (subsyndromal adj2 depressi*).tw.
53. (subthreshold adj2 depressi*).tw.
54. (subdiagnostic adj2 depressi*).tw.
55. exp depression/
56. (perinatal adj2 depressi*).tw.
57. (postnatal adj2 depressi*).tw.
58. (postpartum adj2 depression).mp.
59. or/50-58
60. mass screening/
61. screen*.mp.
62. case finding.tw.
63. casefinding.tw.
64. or/60-63
65. 59 and 64
66. 49 and 65
67. limit 66 to yr="1994 -Current"
68. limit 67 to (english or french)
69. 68 or 41

PsycINFO-OVID
May 23, 2012
1. (depressi* or dysythm* or affective or mood).ti.
2. (subclinical adj2 depressi*).tw.
3. (subsyndromal adj2 depressi*).tw.
4. (subthreshold adj2 depressi*).tw.
5. (subdiagnostic adj2 depressi*).tw.
6. (postnatal adj2 depressi*).tw.
7. (postpartum adj2 depression).mp.
8. (perinatal adj2 depressi*).tw.
9. exp depression/
10. exp major depression/
11. or/1-10
12. mental disorders/
13. depress*.ti,ab.
14. 12 and 13
15. 11 or 14
16. health screening/ or screening/
17. screen*.ti,ab.
18. case finding.ti,ab.
19. casefinding.ti,ab.
20. or/16-19
21. clinical trials/
22. random sampling/

23. ((singl$ or doubl$ or trebl$ or tripl$) adj25 (blind$ or mask$)).ti,ab.
24. placebo$.ti,ab.
25. random$.ti,ab.
26. trial.tw.
27. or/21-26
28. (adverse adj2 (effects or events)).tw.
29. harm*.tw.
30. label*.tw.
31. unnecessary.tw.
32. overtreatment.tw.
33. over treatment.tw.
34. ((inappropriat* or unnecess* or unneed*) adj3 (treat* or Surg* or therap* or regimen*)).mp.
35. or/28-34
36. 15 and 20
37. 27 and 36
38. (adverse adj2 (effects or events)).tw.
39. harm*.tw.
40. label*.tw.
41. unnecessary.tw.
42. overtreatment.tw.
43. over treatment.tw.
44. ((inappropriat* or unnecess* or unneed*) adj3 (treat* or Surg* or therap* or regimen*)).mp.
45. psychological stress/
46. or/38-45
47. 36 and 46
48. 37 or 47
49. limit 48 to yr="1994 -Current"
50. limit 49 to (english or french)

EBM Reviews - Cochrane Central Register of Controlled Trials
May 23, 2012
1. depressive disorder/ or depression, postpartum/ or depressive disorder, major/ or dysthymic disorder/ or seasonal affective disorder/
2. Depression/
3. dysthym*.tw.
4. (subclinical adj2 depressi*).tw.
5. (subsyndromal adj2 depressi*).tw.
6. (subthreshold adj2 depressi*).tw.
7. (subdiagnostic adj2 depressi*).tw.
8. (postnatal adj2 depressi*).tw.
9. (postpartum adj2 depression).mp.
10. (perinatal adj2 depressi*).tw.
11. or/1-10
12. mass screening/
13. screen*.mp.
14. case finding.tw.

15. casefinding.tw.
16. or/12-15
17. 11 and 16
18. depressive disorder/ or depression, postpartum/ or depressive disorder, major/ or dysthymic disorder/ or seasonal affective disorder/
19. Depression/
20. dysthym*.tw.
21. (subclinical adj2 depressi*).tw.
22. (subsyndromal adj2 depressi*).tw.
23. (subthreshold adj2 depressi*).tw.
24. (subdiagnostic adj2 depressi*).tw.
25. (postnatal adj2 depressi*).tw.
26. (postpartum adj2 depression).mp.
27. (perinatal adj2 depressi*).tw.
28. or/18-27
29. mass screening/
30. screen*.mp.
31. case finding.tw.
32. casefinding.tw.
33. or/29-32
34. 28 and 33
35. (adverse adj2 (effects or events)).tw.
36. harm*.tw.
37. label*.tw.
38. adverse effects.fs.
39. ((inappropriat* or unnecess* or unneed*) adj3 (treat* or Surg* or therap* or regimen*)).mp.
40. over treatment.mp.
41. overtreatment.mp.
42. or/35-41
43. 34 and 42
44. 17 or 43
45. limit 44 to yr="1994 -Current"

