

SUPPLEMENTARY FILE 1: Search Strategies (through September 30, 2015)

MEDLINE

1. Depression/
2. Depressive Disorder/
3. Depressive Disorder, Major/
4. depress*.ti.
5. or/1-4
6. exp Psychological Tests/
7. exp Psychiatric Status Rating Scales/
8. *Questionnaires/
9. Self-report/
10. Mass Screening/
11. case finding*.ti,ab.
12. casefinding*.ti,ab.
13. ((screen* or evaluat* or assess* or diagnos* or rating or rate or measur*) adj5 (index or indices or score* or scoring or test* or instrument* or inventor* or battery or batteries or tool* or scale* or checklist* or schedule*)).ti,ab.
14. ("mood and feelings questionnaire*" or MFQ or SMFQ or MFQ-C).ti,ab.
15. reynold* child* depression.ti,ab.
16. Reynold* adolesc* depression.ti,ab.
17. kutcher* adolesc*.ti,ab.
18. "depression scale for child*".ti,ab.
19. (Children's Depression Inventory or CDI).ti,ab.
20. (Brief Patient Health Questionnaire-9 or PHQ-9).ti,ab.
21. (Brief Patient Health Questionnaire-2 or PHQ-2).ti,ab.
22. (Brief Patient Health Questionnaire-A or PHQ-A).ti,ab.
23. (Center for Epidemiologic Studies Depression Scale or CES-D).ti,ab.
24. (Beck Depression Inventory or BDI or BDI-PC).ti,ab.
25. (Hopkins Symptom Checklist or HSCL or HSCL-10).ti,ab.
26. ("Strengths and Difficulties Questionnaire" or SDQ).ti,ab.
27. (Youth Self Report or YSR).ti,ab.
28. DesTeen.ti,ab.
29. or/6-28
30. 5 and 29
31. limit 30 to validation studies
32. (valid* or predict* or identif* or detect*).ti.
33. "Sensitivity and Specificity"/
34. "Predictive Value of Tests"/
35. (cut?off* or accura* or gold standard or reference standard).ab.
36. or/32-35
37. 30 and 36
38. 31 or 37
39. exp Child/
40. child.mp.
41. exp Pediatrics/

42. pediatric*.mp.
43. paediatric*.mp.
44. (boy or boys).mp.
45. girl*.mp.
46. (kid or kids).mp.
47. school?age*.mp.
48. juvenil*.mp.
49. under?age*.mp.
50. teen*.mp.
51. minor*.mp.
52. pubescen*.mp.
53. adolescen*.mp.
54. (youth or youths).mp.
55. child*.jw.
56. pediatric*.jw.
57. paediatric*.jw.
58. adolescen*.jw.
59. or/39-58
60. 38 and 59
61. limit 60 to yr="2006 -Current"
62. Remove duplicates from 61
63. (201409* or 20141* or 2015*).ed.
64. 62 and 63

MEDLINE In-Process

1. MDD.ti,ab.
2. depress*.ti,ab.
3. 1 or 2 [Depression]
4. Psych* Test*.ti,ab.
5. case finding*.ti,ab.
6. casefinding*.ti,ab.
7. ((screen* or evaluat* or assess* or diagnos* or rating or rate or measur*) adj5 (index or indices or score* or scoring or test* or instrument* or inventor* or battery or batteries or tool* or scale* or checklist* or schedule*)).ti,ab.
8. ("mood and feelings questionnaire*" or MFQ or SMFQ or MFQ-C).ti,ab.
9. reynold* child* depression.ti,ab.
10. Reynold* adolesc* depression.ti,ab.
11. kutcher* adolesc*.ti,ab.
12. "depression scale for child*".ti,ab.
13. (Children's Depression Inventory or CDI).ti,ab.
14. (Brief Patient Health Questionnaire-9 or PHQ-9).ti,ab.
15. (Brief Patient Health Questionnaire-2 or PHQ-2).ti,ab.
16. (Brief Patient Health Questionnaire-A or PHQ-A).ti,ab.
17. (Center for Epidemiologic Studies Depression Scale or CES-D).ti,ab.
18. (Beck Depression Inventory or BDI or BDI-PC).ti,ab.
19. (Hopkins Symptom Checklist or HSCL or HSCL-10).ti,ab.

20. ("Strengths and Difficulties Questionnaire" or SDQ).ti,ab.
21. (Youth Self Report or YSR).ti,ab.
22. DesTeen.ti,ab.
23. or/4-22 [Screening]
24. (valid* or predict* or identif* or diagnos* or accura* or detect* or cut?off* or gold standard).ti,ab.
25. child.mp.
26. pediatric*.mp.
27. paediatric*.mp.
28. (boy or boys).mp.
29. girl*.mp.
30. (kid or kids).mp.
31. school?age*.mp.
32. juvenil*.mp.
33. under?age*.mp.
34. teen*.mp.
35. minor*.mp.
36. pubescen*.mp.
37. adolescen*.mp.
38. (youth or youths).mp.
39. child*.jw.
40. pediatric*.jw.
41. paediatric*.jw.
42. adolescen*.jw.
43. or/25-42
44. 3 and 23 and 24 and 43
45. limit 44 to ("in data review" or in process or "pubmed not medline")

