Systematic Reviews

Screening for depression in children and adolescents: a protocol for a systematic review update

Andrew Beck[1*] , John C. LeBlanc[2], Kate Morissette[3], Candyce Hamel[1], Becky Skidmore[1], Heather Colquhoun[4], Eddy Lang[5,6], Ainsley Moore[7,8], John J. Riva[7,8], Brett D. Thombs[9,10], Scott Patten[11], Heather Bragg[12], Ian Colman[13], Gary S. Goldfield[14], Stuart Gordon Nicholls[15], Kathleen Pajer[16], Beth K. Potter[13], Robert Meeder[17], Priya Vasa[18], Brian Hutton[1], Beverley J. Shea[1,13], Eva Graham[3], Julian Little[13], David Moher[1,13] and Adrienne Stevens[1]

Abstract

Background: Major depressive disorder is common, debilitating, and affects feelings, thoughts, mood, and behaviors. Childhood and adolescence are critical periods for the development of depression and adolescence is marked by an increased incidence of mental health disorders. This protocol outlines the planned scope and methods for a systematic review update that will evaluate the benefits and harms of screening for depression in children and adolescents.

Methods: This review will update a previously published systematic review by Roseman and colleagues. Eligible studies are randomized controlled trials (RCTs) assessing formal screening in primary care to identify children or adolescents not already self-reporting symptoms of, diagnosed with, or treated for depression. If no or only a single RCT is available, we will consider controlled studies without random assignment. Studies of participants with characteristics associated with an elevated risk of depression will be analyzed separately. Outcomes of interest are symptoms of depression, classification of major depressive disorder based on a validated diagnostic interview, suicidality, health-related quality of life, social function, impact on lifestyle behavior (e.g., substance use, school performance, lost time at work, or school), false-positive results, overdiagnosis, overtreatment, labeling, and other harms such as those arising from treatment. We will search MEDLINE, Embase, PsycINFO, CINAHL, the Cochrane Library, and grey literature sources. Two reviewers will independently screen the titles and abstracts using the liberal accelerated method. Full-text screening will be performed independently by two reviewers using pre-specified eligibility criteria. Data extraction and risk of bias assessments will be performed independently by two reviewers. Pre-planned analyses, including subgroup and sensitivity analyses, are detailed within this protocol. Two independent reviewers will assess and finalize through consensus the certainty of evidence using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach, and prepare GRADE evidence profiles and summary of findings tables for each outcome of interest.

(Continued on next page)

* Correspondence: andrew.beck@uottawa.ca
[1]Knowledge Synthesis Group, Clinical Epidemiology Program, Ottawa
Hospital Research Institute, Centre for Practice-Changing Research, 501
Smyth Road, Box 201, Ottawa, ON K1H 8L6, Canada
Full list of author information is available at the end of the article

BMC

Beck *et al. Systematic Reviews* (2021) 10:24

Page 2 of 13

(Continued from previous page)

Discussion: The systematic review will provide a current state of the evidence of benefits and harms of depression screening in children and adolescents. These findings will be used by the Canadian Task Force on Preventive Health Care to inform the development of recommendations on depression screening.

Systematic review registration: PROSPERO CRD42020150373

Keywords: Depression, screening, systematic review, child, children, adolescent, youth

Background

Major depressive disorder (MDD) is a common, debilitating mood disorder characterized by negative feelings, thoughts, and behaviors that causes significant impairment in social, occupational and educational functioning, and quality of life, and is related to an increased risk of suicide and death [1, 2]. During the transition from childhood to adolescence, there is an increase in prevalence of some psychiatric disorders including depression [3]. This transition involves physical, psychological, and emotional changes typical of this developmental period, which may increase an individual's sensitivity and reactivity to stress exposure [4, 5]. As with the adult population, diagnoses of depressive episodes in children and adolescents are established by one of the two commonly used diagnostic classification systems for psychiatric diagnoses, the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) [6], and International Classification of Diseases, 10th Revision (ICD-10) [7]. Each diagnostic system provides a minimum number of criteria that must be met over a 2-week period for an episode to be considered a depressive episode. Also, the DSM-5 includes further criteria to specifically define MDD for children and youth [6]. Symptoms of irritability can be considered in place of depressed mood, and a failure to meet expected weight gain can be considered instead of weight loss (see Additional file 1).

Prevalence and burden

From the 2014 Ontario Child Health Study, the 6-month prevalence of possible major depressive episodes (MDE) was 1.1% for children (4- to 11-year-old) and 5.2% or 7.5% for adolescents (12- to 17-years-old) based on parent or adolescent report, respectively [8]. In pooled estimates from the Canadian Community Health Survey, a series of cross-sectional surveys from 2000 to 2014, 5.5% of 12- to 19-year-olds reported experiencing MDE-like episodes in the past year, with little change in prevalence from 2000 to 2014 [9]. Rates were higher among females and among those aged 15 to 19 years (10.1% for females; 4.1% for males) compared to those aged 12 to 14 years (4.1% for females; 0.6% for males) [9], with similar findings supported by other literature [10–12].

The burden of depression is high among children and adolescents. Unipolar depressive disorders (i.e., major depressive episode, dysthymia) are a leading cause of years lost to disability among both the 10- to 14-year-old and 15- to 19-year-old age groups [13, 14]. Poor long-term social outcomes are also a consequence of depression in adolescence. Those with depression are at an increased risk of leaving secondary school, unemployment, early pregnancy, and parenthood [15]. As well, they have a lower likelihood of entering post-secondary education [15]. Depression with onset in childhood and adolescence can continue into adulthood, posing a burden on individuals, families, and communities [15–18]. A recent systematic review found that adolescents who suffer from depression have around 2.5 (2.78 [95% CI 1.97, 3.93]) times the odds of developing depression in adulthood compared with adolescents without depression [16]. Additionally, those who suffer from depression in adolescence are at an increased risk for suicidal ideation, attempts, and completion in adulthood [19–21].

Risk factors

There are several risk factors associated with depression in children and adolescents. As shown above, females are at a higher risk, particularly in adolescence, with the difference between sexes becoming smaller later in adulthood [12, 22]. A family history of depression and exposure to adverse events such as illness or death of a family member, physical or sexual abuse, are also strong risk factors [23, 24]. Parental behaviors associated with an increased risk include aversive behaviors toward the child or adolescent (e.g., criticism, punishment, and conflict), lack of autonomy given to the child or adolescent, lack of warmth, inconsistent parental discipline, and parental over-involvement [25]. Other influential factors include aspects related to the school environment such as bullying, low connectedness with peers and teachers [26, 27], poor academic achievement [28], learning disabilities (e.g., attention deficit hyperactivity disorder, dyslexia) [29–31], and community environment factors such as safety, ethnicity, and prevalence of discrimination [32]. Additional risk factors include substance abuse (e.g., alcohol, tobacco, cannabis, other illicit drugs), poor sleep, screen time, unhealthy diet, and weight problems [33–35].

Screening of MDD in children and adolescents

A screening program for depression must identify symptomatic disease that would not otherwise be reported (e.g., by spontaneous patient self-report, parent/caregiver report, or clinical inquiry). If effective, screening for depression would be expected to lead to interventions that improve future health outcomes in those who otherwise would not have been identified [36]. However, as noted by Cosgrove et al. [37], without evidence on the benefits and harms of screening, there are several assumptions that are questionable in the case of depression screening. First, unlike other disorders, depression does not have a detectable asymptomatic early stage. It manifests as one or more discrete episodes and many patients remit after an initial episode. Screening tools rely on identifying symptoms of depression itself and therefore can only be effective at early detection if the use of the tool prompts consideration of whether or not symptoms of depression are present. Second, there is currently little evidence that adding screening questionnaires to primary care reduces depressive symptoms. Third, optimal treatment for screen-detected depression is not clear. Most people identified as depressed via screening will have mild symptoms that may resolve without intervention. Many are treated with antidepressant medications. However, the majority of antidepressant medications have not been shown to be as effective in adolescents as in adults, may increase risk for suicidality, and may be even less likely to be effective for the mild cases likely overrepresented in patients identified through screening questionnaires [38, 39]. Psychological therapies are a reasonable option to medications with some positive effects [40, 41].