Context Questions for Treatment
Medline-OVID (Treatment)
April 6 2011
1. *Depression/dh, de, dt, mo, px, rh, th [Diet Therapy, Drug Effects, Drug Therapy, Mortality, Psychology, Rehabilitation, Therapy]
2. exp *Depressive Disorder/dh, de, dt, mo, px, rh, th
3. or/1-2
4. exp *Antidepressive Agents/
5. exp *Exercise Therapy/
6. exp *Behavior Therapy/
7. or/4-6
8. 3 and 7
9. meta-analysis/

10. exp meta-analysis as topic/
11. (meta analy* or metaanaly* or met analy* or metanaly*).tw.
12. review literature as topic/
13. (collaborative research or collaborative review* or collaborative overview*).tw.
14. (integrative research or integrative review* or intergrative overview*).tw.
15. (quantitative adj3 (research or review* or overview*)).tw.
16. (research integration or research overview*).tw.
17. (systematic* adj3 (review* or overview*)).tw.
18. (methodologic* adj3 (review* or overview*)).tw.
19. exp technology assessment biomedical/
20. (hta or thas or technology assessment*).tw.
21. ((hand adj2 search*) or (manual* adj search*)).tw.
22. ((electronic adj database*) or (bibliographic* adj database*)).tw.
23. ((data adj2 abstract*) or (data adj2 extract*)).tw.
24. (analys* adj3 (pool or pooled or pooling)).tw.
25. mantel haenszel.tw.
26. (cohrane or pubmed or pub med or medline or embase or psycinfo or psyclit or psychinfo or psychlit or cinahl or science citation index).ab.
27. or/9-26
28. 8 and 27
29. limit 28 to yr="2005 -Current"
30. limit 29 to (english or french)

Medline-OVID (Adverse Events Treatment)
March 24 2011
1. depression/
2. exp depressive disorder/
3. or/1-2
4. exp *Antidepressive Agents/ae, ct, po, to [Adverse Effects, Contraindications, Poisoning, Toxicity]
5. exp *Exercise Therapy/ae, ct, po, to
6. exp *Behavior Therapy/ae, ct
7. (adverse or harm?).ti.
8. 5 or 6 or 7
9. 3 and 8
10. 4 or 9
11. meta-analysis/
12. exp meta-analysis as topic/
13. (meta analy* or metaanaly* or met analy* or metanaly*).tw.
14. review literature as topic/
15. (collaborative research or collaborative review* or collaborative overview*).tw.
16. (integrative research or integrative review* or intergrative overview*).tw.
17. (quantitative adj3 (research or review* or overview*)).tw.
18. (research integration or research overview*).tw.
19. (systematic* adj3 (review* or overview*)).tw.
20. (methodologic* adj3 (review* or overview*)).tw.

21. exp technology assessment biomedical/
22. (hta or thas or technology assessment*).tw.
23. ((hand adj2 search*) or (manual* adj search*)).tw.
24. ((electronic adj database*) or (bibliographic* adj database*)).tw.
25. ((data adj2 abstract*) or (data adj2 extract*)).tw.
26. (analys* adj3 (pool or pooled or pooling)).tw.
27. mantel haenszel.tw.
28. (cohrane or pubmed or pub med or medline or embase or psycinfo or psyclit or psychinfo or psychlit or cinahl or science citation index).ab.
29. or/11-28
30. 10 and 29
31. animals/ not (animals/ and humans/)
32. 30 not 31
33. limit 32 to (english or french)
34. limit 33 to yr="2005 -Current"