HaPI

1. MDD.ti,ab.
2. depress*.ti,ab.
3. ("mood and feelings questionnaire*" or MFQ or SMFQ or MFQ-C).ti,ab.
4. reynold* child* depression.ti,ab.
5. Reynold* adolesc* depression.ti,ab.
6. kutcher* adolesc*.ti,ab.
7. "depression scale for child*".ti,ab.
8. (Children's Depression Inventory or CDI).ti,ab.
9. (Brief Patient Health Questionnaire-9 or PHQ-9).ti,ab.
10. (Brief Patient Health Questionnaire-2 or PHQ-2).ti,ab.
11. (Brief Patient Health Questionnaire-A or PHQ-A).ti,ab.
12. (Center for Epidemiologic Studies Depression Scale or CES-D).ti,ab.
13. (Beck Depression Inventory or BDI or BDI-PC).ti,ab.
14. (Hopkins Symptom Checklist or HSCL or HSCL-10).ti,ab.
15. ("Strengths and Difficulties Questionnaire" or SDQ).ti,ab.
16. (Youth Self Report or YSR).ti,ab.
17. DesTeen.ti,ab.

18. or/1-17
19. (valid* or predict* or identif* or diagnos* or accura* or detect* or cut?off* or gold standard).ti,ab.
20. child.mp.
21. pediatric*.mp.
22. paediatric*.mp.
23. (boy or boys).mp.
24. girl*.mp.
25. (kid or kids).mp.
26. school?age*.mp.
27. juvenil*.mp.
28. under?age*.mp.
29. teen*.mp.
30. minor*.mp.
31. pubescen*.mp.
32. adolescen*.mp.
33. (youth or youths).mp.
34. child*.jw.
35. pediatric*.jw.
36. paediatric*.jw.
37. adolescen*.jw.
38. or/20-37
39. 18 and 19 and 38
40. limit 39 to yr="2006-Current"
41. (201409* or 20141* or 2015*).up.
42. 40 and 41

EMBASE

1. Depression/
2. Depressive Disorder, Major/
3. depress*.ti.
4. or/1-3
5. exp Psychologic Test/
6. Psychiatric Diagnosis/
7. *Questionnaire/
8. Self-report/
9. Mass Screening/
10. case finding*.ti,ab.
11. casefinding*.ti,ab.
12. ((screen* or evaluat* or assess* or diagnos* or rating or rate or measur*) adj5 (index or indices or score* or scoring or test* or instrument* or inventor* or battery or batteries or tool* or scale* or checklist* or schedule*)).ti,ab.
13. ("mood and feelings questionnaire*" or MFQ or SMFQ or MFQ-C).ti,ab.
14. reynold* child* depression.ti,ab.
15. Reynold* adolesc* depression.ti,ab.
16. kutcher* adolesc*.ti,ab.

17. "depression scale for child*".ti,ab.
18. (Children's Depression Inventory or CDI).ti,ab.
19. (Brief Patient Health Questionnaire-9 or PHQ-9).ti,ab.
20. (Brief Patient Health Questionnaire-2 or PHQ-2).ti,ab.
21. (Brief Patient Health Questionnaire-A or PHQ-A).ti,ab.
22. (Center for Epidemiologic Studies Depression Scale or CES-D).ti,ab.
23. (Beck Depression Inventory or BDI or BDI-PC).ti,ab.
24. (Hopkins Symptom Checklist or HSCL or HSCL-10).ti,ab.
25. ("Strengths and Difficulties Questionnaire" or SDQ).ti,ab.
26. (Youth Self Report or YSR).ti,ab.
27. DesTeen.ti,ab.
28. or/5-27
29. (valid* or predict* or identif* or detect*).ti.
30. "Sensitivity and Specificity"/
31. Predictive Value/
32. (cut?off* or accura* or gold standard or reference standard).ab.
33. Validity/
34. or/29-33
35. exp Child/
36. child.mp.
37. Pediatrics/
38. Child Psychology/
39. Child Psychiatry/
40. pediatric*.mp.
41. paediatric*.mp.
42. (boy or boys).mp.
43. girl*.mp.
44. (kid or kids).mp.
45. school?age*.mp.
46. juvenil*.mp.
47. under?age*.mp.
48. teen*.mp.
49. minor*.mp.
50. pubescen*.mp.
51. adolescen*.mp.
52. (youth or youths).mp.
53. child*.jx.
54. pediatric*.jx.
55. paediatric*.jx.
56. adolescen*.jx.
57. or/35-56
58. 4 and 28 and 34 and 57
59. limit 58 to yr="2006 -Current"
60. Remove duplicates from 59
61. (201437* or 201438* or 201439* or 20144* or 20145* or 2015*).em.
62. 60 and 61