In considering the potential benefits and harms of screening, the proper design of trials is paramount [42, 43]. Experts have criticized previous systematic reviews on depression screening that have not explicitly defined the characteristics of screening trials to isolate the effect of screening from, for example, diagnostic suspicion by patients or clinicians [44, 45]. For example, a screening trial should separate the effect of screening from the effect of providing additional treatment resources not otherwise available. Since screening programs should only identify previously unrecognized cases, a screening trial should also exclude patients already diagnosed or under care for depression [42]. Accordingly, we have specified criteria that would need to be met to reduce selection and confounding bias in the inclusion of evidence. These are detailed in the "Methods" section.

Previous guideline recommendations using systematic reviews from the Canadian Task Force on Preventive Health Care and other guideline developers

In 2005, the Canadian Task Force on Preventive Health Care ("Task Force") published recommendations on screening for depression in the adult population as well as children and adolescents. The Task Force guideline concluded that there was insufficient evidence to recommend either for or against depression screening in children and adolescents [46]. In 2016, the U.S. Preventive Services Task Force (USPSTF) updated their 2009 guideline and systematic review and recommended routinely screening for depression in adolescents (age 12 to 18 years) in primary care settings "with adequate systems in place to ensure accurate diagnosis, effective treatment, and appropriate follow-up," but not in children (aged ≤ 11 years) due to insufficient evidence [47, 48]. The USPSTF recommendation for screening relied on indirect and linked evidence that screening instruments for depression can accurately identify MDD in adolescents and that treatment of MDD detected through screening in adolescents is associated with moderate benefit. However, this was not based on screening trials, as no trials were identified that directly assessed the effects of screening compared with no screening [48]. The recommendation was based on eight "fair- "or "good-quality" placebo-controlled intervention trials examining the effectiveness of antidepressants, psychotherapy, and collaborative care interventions (e.g., the Massachusetts Child Psychiatry Access Project, an inter-professional collaboration among primary care providers and mental health specialists) for children and adolescents with MDD. They found that one trial had shown effectiveness with fluoxetine as well as combined fluoxetine and cognitive behavioral therapy. Furthermore, one of the two escitalopram trials and the one collaborative care trial had demonstrated benefits. The USPSTF considered the evidence on screening test accuracy and treatment of MDD adequate and the harms of pharmacotherapy treatment to be minimal if patients were closely monitored. Similar to the USPSTF, the Guidelines for Adolescent Depression in Primary Care (GLAD-PC) Steering Group updated their 2007 systematic literature review and guideline [49] and found no evidence comparing screening to no screening or usual care in primary care settings [50]. They relied upon indirect evidence surrounding the validity and feasibility of screening tools and the feasibility and effectiveness of treatment for individuals with depression. GLAD-PC recommended annual screening for depression in adolescents (ages 12 and older) based on psychometric data on depression screening tools as well as research on screening issues (e.g., whether screening is occurring and whether screening impacts follow-up procedures or treatment to the specifics of screening). They considered that early identification and treatment of adolescent depression is important because of the high prevalence of depression and its persistence, its potential occurrence during a crucial brain development period, and the potential

Beck *et al. Systematic Reviews*   (2021) 10:24

Page 4 of 13

consequences with transition into adulthood. Among those with depression risk factors ages 10 to 21 years, GLAD-PC recommended not only annual screening but also more frequent targeted screening (i.e., systematically monitoring individuals over time for the development of a depressive disorder) during other health care visits. Their justification for this recommendation was that these individuals are likely to experience future depressive episodes and those who are not diagnosed with depression but report high scores on the screening tools may be at risk for depression within 6 months [50].

Rationale

There are inconsistencies among the current screening guidelines reviewed above, notably their reliance on indirect evidence and the failure to include RCTs where the potential impacts of screening and treatment are clearly separated. A recent review by Roseman and colleagues [51] was selected by the Task Force working group as an appropriate review to update but with modifications, such as also addressing patients that may have an elevated risk of depression, including other outcomes of relevance to decision making, other study design filters, and expanding the search approach. This review update will provide a current assessment of the evidence for the Task Force guideline recommendations.

Review objective and key questions

The Task Force is undertaking a systematic evaluation of the literature to inform its upcoming guideline recommendation on depression screening in children (6 to 11 years old) and adolescents (12 to 17 years old) in primary care. A depression working group of task force members (i.e., guideline panel) and external clinical experts was formed. The development of the topic, refinement of the key questions, and the structured Population, Intervention, Comparator, Outcome, Study Design (PICOS) eligibility framework involved discussions among the working group and members of the Public Health Agency of Canada (PHAC) and the Ottawa Evidence Review and Synthesis Centre (ERSC). For more information on the process of topic development and selection of working group members and clinical experts, please refer to the Task Force Procedure Manual (https://canadiantaskforce.ca/methods/).

The purpose of this systematic review is to examine the evidence on screening for depression in children and adolescents within primary care, and the findings will be used by the Task Force to inform the development of guideline recommendations. This protocol outlines the methodological process for performing a review update of Roseman et al. (2017) [51].

Our systematic review update will be guided by the following key questions (KQ):

KQ1: What are the benefits and harms of screening for depression in children (6 to 11 years old) and adolescents (12 to 17 years old) in primary care or other non-mental health clinic settings?

KQ1a: What are the benefits and harms of screening for depression in children (6 to 11 years old) and adolescents (12 to 17 years old) in primary care or other non-mental health clinic settings for patients targeted because they have characteristics that may suggest elevated risk of depression?

Figure 1 presents an analytic framework of screening for depression in children and adolescents in primary care and the relevant health outcomes.

After reviewing the evidence pertinent to KQ1 and KQ1a, the working group may consider a separate systematic review on the patient values and preferences to inform their guideline beyond what is known from their usual patient engagement activities described in the following sections. The decision to undertake a systematic review on values and preferences will be directed by the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) methods and will weigh all a priori identified critical outcomes for both benefits and harms of screening. If the decision is made to pursue a values and preferences review on KQ2 and KQ2a, a separate protocol will be developed at that time that would include topic refinement with the working group. The potential key questions are:

KQ2. In children and adolescents without characteristics that may suggest elevated risk of depression, how do they (or their parents/guardians) value outcomes that may occur from screening for depression, and how do these values influence decisions about being screened?