Medline-OVID (Patient Preferences Treatment)
March 28 2011
1. *"patient acceptance of health care"/ or *patient compliance/ or *patient participation/ or patient satisfaction/ or patient preference/ or *treatment refusal/
2. (women? adj3 (acceptance or preference? or satisfaction or experience?)).tw.
3. (consumer? adj3 (acceptance or preference? or satisfaction or experience?)).tw.
4. (patient? adj3 (acceptance or preference? or satisfaction or experience?)).tw.
5. willingness to pay.tw.
6. ((conjoint or contingent) adj3 (valuation or analysis)).tw.
7. or/1-6
8. *Depression/dh, de, dt, mo, px, rh, th [Diet Therapy, Drug Effects, Drug Therapy, Mortality, Psychology, Rehabilitation, Therapy]
9. exp *Depressive Disorder/dh, de, dt, mo, px, rh, th
10. or/8-9
11. exp *Antidepressive Agents/
12. exp *Exercise Therapy/
13. exp *Behavior Therapy/
14. or/11-13
15. 10 and 14
16. 7 and 15
17. limit 16 to (english or french)
18. limit 17 to yr="2005 -Current"

APPENDIX B

Reviewers of Protocol

Appendix B. Reviewers of Protocol – Screening for Depression

Dan Bilsker, PhD
Adjunct Professor, Faculty of Health Sciences, Simon Fraser University
Clinical Assistant Professor, Faculty of Medicine, University of B.C.
Vancouver, BC Canada

Gerald Devins, PhD
Senior Scientist, Ontario Cancer Institute (OCI)
University of Toronto
Toronto, ON Canada

Rachel Fleurence, Ph.D
Research Scientist, Managing Director of Health Economics
Center for Health Economics, Epidemiology and Science Policy
United BioSource Corporation, Bethesda, MD, USA

Elizabeth O'Connor, PhD
Oregon Health Sciences University
Research Associate, Oregon Evidence-based Practice Center
Portland, Oregon USA

Heather Stuart, Ph.D
Professor, Department of Psychiatry, Queen's University
Kingston, ON Canada

APPENDIX C

Evidence Tables
for General Population – Elderly

C1. Characteristics of Included Studies

Study	Description of Study Population	Definition of Population	Evaluation of Population & Follow-up	Outcomes Descriptions
Oyama, H.[87] 2006 Design: Quasi-experimental Screening Setting: Matsudai, Japan (rural)	**Intervention** Before: 11,567 PY After: 13,330 PY **Control** Before: 15,055 PY After: 19,333 PY Age Mean: NR Age Range: ≥65 Age Median: NR Female: 57.6% Ethnicity: Japanese Education: NR Dx: Major and minor depression	Elderly (≥65 years old) residents living in six rural municipalities of southwest and central Japan Int: mental health workshop, referral to general practitioner or followup interview with public health nurse Exclusions: severely disabled or hospitalized cases were excluded from the study	Screening Instrument: SDS Other Rating: RDC Confirmatory Exam: ICD-9 Number of followups: 10 Number of stages: 2 ten-year **Follow-up**:10-year period prior to intervention, compared to 10-year period after beginning of intervention. Of eligible older adults, more than 90% in both men and in women participated in program	Main Outcome: Changes in suicide risk (Age-adjusted IRRs of completed suicide before and after) The female risk of completing suicide in the intervention area was reduced by 70%, while there was no change in the risk for males in the intervention area. Intervention: 1.02 (95% CI 0.49-2.13) in men, and 0.30 (95% CI 0.14-0.67) in women Control: No significant change

C1. Characteristics of Included Studies

Study	Description of Study Population	Definition of Population	Evaluation of Population & Follow-up	Outcomes Descriptions
Oyama, H.[85] 2006 Design: Quasi-experimental Screening Setting: Yasuzuka, Japan (rural)	**Intervention** Before: 9,791 PY After: 11,901 PY **Control** Before: 16,032 PY After: 20,909 PY Age Mean: NR Age Range: ≥65 Age Median: NR Female: NR Ethnicity: Japanese Education: NR Dx: Major and minor depression	Elderly (≥65 years old) residents of an agricultural rural area in Japan with a high suicide rate Int: The intervention included (a) public health education from 1991 to 2000 and (b) screening for depression with followup from 1991 to 1997 Exclusions: NR	Screening Instrument: SDS Other Rating: RDC Confirmatory Exam: ICD-9 Number of followups: 7 Number of stages: 2 ten-year **Follow-up**:10-year period prior to intervention, compared to 10-year period after beginning of intervention. Of eligible older adults, more than 90% participated in program	Main Outcome: Changes in suicide risk (Age-adjusted IRRs of completed suicide before and after) The risk for women in the intervention area was reduced by 64% whereas there was no significant change for men in the intervention area. Intervention: 0.51 (95% CI 0.22-1.19) in men, and 0.36 (95% CI 0.14-0.93) in women Control: No significant change