PsycINFO

1. "Depression (Emotion)"/
2. Major Depression/
3. depress*.ti.
4. or/1-3
5. ("mood and feelings questionnaire*" or MFQ or SMFQ or MFQ-C).ti,ab.
6. reynold* child* depression.ti,ab.
7. Reynold* adolesc* depression.ti,ab.
8. kutcher* adolesc*.ti,ab.
9. "depression scale for child*".ti,ab.
10. (Children's Depression Inventory or CDI).ti,ab.
11. (Brief Patient Health Questionnaire-9 or PHQ-9).ti,ab.
12. (Brief Patient Health Questionnaire-2 or PHQ-2).ti,ab.
13. (Brief Patient Health Questionnaire-A or PHQ-A).ti,ab.
14. (Center for Epidemiologic Studies Depression Scale or CES-D).ti,ab.
15. (Beck Depression Inventory or BDI or BDI-PC).ti,ab.
16. (Hopkins Symptom Checklist or HSCL or HSCL-10).ti,ab.
17. ("Strengths and Difficulties Questionnaire" or SDQ).ti,ab.
18. (Youth Self Report or YSR).ti,ab.
19. DesTeen.ti,ab.
20. Psychological Assessment/
21. Testing/
22. Psychodiagnosis/
23. Psychiatric Evaluation/
24. Screening/
25. Screening Tests/
26. Health Screening/
27. Rating Scales/
28. Self Report/
29. Self Evaluation/
30. Questionnaires/
31. Inventories/
32. Symptom Checklists/
33. case?finding*.ti,ab.
34. ((screen* or evaluat* or assess* or diagnos* or rating or rate or measur*) adj5 (index or indices or score* or scoring or test* or instrument* or inventor* or battery or batteries or tool* or scale* or checklist* or schedule*)).ti,ab.
35. or/5-34
36. (valid* or predict* or identif* or diagnos* or accura* or detect* or cut?off* or gold standard).ti,ab.
37. Prediction/
38. Test Validity/
39. Misdiagnosis/
40. Test Interpretation/
41. Cutting Scores/

42. or/36-41
43. child.mp.
44. Pediatrics/
45. Child Psychology/
46. Child Psychiatry/
47. pediatric*.mp.
48. paediatric*.mp.
49. (boy or boys).mp.
50. girl*.mp.
51. (kid or kids).mp.
52. school?age*.mp.
53. juvenil*.mp.
54. under?age*.mp.
55. teen*.mp.
56. minor*.mp.
57. pubescen*.mp.
58. adolescen*.mp.
59. (youth or youths).mp.
60. child*.jx.
61. pediatric*.jx.
62. paediatric*.jx.
63. adolescen*.jx.
64. or/43-63
65. 4 and 35 and 42 and 64
66. limit 65 to yr="2006 -Current"
67. remove duplicates from 66
68. (201409* or 20141* or 2015*).up.
69. 67 and 68

LILACS

(ti:Depress\$ OR mh:depression OR mh:"depressive disorder" OR mh:"depressive disorder, major") AND (test\$ OR instrument\$ OR measure\$ OR inventor\$ OR scale\$ OR evaluat\$ OR screen\$ OR assess\$ OR rating\$ OR rate\$ OR score\$ OR checklist\$ OR batter\$ OR score\$ OR scoring OR diagnos\$) AND (child\$ OR teen\$ OR adolescen\$ OR youth\$ OR pediatric\$ OR paediatric\$) AND (valid\$ OR predict\$ OR accura\$ OR detect\$ OR cut off\$ OR gold standard\$ OR identif\$ OR diagnos\$) AND (da:2006\$ OR da:2007\$ or da:2008\$ OR da:2009\$ or da:2010\$ OR da:2011\$ OR da:2012\$)

depress\$ AND (test\$ OR instrument\$ OR measure\$ OR inventor\$ OR scale\$ OR evaluat\$ OR screen\$ OR assess\$ OR rating\$ OR rate\$ OR score\$ OR checklist\$ OR batter\$ OR score\$ OR scoring OR diagnos\$) AND (child\$ OR teen\$ OR adolescen\$ OR youth\$ OR pediatric\$ OR paediatric\$) AND (valid\$ OR predict\$ OR accura\$ OR detect\$ OR cut off\$ OR gold standard\$ OR identif\$ OR diagnos\$)

in Title, Abstract, Subject
Filter: LILACS, 2014, 2015

SUPPLEMENTARY FILE 2: Variables Included in Data Extraction Form

Authors
Year
Country
Age group
Setting
Sample characteristics notes
Key inclusion criteria
Key exclusion criteria
Number of patients
Recruitment rate of eligible patients
Mean age
Percent male
Structured interview used (MDD criterion standard)
Number (%) with major depressive disorder
Screening tool and cut-off threshold
Derivation of cut-off (e.g., literature, exploratory)
Range of cut-offs reported
Number (%) above threshold on screening tool
Interviewer blinded to screening results?
Order of administration (structured interview versus screening tool)
Number positive MDD/positive screening tool
Number positive MDD/negative screening tool
Number negative MDD/positive screening tool
Number negative MDD/negative screening tool
Sensitivity
Specificity
Positive predictive value
Negative predictive value
Notes

SUPPLEMENTARY FILE 3: QUADAS-2 Risk of Bias and Applicability Judgments

Domain	Signaling Questions ^a	Risk of Bias ^b	Applicability Concerns ^b
Patient Selection	(1) Was a consecutive or random sample of patients enrolled? ^c (2) Was a case-control design avoided? (3) Did the study avoid inappropriate exclusions?	Could the selection of patients have introduced bias?	Are there concerns that the included patients and setting do not match the review question? ^d
Index Test	(1) Were the index test results interpreted without knowledge of the results of the reference standard? (2) If a threshold was used, was it pre-specified? ^e	Could the conduct or interpretation of the index test have introduced bias?	Are there concerns that the index test, its conduct, or its interpretation differ from the review question?
Reference Standard	(1) Is the reference standard likely to correctly classify the target condition? ^f (2) Were the reference standard results interpreted without knowledge of the results of the index test?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Are there concerns that the target condition as defined by the reference standard does not match the review question? ^g
Flow and Timing	(1) Was there an appropriate interval between index test and reference standard? ^h (2) Did all patients receive a reference standard? ⁱ (3) Did all patients receive the same reference standard? (4) Were all patients included in the analysis? ^j	Could the patient flow have introduced bias?	Not applicable

^a Signaling questions are rated “yes”, “no”, or “unclear”.