KQ2a. In children and adolescents with characteristics that may suggest elevated risk of depression, how do they (or their parents/guardians) value outcomes that may occur from screening for depression in children and adolescents, and how do these values influence decisions about being screened?
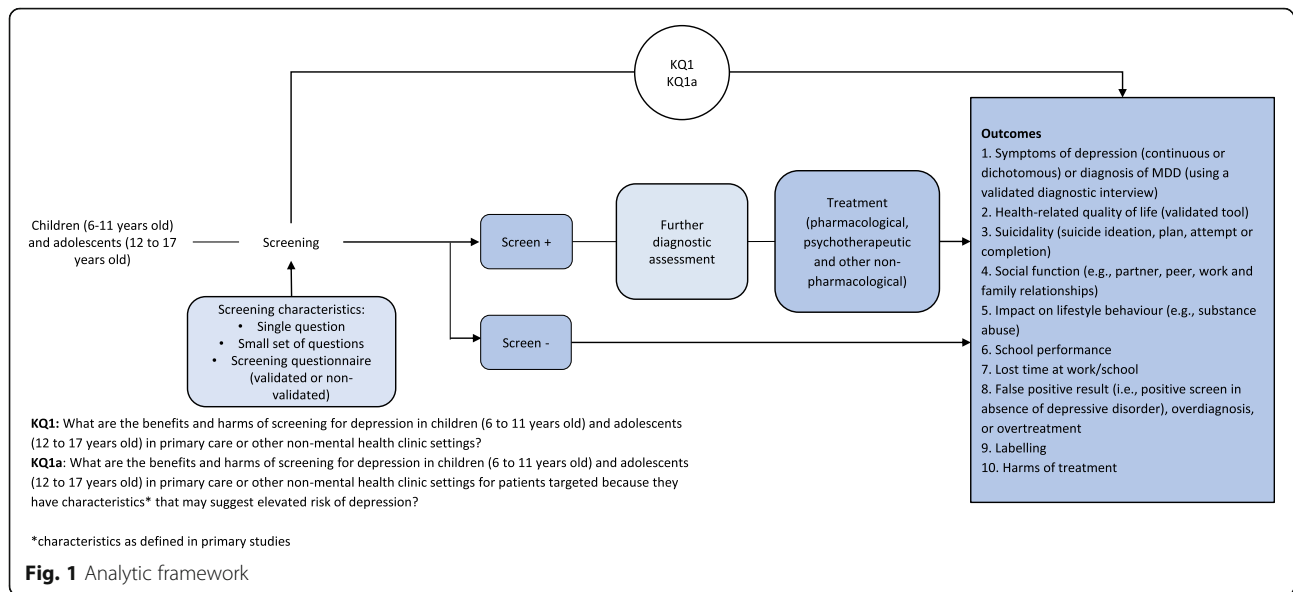
Methods

The review will be completed by the ERSC at the Ottawa Hospital Research Institute. Internal methodological experts, clinical experts, and a patient advisor will be consulted at key points during the conduct of the review. The systematic review will be developed, conducted, and prepared according to the Task Force Procedure Manual [52].

Protocol and registration

This protocol was developed by the ERSC and reviewed by their clinical experts and patient representative, the Depression working group members and their external

Beck *et al. Systematic Reviews* (2021) 10:24

Page 5 of 13

KQ1: What are the benefits and harms of screening for depression in children (6 to 11 years old) and adolescents (12 to 17 years old) in primary care or other non-mental health clinic settings?
KQ1a: What are the benefits and harms of screening for depression in children (6 to 11 years old) and adolescents (12 to 17 years old) in primary care or other non-mental health clinic settings for patients targeted because they have characteristics* that may suggest elevated risk of depression?

*characteristics as defined in primary studies

Fig. 1 Analytic framework

clinical experts, the PHAC, Task Force members, and stakeholder organizations.

The protocol was prepared following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA-P checklist) [53] (Additional file 2). The project is registered with the International Prospective Registry of Systematic Reviews (PROSPERO) database (CRD42020150373). The final review will be reported using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist and the A Measurement Tool to Assess Systematic Reviews (AMSTAR 2) tool will be used for additional quality control [54, 55]. We will document any amendments to this study protocol in PROSPERO and present them in the final manuscript.

Eligibility criteria

Studies for the review will be selected for inclusion according to the eligibility criteria presented in Table 1. Randomized controlled trials (RCTs) examining depression screening tools among patients who are up to 17 years of age will be eligible. If no or only a single RCT is available, controlled studies without random assignment will be sought. Studies that include other ages (e.g., 6 to 20 years of age), but present results separately for those up to 17 will be included and those that do not present results separately for the relevant ages will be excluded. Children will be defined as 6 to 11 years of age and adolescents will be defined as 12 to 17 years of age. For KQ1a, the target population will include patients who are selected for screening because they have characteristics that suggest elevated risk of depression. In all cases, we will include studies where caregivers respond to screening tool questions on behalf of children.

The depression screening tool can be a single question, a small set of questions, or a screening questionnaire (validated or not) with a pre-defined cut-off score to identify patients who may have depression but who have not reported their symptoms to healthcare providers or who have otherwise not been identified as possibly depressed by healthcare providers. The eligible comparator will be no depression screening. Both the screening and no screening arms will offer the same interventions for patients identified with depression symptoms. We will consider "no screening" to reflect normal clinical care that requires ongoing active observation and conversation with patients around symptoms suggestive of depression but no use of a formalized screening instrument. Depression screening should involve administering a depression symptom questionnaire to all patients and using a pre-identified cut-off threshold to classify patients as having positive or negative screening results. Accordingly, the use of a screening tool without the consideration of a pre-defined cut-off threshold would not be considered screening. Those with positive screening results can then be evaluated to determine if they have depression and, if appropriate, be offered treatment.

Studies in primary care or non-mental health clinic settings (e.g., medical specialist clinics, schools, or recreational/community settings) and online settings where screening is administered by a health practitioner will be eligible. Since this is a review update from Roseman et al. [51], we will continue to limit to RCTs and include reports from January 2017 onwards. If non-randomized controlled studies are sought, we will search from January 2015 onwards and replace the RCT study filter with a non-randomized controlled study design filter. The

Table 1 Inclusion and exclusion criteria for key question 1 and 1a

	Inclusion	Exclusion
Population	KQ1: Patients who are up to and including 17 years of age. KQ1a: Patients who are up to and including 17 years of age selected for screening because they have characteristics that may suggest elevated risk of depression. For both key questions, caregivers may respond to screening questions on behalf of children. Onset of adolescence will be considered as being age 12. *Characteristics as defined in primary studies*	• > 20% of the study sample are adults (18 years and older), have a recent history of depression, have a current diagnosis, or are receiving treatment for depression or other mental disorders (unless results are provided separately from the sample of interest) • Members of the study sample are seeking services due to symptoms of mental disorders • Members of the study sample are receiving assessment or care in psychiatric or mental health settings • Members of the study sample are currently pregnant or have given birth in the past year
Intervention	Screening tools that use a single question, a small set of questions, or a screening questionnaire (validated or non-validated) with a pre-defined cut-off score to identify patients who may have depression, but who have not previously reported their symptoms to healthcare providers or who have otherwise not been identified as possibly depressed by healthcare providers. *Patients and/or their guardians have the ability to answer screening tool questions*	Screening tools that, in addition to screening, include depression care referral or treatment options not available to patients identified as depressed in the non-screening trial arm
Comparator	No depression screening. • Patients in comparator arms may be administered depression symptom questionnaires for the purpose of baseline or outcome assessments as long as scores are not provided to the patients or healthcare providers prior to start of intervention.	
Outcomes	<u>Critical</u> 1. Symptoms of depression (measured continuously or dichotomously) or diagnosis of MDD (using a validated diagnostic interview) 2. Health-related quality of life (validated tool) 3. Suicidality (suicide ideation, plan, attempt or completion) 4. Social function (e.g., partner, peer, work and family relationships) 5. Impact on lifestyle behavior (e.g., substance abuse) <u>Important</u> 6. School performance 7. Lost time at work/school 8. False-positive result (i.e., positive screen in absence of depressive disorder), overdiagnosis, or overtreatment 9. Labeling 10. Harms of treatment	
Setting	Primary care or non-mental health clinic settings, such as medical specialist clinics, schools or recreational/community settings, and online settings (e.g., online depression screening), where screening is administered by a health practitioner	Studies conducted in mental health, or psychiatric settings. Studies in non-mental health clinic settings where screening is administered by a non-health practitioner.
Study design	Randomized controlled trials (RCTs), including cluster-randomized trials[a] If no or only a single RCT is available, then controlled studies without random assignment[b]	RCTs where patient eligibility is determined, and patients are enrolled after randomization. Interrupted times series, single cohort studies, case-control studies, cross-sectional studies, case series, case reports, and other publication types (editorials, commentaries, notes, letter, opinions).
Publication language	English or French	Languages other than English and French.
Dates of publication	January 2017 to present (RCTs) January 2015 to present (non-randomized controlled studies)	