C1. Characteristics of Included Studies

Study	Description of Study Population	Definition of Population	Evaluation of Population & Follow-up	Outcomes Descriptions
Oyama, H.[86] 2004 Design: Quasi-experimental Screening Setting: Joboji town, Japan (rural)	I**ntervention** Before: 9,721PY After: 13,032 PY **Control** Before: 17,166 PY After: 25,333 PY Age Mean: NR Age Range: ≥65 Age Median: NR Female: 50.8% Ethnicity: Japanese Education: NR Dx: Depression (unspecified)	Elderly (≥65 years old) residents of an agricultural rural area in Japan with a high suicide rate Int: Two-step depression screening performed by PHN and psychiatrist and followup conducted by psychiatrist every three years in targeted district of an intervention municipality, health education and emphasis on suicide taboo every year in 10-year period from 1990 Exclusions: Elderly people receiving social welfare	Screening Instrument: SDS Other Rating: SADD Confirmatory Exam: ICD-9 Number of followups: 10 Number of stages: 3 five-year **Follow-up:**10-year period prior to intervention, compared to 10-year period after beginning of intervention. Of eligible older adults, 78% in all districts in initial year and after that approximately 60–89% in targeted district participated in program	Main Outcome: Changes in suicide risk (Age-adjusted IRRs of completed suicide before and after) In the intervention area, a 73% reduced risk of suicidal mortality among males, and a 76% reduced risk of suicidal mortality among females during the implementation decade, compared with the pre-implementation decade Intervention: 0.27 (95% CI 0.08-0.88) in men, and 0.24 (95% CI 0.11-0.52) in women Control: No significant change

C1. Characteristics of Included Studies

Study	Description of Study Population	Definition of Population	Evaluation of Population & Follow-up	Outcomes Descriptions
Oyama, H.[88] 2006 Design: Quasi-experimental Screening Setting: Nagawa town, Japan (rural)	**Intervention** Before: 1,982 PY After: 2,634 PY **Control** Before: 16,754 PY After: 19,686 PY Age Mean: NR Age Range: ≥65 Age Median: NR Female: 59-60.8% Ethnicity: Japanese Education: NR Dx: Depression (unspecified)	Elderly (≥65 years old) residents of an agricultural rural area in Japan with a high suicide rate Int: SUPPRESS program (two-stepped screening for depression and followup by PHN, mental health workshop 3 to 4 times a year, and a group activity program once a month Exclusions: NR	Screening Instrument: SDS Other Rating: RDC Confirmatory Exam: ICD-9 Number of followups: 6 Number of stages: 2 six-year **Follow-up:** 6-year period prior to intervention, compared to 6-year period after beginning of intervention. Of eligible older adults, approximately 60–89% in men and 80–95% in women participated in program	Main Outcome: Changes in suicide risk (Age-adjusted IRRs of completed suicide before and after) The risk for elderly females was reduced by 74% while there was no change in the risk for males in the intervention area. Intervention: 0.48 (90% CI 0.10-2.31) in men, and 0.26 (90% CI 0.07-0.98) in women Control: No significant change