^b Risk of bias judgments and concerns related to applicability are rated “high”, “low” or “unclear”.

^c Rated “yes” if a consecutive or random sample of children/adolescents were recruited for the study. Rated “unclear” if the study indicates that consecutive children/adolescents were included, but does not specify the total number of eligibles and the number consented, or if a consecutive or random sample of children/adolescents was recruited but a low rate of eligible children/adolescents (below approximately 70% of eligibles) were included in the study, as it is unclear if the sample maintains the properties of a consecutive or random sample.

^d Rated “high” if study includes patients with already diagnosed or treated depression, as they would not be screened or constitute newly identified cases in clinical practice, and “unclear” if the study does not exclude patients with already diagnosed or treated depression, but does not specify the number of such patients that were included. Studies were not downgraded for only sampling children or adolescents from one age group or setting.

^e Rated “yes” if the study only tested a standard cutoff of the screening questionnaire, and the objective of the study was not to test the performance of the screening questionnaire at different cutoffs.

^f Rated “yes” for all studies as the systematic review inclusion criteria required the reference standard to be a standardized structured or semi-structured diagnostic interview based on DSM or ICD criteria.

^g Rated “unclear” or “high” if the study compares the screening tool to a reference standard of depressive disorders, which includes MDD/depressive episode and others (e.g., dysthymic disorder), depending on the proportion of cases of other depressive disorders included, but MDD/depressive episode cases comprise at least 80% of total cases.

^h Rated “yes” if the index test and reference standard were administered within 1 week of each other, “no” if longer, and “unclear” if not specified in the published report. Studies in which some patients were administered the index test and reference standard more than 2 weeks apart were excluded. When the interval between the administration of the screening instrument and the diagnostic interview was not specified in the published report, study authors were contacted for clarification to determine study eligibility. Quality assessment ratings, however, were based only on published information, with the exception of information on the interval between index test and reference standard, for which information was obtained from study authors to determine study eligibility.

ⁱ Rated “no” if all positive screens, but only a random sample of negative screens, were administered the reference standard, as the index score influenced the likelihood of receiving the reference standard.

^j Rated “no” or “unclear” if dropout occurred between the index test and reference standard (i.e., among children/adolescents who were supposed to receive both the index test and reference standard), depending on the proportion of missing data.

Reference: Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155:529-536

SUPPLEMENTARY FILE 4: QUADAS-2 Coding Notes

Adewuya, 2007

	Authors' Judgment	Support for Judgment
Domain 1: Patient Selection		
Signaling Questions:		
(1) Was a consecutive or random sample of patients enrolled?	Yes	Cluster random sampling. 1095 of 1200 sampled (91%) successfully recruited.
(2) Was a case-control design avoided?	Yes	
(3) Did the study avoid inappropriate exclusions?	Yes	
Risk of Bias:		
Could the selection of patients have introduced bias?	Low risk	
Applicability Concerns:		
Are there concerns that the included patients and setting do not match the review question?	Unclear	Study did not exclude adolescents with already diagnosed or treated depression and did not specify the number of such adolescents that were included.
Domain 2: Index Test		
Signaling Questions:		
(1) Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Sample-specific index test threshold based on exploratory data analysis methods.
(2) If a threshold was used, was it pre-specified?	No	
Risk of Bias:		
Could the conduct or interpretation of the index test have introduced bias?	High risk	
Applicability Concerns:		
Are there concerns that the index test, its conduct, or its interpretation differ from the review question?	Low	
Domain 3: Reference Standard		
Signaling Questions:		
(1) Is the reference standard likely to correctly classify the target condition?	Yes	Reference standard interviewers blind to index test results.
(2) Were the reference standard results interpreted without knowledge of the results of the index test?	Yes	
Risk of Bias:		
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk	
Applicability Concerns:		
Are there concerns that the target condition as defined by the reference standard does not match the review question?	Low	Index test compared to reference standard of DSM-IV major depressive disorder.
Domain 4: Flow and Timing		
Signaling Questions:		
(1) Was there an appropriate interval between index test and reference	Yes	Concurrent index test and reference standard administration

standard?			
(2) Did all patients receive a reference standard?	No		
(3) Did all patients receive the same reference standard?	Yes		
(4) Were all patients included in the analysis?	Yes		
Risk of Bias:			
Could the patient flow have introduced bias?	Unclear risk		per author contact. All students with BDI ≥ 10 (N=383) and 10% of students with BDI < 10 (N=71) were administered reference standard. All selected per the above were administered the index test and reference standard and included in analyses. Unclear risk of bias due to not administering reference standard to all students.