[a]Eligible RCTs will need to meet the following criteria: wherein patient eligibility is determined and then patients are enrolled prior to randomization (i.e., to screening or to no screening); similar resources for depression management and treatment must be available both to patients in the screening arm of the trial and to patients in the non-screening arm of the trial who are identified as depressed via other methods (e.g., unaided clinician diagnosis, patient report) [43]
[b]Eligible nonrandomized controlled studies will need to meet the following criteria: Similar resources for depression management and treatment must be available both to patients in the screening arm of the trial and to patients in the non-screening arm of the trial who are identified as depressed via other methods (e.g., unaided clinician diagnosis, patient report) [43]

search date will be based on the latest search from the 2016 USPSTF review (February 2015), which did not find any non-randomized controlled studies assessing the effects of screening [48]. Their search would be difficult to replicate as it used one search for multiple key questions that were not relevant for this project. However, the USPSTF had similar eligibility criteria to ours for their key question on the effectiveness of

screening and we will consider their list of excluded studies to ensure all studies of interest are captured. Reports will be restricted to those written in English and French. Previous literature suggests including other languages when reviewing topics such as complementary and alternative medicines interventions [56]. We will use the following inclusion criteria from a recent depression screening systematic review to determine eligible RCTs and controlled studies without random assignment, where possible [43, 57]: (1) the patient population is clearly defined and randomized prior to administering the screening test (only applies to RCTs); (2) patients known to have a diagnosis of depression, or who are being treated for depression are excluded (screening is intended to identify undetected cases, and those who are known to have depression should not be screened in clinical practice); and (3) similar resources for depression management and treatment must be available both to patients in the screening arm of the trial and to patients in the non-screening arm of the trial who are identified as depressed via other methods (e.g., unaided clinician diagnosis, patient report).

Outcome rating
Members of the Depression working group developed the list of preliminary outcomes of interest, presented in Table 1. Through consensus, those outcomes were rated by the six working group members according to GRADE methodology as critical (rated 7 to 9 out of 9), important (rated 4 to 6 out of 9), or of limited importance (rated 1 to 3 out of 9) for making guideline recommendations [58]; only critical and important outcomes will be included. These outcomes will also be rated, according to GRADE methodology, by relevant groups of patients who are unaware of the working group ratings through the Task Force patient engagement activities conducted by the Knowledge Translation Program at St. Michael's Hospital in Toronto, Ontario. The list of outcomes will be finalized before final study selection and data extraction, and the PROSPERO record will be updated, as required. For more details on the outcome rating process, please see https://canadiantaskforce.ca/methods/ for the working group outcome rating and https://canadiantaskforce.ca/methods/patient-preferences-protocol/ for rating through patient engagement.

Outcomes will not be used in the selection of studies. Based on ratings to date by the six working group members, outcomes of interest deemed of critical importance for guideline development and decision-making are the symptoms of depression or diagnosis of MDD (average rating of 8.7), suicidality (average rating of 7.5), health-related quality of life (average rating of 6.8), social function (average rating of 6.5), and impacts on lifestyle behavior such as substance use (average rating of 6.5).

Outcomes rated as important are school performance (average rating of 6.2), harms of treatment (average rating of 6.2), lost time at work or school (average rating of 5.5), false-positive result, overdiagnosis, or overtreatment (average rating 5.5), and labeling (average rating of 4.7). No outcomes were rated as limited importance. The outcomes are listed in Table 1.

Information sources and search strategy
The search strategies were developed and tested through an iterative process by an experienced medical information specialist in consultation with the review team. Using the OVID platform, we will search Ovid MEDLINE®, including Epub Ahead of Print and In-Process & Other Non-Indexed Citations, Embase Classic+Embase, and PsycINFO. We will also search the Cumulative Index to Nursing and Allied Health Literature (CINAHL) (Ebsco) and CENTRAL (Wiley). The searches will utilize a combination of controlled vocabulary (e.g., "Depressive Disorder," "Mass Screening," "Adolescent") and keywords (e.g., MDD, screen, pediatric). We will apply an amended version of the 2008 Cochrane Highly Sensitive Search Strategy, sensitivity- and precision-maximizing version, to identify randomized controlled trials, and an additional filter to identify non-randomized controlled studies. Vocabulary and syntax will be adjusted across the databases. There will be no language restrictions, but results will be limited to the period of 2017 to the present for RCTs and 2015 for non-randomized controlled studies. Animal-only and opinion pieces will be removed from the results.

The draft search strategy for the MEDLINE database was peer-reviewed by another experienced librarian using the Peer Review of Electronic Search Strategies (PRESS) checklist [59], modified where necessary, and then translated to other databases. The draft strategies are available in Additional file 3.

We will search grey literature sources to identify studies not formally published in journals and any ongoing trials using the Canadian Agency for Drugs and Technologies in Health (CADTH) Grey Matters checklist [60]. The CADTH checklist includes public health websites, health technology assessment agencies, clinical practice guideline organizations, clinical trials registries, search engines, and additional databases. We will also search the websites of the following organizations: American Academy of Child and Adolescent Psychiatry, American Academy of Family Physicians, American Academy of Pediatrics, American College of Physicians, American Nurses Association, American Psychological Association, Anxiety and Depression Association of America, Canadian Academy of Child and Adolescent Psychiatry, Canadian Mental Health Association, Canadian Nurses Association, Canadian Paediatric Society,

Canadian Psychiatric Association, Centre for Addiction and Mental Health, Child Mind Institute, and the College of Family Physicians of Canada. Searches for grey literature will be restricted to English and French language documents and limited to a 5-day maximum search period by one person. The reference list of included studies and secondary evidence reports (e.g., evidence-based clinical practice guidelines, systematic reviews, and meta-analyses) will also be reviewed. Using Robinson et al. (2014) as guidance [61], a review would need to meet the following criteria to be considered systematic: (1) At least one database was searched; (2) authors report selection criteria; (3) risk of bias (or intended analogous) of included studies is reported; and (4) authors report a list and synthesis of included studies.

Study selection
Citations from bibliographic databases, grey literature, and supplemental source searching will be collated in Reference Manager [62]. Those citations will be de-duplicated and uploaded to an online systematic review management software package, DistillerSR [63]. A pilot test will be conducted by two reviewers for title and abstract (random sample of 50 citations) and full-text levels (random sample of 25 articles) to test the screening questions and to ensure consistent application of the selection criteria among reviewers. Adjustments to the forms will be made as needed.

Title and abstract screening will be performed by two reviewers using the liberal accelerated method [64]. This approach involves one reviewer screening citations and a second reviewer verifying the citations excluded by the first reviewer. References will be screened in random order, and screening will be done concurrently among reviewers to reduce the likelihood that a citation has been already considered by another reviewer. Conflict resolution will not be conducted at this stage.

At the full-text screening stage, two reviewers will independently and in duplicate review the full text of potentially relevant articles to determine eligibility. Conflicts will be resolved by consensus or by consulting with a senior team member. Reviewers will agree on the reasons for exclusion at the full-text stage.