C1. Characteristics of Included Studies

Study	Description of Study Population	Definition of Population	Evaluation of Population & Follow-up	Outcomes Descriptions
Oyama, H.[89] 2010[89] Design: Quasi-experimental Screening Setting: Six rural municipalities of the Sanpachi Second Medical Zone, Japan (rural)	**Intervention** Before: 28,838PY After: 29,156 PY **Control** Before: 27,633PY After: 28,063 PY Age Mean: NR Age Range: ≥60 Age Median: NR Female: 57.5% Ethnicity: Japanese Education: NR Dx: Depression (unspecified)	Elderly (≥60 years) residents living in six rural municipalities of the Sanpachi Second Medical Zone (a mostly agricultural region with a high suicide rate Int: The intervention included (1) health education and (2) screening for depression with followup, using the community resources of primary care and public health nursing Exclusions: NR	Screening Instrument: CES-D, DSS Other Rating: Zung-SDS, GDS-5, CIDI Confirmatory Exam: ICD-10 Number of followups: 2 Number of stages: 2 two-year **Follow-up**:12-year period prior to intervention, compared to 2-year period after beginning of intervention. Of eligible older adults, more than 85% participated in program	Main Outcome: Changes in suicide risk (Age-adjusted IRRs of completed suicide before and after) In the intervention region there was a 61% reduction in risk of suicide among men aged 60 and over. The 51% reduction in risk in women aged 60 and over did not reach statistical significance. Intervention: 0.39 (90% CI 0.18-0.87) in men, and 0.49 (90% CI 0.19-1.22) in women Control: No significant change

Abbreviations: 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 = Geriatric Depression Scale of five items; ICD = International Statistical Classification of Diseases; Int = Intervention; 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**;** Zung-SDS = Zung Self-Rating Depression Scale

C2. Characteristics of Included Studies - Outcomes

Study Setting Study Design	Description of Study Population Screening Tool(s)	Number and Length of followups Study Duration	Approach to intervention beyond screening results feedback	N (%) of Int depressed at followup Screening Tool	% reduced risk of suicide *p<0.05 **p<0.01 Diagnostic Method
Oyama, H. 2006[87] Matsudai town, Japan (rural) Quasi-experimental	N (residents) = 6,015 (19% elderly aged ≥65) Age Range: ≥65 Female: 3,016 (50.1%) Male: 2,999 (49.9%) Int: 11,567 person-years and 35 suicide victims during the baseline, and 13,330 person-years and 24 suicide victims during the implementation. Ctrl.: 15,055 person-years and 27 suicide victims during the baseline, and 19,333 person-years and 25 suicide victims during the intervention Screening Tool(s): SDS; RDC	Screened and followed up annually for 10 years (April 1988-March 1998) 10-year baseline period (April 1978-March 1988) 10-year intervention period (April 1988-March 1998) Total: 20 years	Annual two-step depression screening performed by PHN and psychiatrist and followup conducted by GP, and health education in 10-year period from 1988.	218 (1.6%) RDC	Int: 70%* reduced risk of suicide in women ≥ 65 years; no significant change in risk of suicide for men ≥65 Ctrl.: No significant change in risk of suicide for men and women ICD-9
Oyama, H. 2006[85] Yasuzuka, Japan (rural) Quasi-experimental	N (residents) = 4,940 (21% elderly aged ≥65) Age Range: ≥65 Female: 2,444 (49.5%) Male: 2,496 (50.5%) Int: 9,791 person-years and 28 suicide victims during the baseline, and 11,901 person-years and 16 suicide victims during the implementation. Ctrl.: 16,032 person-years and 30 suicide victims during the baseline, and 20,909 person-years and 25 suicide victims during the intervention Screening Tool(s): SDS; RDC	Screened and followed up annually for 7 years (1991-1997) 10-year baseline period (1981-1990) 10-year intervention period (1991-2000) Total: 20 years	Annual two-step depression screening performed by PHN and psychiatrist and followup conducted by GP in 7-year period from 1991, and health education in 10-year period from 1991.	NR	Int: 64% reduced risk of suicide in women ≥65 years; 49% reduction of the risk in men ≥65 years Ctrl.: No significant change in risk of suicide for men and women ICD-9