Araya, 2013

	Authors' Judgment	Support for Judgment
Domain 1: Patient Selection		
Signaling Questions:		
(1) Was a consecutive or random sample of patients enrolled?	Unclear	Method of recruitment and rate of recruitment not reported.
(2) Was a case-control design avoided?	Yes	
(3) Did the study avoid inappropriate exclusions?	Unclear	
Risk of Bias:		
Could the selection of patients have introduced bias?	Unclear risk	
Applicability Concerns:		
Are there concerns that the included patients and setting do not match the review question?	Unclear	Study did not exclude adolescents with already diagnosed or treated depression, and did not specify the number of such adolescents that were included.
Domain 2: Index Test		
Signaling Questions:		
(1) Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Sample-specific index test threshold based on exploratory data analysis methods.
(2) If a threshold was used, was it pre-specified?	No	
Risk of Bias:		
Could the conduct or interpretation of the index test have introduced bias?	High risk	
Applicability Concerns:		
Are there concerns that the index test, its conduct, or its interpretation differ from the review question?	Low	
Domain 3: Reference Standard		
Signaling Questions:		
(1) Is the reference standard likely to correctly classify the target condition?	Yes	Reference standard interviewers blind to index test results.
(2) Were the reference standard results interpreted without knowledge of the	Yes	

results of the index test?

Risk of Bias:

Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk

Applicability Concerns:

Are there concerns that the target condition as defined by the reference standard does not match the review questions? Low Index test compared to reference standard of DSM-IV major depressive disorder.

Domain 4: Flow and Timing

Signaling Questions:

(1) Was there an appropriate interval between index test and reference standard?	Yes	Reference standard administration within 72 hours of index test. Reference test administration based on BDI-II cutoffs. Prevalence of MDD in sample with reference standard over 50%, which could have biased reference standard administration/ interpretation. The authors do not report how many adolescents were initially administered BDI-II versus how many were administered the reference standard. Of 602 included in sample, 571 (95%) included in diagnostic accuracy data analysis.
(2) Did all patients receive a reference standard?	No	
(3) Did all patients receive the same reference standard?	Yes	
(4) Were all patients included in the analysis?	Yes	

Risk of Bias:

Could the patient flow have introduced bias? Unclear risk

Bang, 2015

	Authors' Judgment	Support for Judgment
--	--------------------------	-----------------------------

Domain 1: Patient Selection

Signaling Questions:

(1) Was a consecutive or random sample of patients enrolled?	No	Authors reported potential bias due to non-random sampling. Unlikely that it was consecutive. Rate of recruitment and method not reported.
(2) Was a case-control design avoided?	Yes	
(3) Did the study avoid inappropriate exclusions?	Yes	

Risk of Bias:

Could the selection of patients have introduced bias? Unclear

Applicability Concerns:

Are there concerns that the included patients and setting do not match the review question? Unclear Study did not exclude adolescents with already diagnosed or treated depression, and did not specify the number of such adolescents that were included.

Domain 2: Index Test

Signaling Questions:

(1) Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Sample-specific index test threshold based on exploratory data analysis methods.
(2) If a threshold was used, was it pre-specified?	No	

Risk of Bias:

Could the conduct or interpretation of the index test have introduced bias? High risk

Applicability Concerns:

Are there concerns that the index test, its conduct, or its interpretation differ from the review question? Low

Domain 3: Reference Standard**Signaling Questions:**

(1) Is the reference standard likely to correctly classify the target condition?	Yes	Reference standard interviewers blind to index test results.
(2) Were the reference standard results interpreted without knowledge of the results of the index test?	Yes	

Risk of Bias:

Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk

Applicability Concerns:

Are there concerns that the target condition as defined by the reference standard does not match the review question? Low

Index test compared to reference standard of DSM-IV-TR major depressive disorder.

Domain 4: Flow and Timing**Signaling Questions:**

(1) Was there an appropriate interval between index test and reference standard?	Unclear	Index test and reference standard administered within two weeks, however unclear whether it was administered within 1 week. 468 students (72.1%) were included in the diagnostic accuracy analysis. 54 students (8.3%) were excluded based on incomplete CDI questionnaires responses. 109 students (16.8%) were excluded as they didn't participate in the diagnostic interview, 18 (2.8%) students were excluded based on missing K-SADS items.
(2) Did all patients receive a reference standard?	No	
(3) Did all patients receive the same reference standard?	Yes	
(4) Were all patients included in the analysis?	No	

Risk of Bias:

Could the patient flow have introduced bias? Unclear

Barrera, 1988

	Authors' Judgment	Support for Judgment
--	-------------------	----------------------

Domain 1: Patient Selection**Signaling Questions:**

(1) Was a consecutive or random sample of patients enrolled?	Unclear	No information reported on sampling method, recruitment rate, or inclusion/exclusion criteria.
(2) Was a case-control design avoided?	Yes	
(3) Did the study avoid inappropriate exclusions?	Unclear	

Risk of Bias:

Could the selection of patients have introduced bias? Unclear risk

Applicability Concerns:

Are there concerns that the included patients and setting do not match the review question? Unclear

Study did not exclude children/adolescents with already diagnosed or treated depression,

and did not specify the number of such children/adolescents that were included.

Domain 2: Index Test		
Signaling Questions:		
(1) Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Sample-specific index test threshold based on exploratory data analysis methods.
(2) If a threshold was used, was it pre-specified?	No	
Risk of Bias:		
Could the conduct or interpretation of the index test have introduced bias?	High risk	
Applicability Concerns:		
Are there concerns that the index test, its conduct, or its interpretation differ from the review question?	Low	

Domain 3: Reference Standard		
Signaling Questions:		
(1) Is the reference standard likely to correctly classify the target condition?	Yes	Unclear risk due to lack of information on blinding of interviewers to index test results.
(2) Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear	
Risk of Bias:		
Could the reference standard, its conduct, or its interpretation have introduced bias?	Unclear risk	
Applicability Concerns:		
Are there concerns that the target condition as defined by the reference standard does not match the review question?	Low	Index test compared to reference standard of DSM-III major depressive disorder.