Where study eligibility is unclear, authors will be contacted for additional information twice by email over three weeks. Where input is needed for potentially eligible studies, we will consult with working group members and clinical experts who are blinded to study identification information and outcome data. Articles that are not available electronically will be ordered via interlibrary loan, allowing 30 days for receipt. Potentially relevant studies available only in abstract form will be

excluded and the citation will be annotated in the list of excluded studies. To reduce potential bias, we will consult with working group members and clinical experts on any potentially missing studies once all outstanding queries related to synthesizing the evidence are addressed. Draft screening forms are available in Additional file 4.

Data extraction
Before commencing data extraction, we will pilot the data extraction form in DistillerSR on a random sample of full-text articles (three to six articles, depending on the number of included studies) and modify the form as needed. Two reviewers will independently extract data, compare data sets, discuss, and resolve any conflicts through discussion. If consensus cannot be reached, we will consult with a senior team member. Study authors will be contacted by email twice over 3 weeks for further information as required. We will extract all formats of continuous outcome data whether reported as post-intervention or change from baseline. Draft items for data extraction are provided in Additional file 5, and clinical experts and the working group will be consulted on extraction variables.

Risk of bias in individual studies
Concurrently with data extraction, we will follow Task Force methods and will appraise the risk of bias (ROB) using the Cochrane ROB tool for RCTs [65] and the Newcastle-Ottawa Scale (NOS) for non-randomized controlled studies [66]. Two reviewers will independently assess the ROB of all included studies, compare assessments, and resolve any assessments and their supports for judgments where disagreement has occurred. This will be done by consensus or by consulting with a senior team member when required. The Cochrane ROB tool consists of the following seven domains: (1) random sequence generation, (2) allocation concealment, (3) blinding of participants and personnel, (4) blinding of outcome assessment, (5) incomplete outcome data, (6) selective reporting, and (7) other sources of bias. As recommended by the Cochrane Handbook [67], certain domains are outcome-specific (e.g., blinding of outcome assessors) and will be assessed at the outcome level. The overall ROB will require the consideration of the relative importance of the seven domains accompanied by the known empirical evidence of bias, likely direction of bias, and likely magnitude of bias [65]. For outcome and analysis reporting bias, we will use the methods outlined in the Agency for Healthcare Research and Quality guidance [68]. For assessing cluster randomized trials, we plan to assess recruitment bias (when trial recruitment follows after clusters have been randomized and the knowledge of intervention and control clusters

could affect the types of participants recruited) in the "other risks of bias" domain of the Cochrane ROB tool [69]. We will classify the overall ROB as low if all domains were assessed at low ROB, high if at least one domain was assessed at high ROB, or unclear if at least one domain was assessed at unclear ROB and no domain was at high risk.

The NOS for cohort studies consists of the following three domains: (1) selection of study groups, (2) comparability of study groups, and (3) outcome assessment. Each item within the selection (4 items) and outcome (3 items) domains can be awarded a maximum of one point, and two points for the single item in the comparability domain. A study can be awarded a maximum of nine points for the least risk of bias [66]. We will classify good quality studies as scoring at least 3 stars in the selection domain, at least 1 star in comparability, and at least 2 stars in the outcome domain. A fair quality study will require 2 stars in selection, at least 1 star in comparability, and at least 2 stars in the outcome domain. A poor-quality study will score 0 or 1 star(s) in selection, or 0 stars in comparability, or 0 or 1 star(s) in the outcome domain.

The individual study ROB and NOS assessments will inform the study limitations domain assessment across studies for each outcome using GRADE assessments, described in the following sections [70].

Synthesis of included studies and interpretation

We will describe and present in summary tables the study characteristics, participant characteristics, intervention and comparator details, outcome results, and Cochrane ROB or NOS evaluations. We will convert data (e.g., recovery of intervention group standard deviations from standard errors or 95% confidence intervals) to facilitate consistent presentation and synthesis of the results across studies. We will calculate the appropriate statistics based on the type of outcome data. Relative and absolute effects with 95% confidence intervals will be calculated to facilitate the presentation of outcome data according to the GRADE guidance for developing evidence profile tables and summary of findings. Risk ratios and risk differences will be used to report effects for binary data. When presenting continuous data, GRADE guidance will be used [71]. Where needed, we will use a conservative value for a correlation coefficient of 0.25 to impute standard deviations for means used in change from baseline calculations, as performed in other reviews [72]. Studies with a cluster design (e.g., allocation by clinical setting) with data analyzed at the individual level (e.g., patients) will provide incorrect estimates of precision if correlation within clusters is not taken into account [73]. Where authors of cluster randomized trials have not adjusted for unit-of-analysis errors, we will

attempt to correct these by using intra-class correlation coefficients (ICC) when available and by estimating ICCs when not [67, 74]. Where possible, the number needed to screen will be calculated.

Meta-analysis

Prior to performing a meta-analysis, we will assess the clinical (e.g., patient characteristics) and methodological (e.g., study design) heterogeneity of included studies. We will seek input from clinical experts and the guideline working group on decision-making.

The I^2 statistic will be used to assess the statistical heterogeneity across included studies (low [0–25%], moderate [25–50%], substantial [50–75%], and considerable [> 75%]) [75–77]. We will use the Cochran's Q test (threshold p value < 0.10) to interpret the strength of evidence for heterogeneity [78, 79]. If considerable heterogeneity is observed (defined as I^2 of > 75%), we will present studies in a forest plot, but will not report a pooled estimate. We will make efforts to explain statistical heterogeneity via the performance of subgroup analyses, sensitivity analyses, and/or meta-regression analyses; the optimal approach will be determined once determining what information is available. Meta-regression will be based on random effects models to allow for residual unexplained heterogeneity. We will follow previously published guidance for meta-regression [75]. The meta-regression analyses will be conducted when 10 or more studies are available for a continuous study-level variable. For a categorical subgroup variable, each subgroup should have four or more studies. These numbers serve as the lower bounds for consideration of meta-regression for small studies [75]. We will use univariable meta-regression when an insufficient number of studies is available to conduct multivariable analyses.

When a meta-analysis is appropriate (e.g., studies are similar and most do not have high ROB), we will use a random effects model for effect measures. Risk ratios and risk differences will be used for pooling binary outcomes. Mean differences (outcomes measure in a similar manner or with the same scale) or standardized mean differences (where outcomes measured using different scales) will be used for continuous data. We will not combine RCTs and non-randomized controlled studies in a meta-analysis. If a meta-analysis is not appropriate, we will present the range of effects.

Sparse binary data and studies with zero events

When studies report endpoints with low event rates, the results will be summarized narratively. The risk difference summary measure will be used for the outcomes where at least one intervention group contains zero events (e.g., suicide completion). We will use the median baseline risk for the control group, where available, for

risk difference calculations to enable reporting for GRADE tables. However, we may conduct a sensitivity analysis using different baseline risks if suitable.

Subgroup analysis

If there are sufficient data and adequate reporting, the following subgroup analyses are planned:

- Socioeconomic status of parent(s) (e.g., income, level of education; as assessed by study authors)
- Ethnic group (e.g., Indigenous peoples; will be determined post hoc, depending on populations encountered in studies)
- Geographical location (e.g., rural vs. urban settings, by country/region)
- Age groups
- Gender/sex (according to how reported by study authors)
- Chronic illness
- Immigrant status of child or parent (as reported by study authors)
- Risk factors for depression (to be determined post-hoc, depending on combination of risk factors as reported in studies)
- Type of setting

Sensitivity analyses

When appropriate, a sensitivity analysis will be conducted by restricting analyses to studies assessed as low overall ROB or good quality. Additional analyses may also be performed to explore statistical heterogeneity, address decisions made regarding the handling of data (e.g., multiple observations in a small percentage of patients), the timing of publication, study design issues, or data requiring computation.