C2. Characteristics of Included Studies – Outcomes

Study Setting Study Design	Description of Study Population Screening Tool(s) Diagnostic Method	Number and Length of followups Study Duration	Approach to intervention beyond screening results feedback	N (%) of Int depressed at followup Screening Tool	% reduced risk of suicide *p<0.05 **p<0.01 Diagnostic Method
Oyama, H.[86] 2004 Joboji, Japan (rural) Quasi-experimental	N (residents) = 7,030 (13% elderly aged ≥65) Age Range: ≥65 Female: 3,570 (50.8%) Male: 3,460 (49.2%) Int: 9,721 person-years and 29 suicide victims during the baseline, and 13,032 person years and 10 suicide victims during the implementation. Ctrl.: 17,166 person-years and 33 suicide victims during the baseline, and 25,333 person-years and 43 suicide victims during the intervention Screening Tool(s): SDS; SADD	Screened annually for 12 years (1988-1999), followed up 4 months after each screening 5-year baseline period (1985-1989) 10-year intervention period (1990-1999) Total: 15 years	Two-step depression screening performed by PHN and psychiatrist and followup conducted by psychiatrist, health education and emphasis on suicide taboo every year in 10-year period from 1990.	approximately 133 to 301 (10-20%) SDS	Int: 69% (intention stage); 81% (maintenance stage) reduced risk of suicide in women ≥ 65 years reduced; men NR Ctrl.: No significant change in risk of suicide for men and women ICD-9
Oyama, H. 2006[88] Nagawa town, Japan (rural) Quasi-experimental	N (residents) = 1,685 (25% elderly aged ≥65) Age Range: ≥65 Female: NR Male: NR Int: 1,982 person-years and 8 suicide victims during the baseline, and 2,634 person-years and 4 suicide victims during the implementation. Ctrl.: 16,754 person-years and 21 suicide victims during the baseline, and 1,686 person-years and 25 suicide victims during the intervention Screening Tool(s): Five self-reported items; mental health assessment by PHN	Screened and followed up annually for 6 years (1999-2004) 6-year baseline period (1993-1998) 6-year intervention period (1999-2004) Total: 12 years	Annual two-step depression screening performed by PHN and psychiatrist and followup conducted by GP and to a lesser extent psychiatrist, group activity and health education in 6-year period from 1999.	27 to 51 (6.7-11%) Mental health assessment by PHN	Int: 74%* reduced risk of suicide in women ≥ 65 years; no significant change in risk of suicide for men ≥ 65 Ctrl.: No significant change in risk of suicide for men and women ICD-9

C2. Characteristics of Included Studies – Outcomes

Study Setting Study Design	Description of Study Population Screening Tool(s) Diagnostic Method	Number and Length of followups Study Duration	Approach to intervention beyond screening results feedback	N (%) of Int depressed at followup Screening Tool	% reduced risk of suicide *p<0.05 **p<0.01 Diagnostic Method
Oyama, H. 2010[89] Six rural municipalities of the Sanpachi Second Medical Zone, Japan (rural) Quasi-experimental	N (residents) = 41,337 (35% elderly aged ≥60) Age Range: ≥60 Female ≥60: 8,556 (59.0% of 35%) Male ≥60: 5,948 (41.0% of 35%) Int: 28,836 person-years and 25 suicide victims during the baseline, and 29,156 person-years and 11 suicide victims. Ctrl.: 27,183 person-years and 21 suicide victims during the baseline, and 27,973 person-years and 27 suicide victims. Screening Tool(s): CES-D; DSS; Zung-SDS; GDS-5; CIDI	Screened and followed up annually for 2 years (2005-2006) 2-year baseline period (2003-2004) 2-year intervention period (2005-2006) Total: 4 years	Two-step depression screening performed by PHN and psychiatric social workers and followup conducted by psychiatrist or PHN, group activity and health education in 2-year period from 2005.	19/420 (4.5%) CIDI	Int: 51% reduction of the risk in women ≥ 60 years; 61%* reduced risk of suicide in men ≥ 60 years Ctrl.: No significant change in risk of suicide for men and women ICD-10

Note: % currently or recently treated for depression = NR
Abbreviations: CES-D = Center for epidemiologic studies depression scale; CIDI = Composite International Diagnostic Interview; Ctrl = Control; DSS = Depression suicide screen; GDS-5 = Geriatric Depression Scale of five items; Gp = general practitioner; ICD = International Statistical Classification of Diseases; Int = Intervention; NR = not reported; PHN = public health nurse; RDC = Research Diagnostic Criteria; SADD = Schedules of Standardized Assessment of Patient ith Depressive Disorders; SDS = Self-rating Depression Scale; Zung-SDS = Zung Self-Rating Depression Scale

APPENDIX D

The Method of Calculation of Ratio of Rate Ratios (RRR)

Appendix D. Methods: Ratio of Rate Ratios (RRR)

Data was presented in the papers as pre (baseline) and post (implementation) analysis for both intervention and control [85-89] groups. Baseline events (n) and implementation data (person-years) were given.