Domain 4: Flow and Timing		
Signaling Questions:		
(1) Was there an appropriate interval between index test and reference standard?	Yes	Index test and reference standard administered within one week. No information reported on missing data. Unclear risk of bias due to lack of information on missing data.
(2) Did all patients receive a reference standard?	Yes	
(3) Did all patients receive the same reference standard?	Yes	
(4) Were all patients included in the analysis?	Unclear	
Risk of Bias:		
Could the patient flow have introduced bias?	Unclear risk	

Butwicka, 2012		
	Authors' Judgment	Support for Judgment

Domain 1: Patient Selection		
Signaling Questions:		
(1) Was a consecutive or random sample of patients enrolled?	Yes	176 of 211 eligibles (83%) agreed to participate in study.
(2) Was a case-control design avoided?	Yes	

(3) Did the study avoid inappropriate exclusions? Yes

Risk of Bias:

Could the selection of patients have introduced bias? Low risk

Applicability Concerns:

Are there concerns that the included patients and setting do not match the review question? Unclear

Study did not exclude children/adolescents with already diagnosed or treated depression, and did not specify the number of such children/adolescents that were included.

Domain 2: Index Test

Signaling Questions:

(1) Were the index test results interpreted without knowledge of the results of the reference standard? Yes

Unclear if pre-specified or sample-specific index test threshold based on exploratory data analysis methods.

(2) If a threshold was used, was it pre-specified? Unclear

Risk of Bias:

Could the conduct or interpretation of the index test have introduced bias? High risk

Applicability Concerns:

Are there concerns that the index test, its conduct, or its interpretation differ from the review question? Low

Domain 3: Reference Standard

Signaling Questions:

(1) Is the reference standard likely to correctly classify the target condition? Yes

Two reference standard interviewers blinded to each other's results. However, it was not reported whether or not interviewers blinded to index test results.

(2) Were the reference standard results interpreted without knowledge of the results of the index test? Unclear

Risk of Bias:

Could the reference standard, its conduct, or its interpretation have introduced bias? Unclear risk

Applicability Concerns:

Are there concerns that the target condition as defined by the reference standard does not match the review question? Low

Index test compared to reference standard of DSM-IV major depressive disorder.

Domain 4: Flow and Timing

Signaling Questions:

(1) Was there an appropriate interval between index test and reference standard? Yes

All assessments conducted at same visit. 163 of 176 (93%) in study included in diagnostic accuracy analysis.

(2) Did all patients receive a reference standard? Yes

(3) Did all patients receive the same reference standard? Yes

(4) Were all patients included in the analysis? Yes

Risk of Bias:

Could the patient flow have introduced bias? Low risk

Canals, 2001

	Authors' Judgment	Support for Judgment
Domain 1: Patient Selection		
Signaling Questions:		
(1) Was a consecutive or random sample of patients enrolled?	Unclear	Appears to be longitudinal cohort study, with 304 participants representing remainder of 579 in original cohort 7-8 years prior. Original sampling frame not described. Inclusion/exclusion criteria not reported.
(2) Was a case-control design avoided?	Yes	
(3) Did the study avoid inappropriate exclusions?	Unclear	
Risk of Bias:		
Could the selection of patients have introduced bias?	Unclear risk	
Applicability Concerns:		
Are there concerns that the included patients and setting do not match the review question?	Unclear	Study did not exclude adolescents with already diagnosed or treated depression, and did not specify the number of such adolescents that were included.
Domain 2: Index Test		
Signaling Questions:		
(1) Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Sample-specific index test threshold based on exploratory data analysis methods.
(2) If a threshold was used, was it pre-specified?	No	
Risk of Bias:		
Could the conduct or interpretation of the index test have introduced bias?	High risk	
Applicability Concerns:		
Are there concerns that the index test, its conduct, or its interpretation differ from the review question?	Low	
Domain 3: Reference Standard		
Signaling Questions:		
(1) Is the reference standard likely to correctly classify the target condition?	Yes	Reference standard interviewers blind to index test results.
(2) Were the reference standard results interpreted without knowledge of the results of the index test?	Yes	
Risk of Bias:		
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk	
Applicability Concerns:		
Are there concerns that the target condition as defined by the reference standard does not match the review question?	Low	Index test compared to reference standard of ICD-10 major depressive episode.
Domain 4: Flow and Timing		
Signaling Questions:		
(1) Was there an appropriate interval between index test and reference	Yes	Reference standard administered same day as index test in most

standard?		cases, and at most a week later.
(2) Did all patients receive a reference standard?	Yes	Limited missing data, as 290 of 304 (95%) included were administered the reference standard.
(3) Did all patients receive the same reference standard?	Yes	
(4) Were all patients included in the analysis?	Yes	