Small-study effects

For outcomes reported in 10 or more studies in a meta-analysis, publication bias will be assessed using a combination of graphical aids (e.g., funnel plots) and/or statistical tests (e.g., Egger regression test, Hedges-Olkin) [77, 80].

Grading the certainty of evidence and interpretation

We will assess the certainty of evidence using the GRADE approach. We will perform separate GRADE assessments for RCTs and non-randomized controlled studies. GRADE evidence profiles and summary of findings tables will be prepared for each of the critical and important outcomes using the GRADE framework to assess each domain (i.e., study limitations, imprecision, inconsistency, indirectness, and other considerations (including publication bias)) [52, 58].

The ERSC will undertake initial assessments of the GRADE domains, except for imprecision, and format the evidence according to GRADE guidance. Members of the guideline working group will have the opportunity to review those draft GRADE tables to determine imprecision thresholds for interpreting the importance of absolute effect sizes observed for outcomes. The ERSC plans to finalize the GRADE assessments incorporating those thresholds. Given the evolution of GRADE guidance at the time of the writing of this protocol, the formulation of GRADE narrative statements to represent the quantity, magnitude, and certainty of the evidence will be informed by available guidance [71, 81, 82].

The GRADE assessments will be piloted among reviewers and a senior methodologist for a sample of three to six outcomes that, if possible, will demonstrate diversity among data/information sets. Following the pilot phase, two reviewers will independently and together assess the certainty of the evidence in GRADEpro GDT, compare GRADE table outputs, and resolve disagreements by discussion or consulting a senior team member.

Software

We will use Cochrane Review Manager 5.3 (RevMan) to calculate effect estimates and conduct meta-analyses, where needed [83]. For all analyses not possible in RevMan, we will use Comprehensive Meta-Analysis [84]. GRADE assessments (including piloting) and corresponding tables will be produced in GRADEpro GDT [85].

Discussion

The results of this review and the manuscript will be published in the "Canadian Task Force on Preventive Health Care Evidence Reviews" thematic series. Our systematic review will provide an evidence base for the Task Force and others to develop guidance on screening for depression in children and adolescents in primary care settings in Canada. We will address considerations for future research and highlight the implications for primary care practice.

Supplementary Information

Additional file 1:. DSM-5 and ICD-10 definition of MDE.docx (19 KB)

Additional file 2:. PRISMA-P 2015 checklist.docx (21 KB)

Additional file 3:. Search strategies.docx (27 KB)

Additional file 4:. Draft screening forms.docx (19 KB)

Additional file 5:. Draft data extraction form.docx (16 KB)

Additional file 6:. Stakeholder and peer review feedback.docx (24 KB)

Abbreviations

DSM-5: Diagnostic and statistical manual of mental disorders, 5th edition; ERSC: Evidence review and synthesis centre; GRADE: Grading of

Authors' contributions
AB: conceptualization, project administration, writing original draft and revisions. CH: project administration, methodology, supervision, writing original draft and revisions. BS: review and editing, search strategy. AS: review and editing, methodology, supervision. BH, BJS, BP, BS, DM, JL, and SN: review and editing, methodology. GG, HB, IC, KP, PV, RM, and SJ: review and editing. Members of the Depression Working Group for the Canadian Task Force on Preventive Health Care (AM, BDT, EL, HC, JCL, JR): conceptualization, review and editing. Public Health Agency of Canada Global Health and Guidelines Division (KM, EG): conceptualization, review, and editing. All authors read and approved the final manuscript.

Availability of data and materials
Not applicable.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Written informed consent to publish was obtained from the stakeholders and peer reviewers who provided feedback on the protocol. A copy of the written consent is available for review by the Editors-in-Chief of this journal. The stakeholder and peer reviewer feedback has been anonymized and included as Additional file 6.

Competing interests
DM is Co-Editor-in-Chief with *Systematic Reviews*. BH has previously received honoraria from Cornerstone Research Group for the provision of methodologic advice related to systematic reviews and meta-analysis. BDT has received grants from the Canadian Institutes of Health Research for the evaluation of the accuracy of depression screening tools and the effects of depression screening.

Author details
[1]Knowledge Synthesis Group, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Centre for Practice-Changing Research, 501 Smyth Road, Box 201, Ottawa, ON K1H 8L6, Canada. [2]Department of Pediatrics, Dalhousie University, Halifax, NS, Canada. [3]Public Health Agency of Canada, Ottawa, ON, Canada. [4]Department of Occupational Science and Occupational Therapy, University of Toronto, Toronto, ON, Canada. [5]University of Calgary Cumming School of Medicine, Calgary, AB, Canada. [6]Alberta Health Services, Calgary, AB, Canada. [7]Department of Health Research Methods, Evidence and Impact, McMaster University, Hamilton, ON, Canada. [8]Department of Family Medicine, McMaster University, David Braley Health Sciences Centre, Hamilton, ON, Canada. [9]Lady Davis Institute of the Jewish General Hospital, Montreal, QC, Canada. [10]Faculty of Medicine, McGill University, Montreal, QC, Canada. [11]Department of Community Health Services and Department of Psychiatry, University of Calgary, Calgary, AB, Canada. [12]Children's Hospital of Eastern Ontario, Ottawa, ON, Canada. [13]School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada. [14]Children's Hospital of Eastern Ontario Research Institute, Ottawa, ON, Canada. [15]Ottawa Hospital Research Institute, Ottawa, ON, Canada. [16]Department of Psychiatry, Children's Hospital of Eastern Ontario, Ottawa Faculty of Medicine, Ottawa, ON, Canada. [17]Waypoint Centre For Mental Health Care, Penetanguishene, ON, Canada. [18]Department of Family and Community Medicine, St. Michael's Hospital, University of Toronto, Toronto, ON, Canada.