The following 2x2 table shows how calculations were performed. Rates for both pre and post groups were calculated as number of events (n) divided by person-years. In the cells from the table r_{11}, r_{12}, r_{21} and r_{22} represent these rates.

	Intervention	**Control**
Before	r_{11}	r_{12}
After	r_{21}	r_{22}

The rate ratios (RR) of the intervention and control groups were estimated as r_{21}/r_{11} for intervention group and r_{22}/r_{12} for the control group. The logarithm of the RR's and corresponding standard error ($\sqrt{1/n_1 + 1/n_2}$) for both intervention and control groups were calculated. The difference between the log (RR) values for the intervention and control groups along with the corresponding standard error $\sqrt{S_I^2 + S_C^2}$ where s_i^2 and s_c^2 are the variances of the RRs for intervention and control, respectively. This process was used to calculate the difference between log (RR) values and standard errors for the intervention and control groups.

Example: Oyama H, Fujita M, Goto M, et al. Outcomes of community-based screening for depression and suicide prevention among Japanese elders. Gerontologist. 2006;46(6):821-6.
Males 65-74

Intervention		Control	
Event Before	9	Event Before	5
Total Before	2765	Total Before	4152
Event After	7	Event After	6
Total After	3341	Total After	5419
Rate Ratio Before	0.003254973	Rate Ratio Before	0.001204239
Rate Ratio After	0.002095181	Rate Ratio After	0.001107215
Ratio of Rate Ratios (RRR)	0.643686188	Ratio of Rate Ratios (RRR)	0.919431629
Log(RRR)	-0.440543957	Log(RRR)	-0.083999594
SE(log(RRR))	0.503952631	SE(log(RRR))	0.605530071

The events and totals before and after from the table are actual data taken from Oyama et al (2006) for males in the age group 65-74. The following are calculations for the rest of the table.

Rate ratios were calculated as:

Rate Ratio Before (Intervention) = 9/2765 = 0.003254973

Rate Ratio After (Intervention) = 7/3341 = 0.002095181

Rate Ratio Before (Control) = 5/4152 = 0.001204239

Rate Ratio After (Control) = 6/5419 = 0.001107215

Ratio of Rate Ratios (RRR) were calculated as:

RRR (Intervention) =
Rate Ratio After/Rate Ratio Before = 0.002095181/0.003254973 = 0.643686188

RRR (Control) =
Rate Ratio After/Rate Ratio Before = 0.001107215/0.001204239 = 0.919431629

Log(RRR) were calculated as:

Log(RRR) (Intervention) = log(0.643686188) = -0.440543957

Log(RRR) (Control) = log(0.919431629) = -0.083999594

SE(log(RRR)) were calculated as:

$$SE(Intervention) = \sqrt{\frac{1}{EB} + \frac{1}{EA}} = \sqrt{\frac{1}{9} + \frac{1}{7}} = 0.503952631$$

$$SE(Control) = \sqrt{\frac{1}{EB} + \frac{1}{EA}} = \sqrt{\frac{1}{5} + \frac{1}{6}} = 0.605530071$$

where EB and EA are the events for before and after respectively.

The final part was to calculate the difference between log(RRR Intervention) and log(RRR Control),

Difference =
log(RRR Intervention) – log(RRR Control) = -0.440543957 – (-0.083999594) = -0.356544363

and the corresponding standard error (SE) was calculated as:

$$SE(Difference) = \sqrt{SE(Intervention)^2 + SE(Control)^2}$$
$$= \sqrt{(0.503952631)^2 + (0.605530071)^2} = 0.787803859$$