Risk of Bias:
 Could the patient flow have introduced bias? Low risk

Fruhe (2012a, 2012b)
Fruhe, 2012a

	Authors' Judgment	Support for Judgment
Domain 1: Patient Selection		
Signaling Questions:		
(1) Was a consecutive or random sample of patients enrolled?	Yes	Of 352 children and adolescents who were eligible and contacted, 247 (70%) consented to participate. Exclusions were appropriate.
(2) Was a case-control design avoided?	Yes	
(3) Did the study avoid inappropriate exclusions?	Yes	
Risk of Bias:		
Could the selection of patients have introduced bias?	Low risk	
Applicability Concerns:		
Are there concerns that the included patients and setting do not match the review question?	Unclear	Study did not exclude children/adolescents with already diagnosed or treated depression, and did not specify the number of such children/adolescents that were included.
Domain 2: Index Test		
Signaling Questions:		
(1) Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Sample-specific index test threshold based on exploratory data analysis methods.
(2) If a threshold was used, was it pre-specified?	No	
Risk of Bias:		
Could the conduct or interpretation of the index test have introduced bias?	High risk	
Applicability Concerns:		
Are there concerns that the index test, its conduct, or its interpretation differ from the review question?	Low	
Domain 3: Reference Standard		
Signaling Questions:		
(1) Is the reference standard likely to correctly classify the target condition?	Yes	Reference standard interviewers blind to index test results.
(2) Were the reference standard results interpreted without knowledge of the results of the index test?	Yes	
Risk of Bias:		
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk	
Applicability Concerns:		

Are there concerns that the target condition as defined by the reference standard does not match the review question?	Unclear	Index test compared to DSM-IV reference standard of “any depressive disorder”, of which 9 of 11 cases had major depressive disorder and 2 had dysthymia.
---	---------	--

Domain 4: Flow and Timing

Signaling Questions:

(1) Was there an appropriate interval between index test and reference standard?	Yes	Reference standard interviews conducted within a week of index test. Only 1 participant of 247 was excluded from diagnostic accuracy analysis due to missing data.
(2) Did all patients receive a reference standard?	Yes	
(3) Did all patients receive the same reference standard?	Yes	
(4) Were all patients included in the analysis?	Yes	

Risk of Bias:

Could the patient flow have introduced bias?	Low risk
--	----------

Fruhe, 2012b

	Authors' Judgment	Support for Judgment
--	-------------------	----------------------

Domain 1: Patient Selection

Signaling Questions:

(1) Was a consecutive or random sample of patients enrolled?	Yes	Of 352 children and adolescents who were eligible and contacted, 247 (70%) consented to participate. Exclusions were appropriate.
(2) Was a case-control design avoided?	Yes	
(3) Did the study avoid inappropriate exclusions?	Yes	

Risk of Bias:

Could the selection of patients have introduced bias?	Low risk
---	----------

Applicability Concerns:

Are there concerns that the included patients and setting do not match the review question?	Unclear	Study did not exclude children/adolescents with already diagnosed or treated depression, and did not specify the number of such children/adolescents that were included.
---	---------	--

Domain 2: Index Test

Signaling Questions:

(1) Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Sample-specific index test threshold based on exploratory data analysis methods.
(2) If a threshold was used, was it pre-specified?	No	

Risk of Bias:

Could the conduct or interpretation of the index test have introduced bias?	High risk
---	-----------

Applicability Concerns:

Are there concerns that the index test, its conduct, or its interpretation differ from the review question?	Low
---	-----

Domain 3: Reference Standard		
Signaling Questions:		
(1) Is the reference standard likely to correctly classify the target condition?	Yes	Reference standard interviewers blind to index test results.
(2) Were the reference standard results interpreted without knowledge of the results of the index test?	Yes	
Risk of Bias:		
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk	
Applicability Concerns:		
Are there concerns that the target condition as defined by the reference standard does not match the review question?	Unclear	Index test compared to ICD-10 reference standard of "any depressive disorder", of which 10 of 12 cases had a major depressive episode and 2 cases had dysthymia.
Domain 4: Flow and Timing		
Signaling Questions:		
(1) Was there an appropriate interval between index test and reference standard?	Yes	Reference standard interviews conducted within a week of index test. There were 19 of 247 (8%) participants excluded from diagnostic accuracy data analysis due to missing data.
(2) Did all patients receive a reference standard?	Yes	
(3) Did all patients receive the same reference standard?	Yes	
(4) Were all patients included in the analysis?	Yes	
Risk of Bias:		
Could the patient flow have introduced bias?	Low risk	
Ganguly, 2013		
	Authors' Judgment	Support for Judgment
Domain 1: Patient Selection		
Signaling Questions:		
(1) Was a consecutive or random sample of patients enrolled?	Yes	91% participation rate from random sample.
(2) Was a case-control design avoided?	Yes	
(3) Did the study avoid inappropriate exclusions?	Yes	
Risk of Bias:		
Could the selection of patients have introduced bias?	Low risk	
Applicability Concerns:		
Are there concerns that the included patients and setting do not match the review question?	Unclear	Study did not exclude adolescents with already diagnosed or treated depression, and did not specify the number of such adolescents that were included.
Domain 2: Index Test		
Signaling Questions:		
(1) Were the index test results interpreted without knowledge of the results of the	Yes	Sample-specific index test threshold based on exploratory

reference standard?			
(2) If a threshold was used, was it pre-specified?	No		data analysis methods.