References
1. Otte C, Gold SM, Penninx BW, Pariante CM, Etkin A, Fava M, et al. Major depressive disorder. Nat Rev Dis Primers. 2016 Sep 15;2:16065.
2. Public Health Agency of Canada (PHAC). A report on mental illnesses in Canada Chapter 2: mood disorders [Internet]. 2002 [cited 2019 Apr 10]. Available from: https://www.canada.ca/content/dam/phac-aspc/migration/phac-aspc/publicat/miic-mmac/pdf/chap_2_e.pdf.
3. Costello EJ, Copeland W, Angold A. Trends in psychopathology across the adolescent years: What changes when children become adolescents, and when adolescents become adults? J Child Psychol Psychiatry. 2011;52(10):1015–25.
4. Patton GC, Viner R. Pubertal transitions in health. Lancet. 200731;369(9567): 1130–9.
5. Eiland L, Romeo RD. Stress and the developing adolescent brain. Neuroscience. 2013;249:162–71.
6. American Psychiatric Assoiation. Diagnostic and statistical manual of mental disorders. 5th ed. Washington, DC: American Psychiatric Publishing; 2013.
7. World Health Organization (WHO). ICD-10: international statistical classification of diseases and related health problems Version 2016. Genève: World Health Organization.
8. Georgiades K, Duncan L, Wang L, Comeau J, Boyle MH. Six-Month Prevalence of Mental Disorders and Service Contacts among Children and Youth in Ontario: Evidence from the 2014 Ontario Child Health Study. Can J Psychiatry. 2019;64(4):246–55.
9. Wiens K, Williams JVA, Lavorato DH, Duffy A, Pringsheim TM, Sajobi TT, et al. Is the prevalence of major depression increasing in the Canadian adolescent population? Assessing trends from 2000 to 2014. J Affect Disord. 2017;210:22–6.
10. Hankin BL, Abramson LY, Moffitt TE, Silva PA, McGee R, Angell KE. Development of depression from preadolescence to young adulthood: emerging gender differences in a 10-year longitudinal study. J Abnorm Psychol. 1998;107(1):128–40.
11. Wade TJ, Cairney J, Pevalin DJ. Emergence of gender differences in depression during adolescence: national panel results from three countries. J Am Acad Child Adolesc Psychiatry. 2002;41(2):190–8.
12. Patten SB, Williams JVA, Lavorato DH, Wang JL, Bulloch AGM, Sajobi T. The association between major depression prevalence and sex becomes weaker with age. Soc Psychiatry Psychiatr Epidemiol. 2016;51(2):203–10.
13. Gore FM, Bloem PJ, Patton GC, Ferguson J, Joseph V, Coffey C, et al. Global burden of disease in young people aged 10–24 years: a systematic analysis. Lancet. 2011;377(9783):2093–102.
14. World Health Organization (WHO). Adolescent health epidemiology. Geneva: World Health Organization; 2014. https://www.who.int/maternal_child_adolescent/epidemiology/adolescence/en/. Accessed 12 Apr 2019.
15. Clayborne ZM, Varin M, Colman I. Systematic review and meta-analysis: adolescent depression and long-term psychosocial outcomes. J Am Acad Child Adolesc Psychiatry. 2019;58(1):72–9.
16. Johnson D, Dupuis G, Piche J, Clayborne Z, Colman I. Adult mental health outcomes of adolescent depression: A systematic review. Depression and Anxiety. 2018;35(8):700–16.
17. Chisholm D, Sanderson K, Ayuso-Mateos JL, Saxena S. Reducing the global burden of depression: population-level analysis of intervention cost-effectiveness in 14 world regions. Br J Psychiatry. 2004;184:393–403.
18. Bodden DHM, Stikkelbroek Y, Dirksen CD. Societal burden of adolescent depression, an overview and cost-of-illness study. J Affect Disord. 2018;241: 256–62.

19. Zisook S, Lesser I, Stewart JW, Wisniewski SR, Balasubramani GK, Fava M, et al. Effect of age at onset on the course of major depressive disorder. Am J Psychiatry. 2007;164(10):1539–46.

20. Rohde P, Lewinsohn PM, Klein DN, Seeley JR, Gau JM. Key characteristics of major depressive disorder occurring in childhood, adolescence, emerging adulthood, adulthood. Clin Psychol Sci. 2013;1(1):41–53.

21. Weissman MM, Wolk S, Goldstein RB, Moreau D, Adams P, Greenwald S, et al. Depressed Adolescents Grown Up. JAMA. 1999;281(18):1707–13.

22. Patten SB, Wang JL, Williams JV, Currie S, Beck CA, Maxwell CJ, et al. Descriptive epidemiology of major depression in Canada. Can J Psychiatry. 2006;51(2):84–90.

23. Thapar A, Collishaw S, Pine DS, Thapar AK. Depression in adolescence. Lancet. 2012;379(9820):1056–67.

24. Oldehinkel AJ, Ormel J. A longitudinal perspective on childhood adversities and onset risk of various psychiatric disorders. Eur Child Adolesc Psychiatry. 2015;24(6):641–50.

25. Yap MBH, Pilkington PD, Ryan SM, Jorm AF. Parental factors associated with depression and anxiety in young people: A systematic review and meta-analysis. J Affect Disord. 2014;156:8–23.

26. Dooley B, Fitzgerald A, Giollabhui NM. The risk and protective factors associated with depression and anxiety in a national sample of Irish adolescents. Irish J Psychol Med. 2015;32(1):93–105.

27. Kidger J, Araya R, Donovan J, Gunnell D. The effect of the school environment on the emotional health of adolescents: a systematic review. Pediatrics. 2012;129(5):925–49.

28. Riglin L, Petrides KV, Frederickson N, Rice F. The relationship between emotional problems and subsequent school attainment: A meta-analysis. J Adolesc. 2014;37(4):335–46.

29. Hendren RL, Haft SL, Black JM, White NC, Hoeft F. Recognizing psychiatric comorbidity with reading disorders. Front Psychiatry. 2018;9. [cited 2020 Apr 23] Available from: https://www.frontiersin.org/articles/https://doi.org/10.3389/fpsyt.2018.00101/full#B1.

30. Meinzer MC, Pettit JW, Viswesvaran C. The co-occurrence of attention-deficit/hyperactivity disorder and unipolar depression in children and adolescents: A meta-analytic review. Clin Psychol Rev. 2014;34(8):595–607.

31. Gundel LK, Pedersen CB, Munk-Olsen T, Dalsgaard S. Longitudinal association between mental disorders in childhood and subsequent depression – A nationwide prospective cohort study. J Affect Disord. 2018;227:56–64.

32. Stirling K, Toumbourou JW, Rowland B. Community factors influencing child and adolescent depression: a systematic review and meta-analysis. Aust N Z J Psychiatry. 2015;49(10):869–86.

33. Ceairns KE, Yap MBH, Pilkington PD, Jorm AF. Risk and protective factors for depression that adolescents can modify: A systematic review and meta-analysis of longitudinal studies. J Affect Disord. 2014;169:61–75.

34. Gunnell KE, Flament MF, Buchholz A, Henderson KA, Obeid N, Schubert N, et al. Examining the bidirectional relationship between physical activity, screen time, and symptoms of anxiety and depression over time during adolescence. Prev Med. 2016;88:147–52.

35. Maras D, Flament MF, Murray M, Buchholz A, Henderson KA, Obeid N, et al. Screen time is associated with depression and anxiety in Canadian youth. Prev Med. 2015;73:133–8.

36. Thombs BD, Saadat N, Riehm KE, Karter JM, Vaswani A, Andrews BK, et al. Consistency and sources of divergence in recommendations on screening with questionnaires for presently experienced health problems or symptoms: a comparison of recommendations from the Canadian Task Force on Preventive Health Care, UK National Screening Committee, and US Preventive Services Task Force. BMC Med. 2017;15(1):150.

37. Cosgrove L, Karter JM, Vaswani A, Thombs BD. Unexamined assumptions and unintended consequences of routine screening for depression. J Psychosom Res. 2018;109:9–11.

38. Cipriani A, Zhou X, Del Giovane C, Hetrick SE, Qin B, Whittington C, et al. Comparative efficacy and tolerability of antidepressants for major depressive disorder in children and adolescents: a network meta-analysis. Lancet. 2016;388(10047):881–90.

39. Hetrick SE, McKenzie JE, Cox GR, Simmons MB, Merry SN. Newer generation antidepressants for depressive disorders in children and adolescents. Cochrane Database Syst Rev. 2012;(11). [cited 2020 Apr 23] Available from: https://www.cochranelibrary.com/cdsr/doi/https://doi.org/10.1002/14651858.CD004851.pub3/abstract.

40. Weisz JR, McCarty CA, Valeri SM. Effects of Psychotherapy for Depression in Children and Adolescents: A Meta-Analysis. Psychol Bull. 2006;132(1):132–49.

41. Dubicka B, Elvins R, Roberts C, Chick G, Wilkinson P, Goodyer IM. Combined treatment with cognitive-behavioural therapy in adolescent depression: meta-analysis. Br J Psychiatry. 2010;197(6):433–40.

42. Thombs BD, Ziegelstein RC, Roseman M, Kloda LA, Ioannidis JP. There are no randomized controlled trials that support the United States Preventive Services Task Force guideline on screening for depression in primary care: a systematic review. BMC Med. 2014;12(1):13.

43. Thombs BD, Arthurs E, Coronado-Montoya S, Roseman M, Delisle VC, Leavens A, et al. Depression screening and patient outcomes in pregnancy or postpartum: a systematic review. J Psychosom Res. 2014;76(6):433–46.

44. Goodyear-Smith FA, van Driel ML, Arroll B, Del Mar C. Analysis of decisions made in meta-analyses of depression screening and the risk of confirmation bias: a case study. BMC Med Res Methodol. 2012;12:76.

45. Bland RC, Streiner DL. Why screening for depression in primary care is impractical. CMAJ. 2013;185(9):753–4.

46. MacMillan HL, Patterson CJS, Wathen CN. Screening for depression in primary care: recommendation statement from the Canadian Task Force on Preventive Health Care. CMAJ. 2005;172(1):33–5.

47. Siu AL, on behalf of the U.S. Preventive services task force. screening for depression in children and adolescents: U.S. preventive services task force recommendation statement. Ann Intern Med. 2016;164(5):360.

48. Forman-Hoffman V, McClure E, McKeeman J, Wood CT, Middleton JC, Skinner AC, et al. Screening for major depressive disorder in children and adolescents: a systematic review for the U.S. preventive services task force. Ann Intern Med. 2016;164(5):342.

49. Zuckerbrot RA, Cheung AH, Jensen PS, Stein REK, Laraque D. Guidelines for adolescent depression in primary care (GLAD-PC): I. identification, assessment, and initial management. Pediatrics. 2007;120(5):e1299–312.

50. Zuckerbrot RA, Cheung A, Jensen PS, REK S, Laraque D, Group G-PS. Guidelines for adolescent depression in primary care (GLAD-PC): Part I. practice preparation, identification, assessment, and initial management. Pediatrics. 2018;141(3):e20174081.

51. Roseman M, Saadat N, Riehm KE, Kloda LA, Boruff J, Ickowicz A, et al. Depression screening and health outcomes in children and adolescents: a systematic review. Can J Psychiatry. 2017;62(12):813–7.

52. Canadian Task Force on Preventive Health Care. Canadian Task Force on Preventive Health Care Procedure Manual [Internet]. 2014. Available from: https://canadiantaskforce.ca/methods/.

53. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015;349:g7647.

54. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLOS Med. 2009;6(7):e1000100.

55. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. Bmj. 2017;358:j4008.

56. Moher D, Pham B, Lawson ML, Klassen TP. The inclusion of reports of randomised trials published in languages other than English in systematic reviews. Health Technol Assess. 2003;7(41):1–90.

57. Hamel C, Lang E, Morissette K, Beck A, Stevens A, Skidmore B, et al. Screening for depression in women during pregnancy or the first year postpartum and in the general adult population: a protocol for two systematic reviews to update a guideline of the Canadian task force on preventive health care. Syst Rev. 2019;8(1):27.

58. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, et al. Grading quality of evidence and strength of recommendations. BMJ. 2004;328(7454):1490.

59. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol. 2016;75:40–6.

60. CADTH. Grey Matters: a practical tool for searching health-related grey literature [Internet]. 2018 [cited 2019 Apr 25]. Available from: https://www.cadth.ca/resources/finding-evidence.

61. Robinson KA, Whitlock EP, Oneil ME, Anderson JK, Hartling L, Dryden DM, et al. Integration of existing systematic reviews into new reviews: identification of guidance needs. Syst Rev. 2014;3:60.

62. Reference Manager 12. New York: Thomson Reuters; 2011.

Beck *et al. Systematic Reviews*   (2021) 10:24

Page 13 of 13

63. Evidence Partners. DistillerSR [Internet]. Ottawa, Canada. https://www.evidencepartners.com/; 2011. Available from: https://www.evidencepartners.com/.

64. Khangura S, Konnyu K, Cushman R, Grimshaw J, Moher D. Evidence summaries: the evolution of a rapid review approach. Syst Rev. 2012;1:10.

65. Higgins JPT, Altman DG, Sterne JA. Chapter 8: Assessing risk of bias in included studies. In: Cochrane Handbook for Systematic Reviews of Interventions [Internet]. Version 5.1.0. The Cochrane Collaboration; 2011. Available from: www.handbook.cochrane.org.

66. Wells G, Shea BJ, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa: Ottawa Hospital Research Institute; 2011. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed 20 Nov 2019.

67. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. [Internet] http://handbook.cochrane.org/. 2011.

68. Balshem H, Stevens A, Ansari M, Norris S, Kansagara D, Shamliyan T, et al. Finding grey literature evidence and assessing for outcome and analysis reporting biases when comparing medical interventions: AHRQ and the effective health care program. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2013 [cited 2018 Jun 20]. (AHRQ Methods for Effective Health Care). Available from: http://www.ncbi.nlm.nih.gov/books/NBK174882/.

69. Higgins J, Green S. Chapter 16: Special topics in statistics. In: The Cochrane Collaboration Cochrane Handbook for Systematic Reviews of Interventions,. 2011.

70. Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol. 2011; 64(1878-5921 (Electronic)):401–6.

71. Guyatt GH, Thorlund K, Oxman AD, Walter SD, Patrick D, Furukawa TA, et al. GRADE guidelines: 13. Preparing summary of findings tables and evidence profiles—continuous outcomes. J Clin Epidemiol. 2013;66(2):173–83.

72. Hong CJ, Tsang AC, Quinn JG, Bonaparte JP, Stevens A, Kilty SJ. Anti-IgE monoclonal antibody therapy for the treatment of chronic rhinosinusitis: a systematic review. Syst Rev. 2015;4:166.

73. Whiting-O'Keefe QE, Henke C, Simborg DW. Choosing the correct unit of analysis in medical care experiments. Med Care. 1984;22(12):1101–14.

74. White IR, Thomas J. Standardized mean differences in individually-randomized and cluster-randomized trials, with applications to meta-analysis. Clin Trials. 2005;2(2):141–51.

75. Fu R, Gartlehner G, Grant M, Shamliyan T, Sedrakyan A, Wilt TJ, et al. Conducting quantitative synthesis when comparing medical interventions: AHRQ and the effective health care program. J Clin Epidemiol. 2011; 64(1878-5921 (Electronic)):1187–97.

76. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. Br Med J. 2003;327(7414):557.

77. Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. Bmj. 2011;343:d4002.

78. Deeks JJ, Higgins JP, Altman DG (editors). Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions. 2nd Edition. Chichester (UK): Wiley; 2019.

79. Sutton AJ. Methods for meta-analysis in medical research. Chichester: Wiley; 2000.

80. Sterne JAC, Egger M, Smith GD. Investigating and dealing with publication and other biases in meta-analysis. BMJ. 2001;323(7304):101–5.

81. Santesso N, Glenton C, Dahm P, Garner P, Akl EA, Alper B, et al. GRADE guidelines 26: informative statements to communicate the findings of systematic reviews of interventions. J Clin Epidemiol. 2020;119:126–35.

82. Cochrane Effective Practice and Organisation of Care (EPOC). Reporting the effects of an intervention in EPOC reviews. EPOC Resources for review authors. [Internet]. 2018. Available from: https://epoc.cochrane.org/sites/epoc.cochrane.org/files/public/uploads/Resources-for-authors2017/how_to_report_the_effects_of_an_intervention.pdf.

83. The Cochrane Collaboration. Review Manager (RevMan). Copenhagen. 2015: The Nordic Cochrane Centre;.

84. Biostat. Comprehensive Meta-Analysis [Software]. Englewood: Biostat; 2014. https://www.meta-analysis.com/.

85. Evidence Prime. GRADEpro GDT. GRADEpro Guideline Development Tool [Software]. Hamilton: McMaster University; 2020. https://gradepro.org/.

Publisher's Note