Risk of Bias:

Could the conduct or interpretation of the index test have introduced bias?	High risk		
---	-----------	--	--

Applicability Concerns:

Are there concerns that the index test, its conduct, or its interpretation differ from the review question?	Low		
---	-----	--	--

Domain 3: Reference Standard

Signaling Questions:

(1) Is the reference standard likely to correctly classify the target condition?	Yes		Reference standard interviewers blind to index test results.
(2) Were the reference standard results interpreted without knowledge of the results of the index test?	Yes		

Risk of Bias:

Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk		
--	----------	--	--

Applicability Concerns:

Are there concerns that the target condition as defined by the reference standard does not match the review question?	Low		Index test compared to reference standard of ICD-10 major depressive episode.
---	-----	--	---

Domain 4: Flow and Timing

Signaling Questions:

(1) Was there an appropriate interval between index test and reference standard?	Yes		Reference standard interview conducted within one week of index test. 233 of 250 who were sent screening questionnaires included in diagnostic accuracy analysis (93%).
(2) Did all patients receive a reference standard?	Yes		
(3) Did all patients receive the same reference standard?	Yes		
(4) Were all patients included in the analysis?	Yes		

Risk of Bias:

Could the patient flow have introduced bias?	Low risk		
--	----------	--	--

Katon, 2008

	Authors' Judgment	Support for Judgment
--	--------------------------	-----------------------------

Domain 1: Patient Selection

Signaling Questions:

(1) Was a consecutive or random sample of patients enrolled?	Unclear		Of 2648 eligible adolescents, 1481 consented (56%) and only 1379 (52%) completed study measures.
(2) Was a case-control design avoided?	Yes		
(3) Did the study avoid inappropriate exclusions?	Yes		

Risk of Bias:

Could the selection of patients have introduced bias?	Unclear risk		
---	--------------	--	--

Applicability Concerns:

Are there concerns that the included patients and setting do not match the review question?	Unclear	Study did not exclude children/adolescents with already diagnosed or treated depression, and did not specify the number of such children/adolescents that were included.
---	---------	--

Domain 2: Index Test

Signaling Questions:

(1) Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Sample-specific index test threshold based on exploratory data analysis methods.
(2) If a threshold was used, was it pre-specified?	No	

Risk of Bias:

Could the conduct or interpretation of the index test have introduced bias?	High risk
---	-----------

Applicability Concerns:

Are there concerns that the index test, its conduct, or its interpretation differ from the review question?	Low
---	-----

Domain 3: Reference Standard

Signaling Questions:

(1) Is the reference standard likely to correctly classify the target condition?	Yes	Unclear risk due to lack of information on blinding of interviewers to index test results.
(2) Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear	

Risk of Bias:

Could the reference standard, its conduct, or its interpretation have introduced bias?	Unclear risk
--	--------------

Applicability Concerns:

Are there concerns that the target condition as defined by the reference standard does not match the review question?	Low	Index test compared to reference standard of DSM-IV “depressive disorders”. However, of 83 cases, 82 had major depressive disorder and 1 case had dysthymia.
---	-----	--

Domain 4: Flow and Timing

Signaling Questions:

(1) Was there an appropriate interval between index test and reference standard?	Yes	Only 4 of 1379 participants in total sample were excluded from diagnostic accuracy analysis due to missing item responses.
(2) Did all patients receive a reference standard?	Yes	
(3) Did all patients receive the same reference standard?	Yes	
(4) Were all patients included in the analysis?	Yes	

Risk of Bias:

Could the patient flow have introduced bias?	Low risk
--	----------

Logsdon, 2010

Authors' Judgment

Support for Judgment

Domain 1: Patient Selection

Signaling Questions:		
(1) Was a consecutive or random sample of patients enrolled?	Unclear	Method of recruitment and rate of recruitment not reported.
(2) Was a case-control design avoided?	Yes	
(3) Did the study avoid inappropriate exclusions?	Yes	
Risk of Bias:		
Could the selection of patients have introduced bias?	Unclear risk	
Applicability Concerns:		
Are there concerns that the included patients and setting do not match the review question?	Unclear	Study did not exclude adolescents with already diagnosed or treated depression, and did not specify the number of such adolescents that were included.

Domain 2: Index Test

Signaling Questions:		
(1) Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Sample-specific index test threshold based on exploratory data analysis methods.
(2) If a threshold was used, was it pre-specified?	No	
Risk of Bias:		
Could the conduct or interpretation of the index test have introduced bias?	High risk	
Applicability Concerns		
Are there concerns that the index test, its conduct, or its interpretation differ from the review question?	Low	

Domain 3: Reference Standard

Signaling Questions:		
(1) Is the reference standard likely to correctly classify the target condition?	Yes	Reference standard interviewers blind to index test results.
(2) Were the reference standard results interpreted without knowledge of the results of the index test?	Yes	
Risk of Bias:		
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk	
Applicability Concerns:		
Are there concerns that the target condition as defined by the reference standard does not match the review question?	Low	Index test compared to reference standard of DSM-IV major depressive disorder.

Domain 4: Flow and Timing

Signaling Questions:		
(1) Was there an appropriate interval between index test and reference standard?	Yes	No information reported on missing data, but all study measures appear to have been administered at same visit. Thus, it unlikely that there were significant missing data.
(2) Did all patients receive a reference standard?	Yes	
(3) Did all patients receive the same reference standard?	Yes	
(4) Were all patients included in the analysis?	Yes	

Risk of Bias:

Could the patient flow have introduced bias? Low risk

Pietsch (2012, 2013)

Pietsch, 2012

	Authors' Judgment	Support for Judgment
Domain 1: Patient Selection		
Signaling Questions:		
(1) Was a consecutive or random sample of patients enrolled?	Yes	Consecutive sample. Of 446 eligibles (not counting other reasons for non-participation, such as released from hospital), 332 consented to participate (74%).
(2) Was a case-control design avoided?	Yes	
(3) Did the study avoid inappropriate exclusions?	Yes	
Risk of Bias:		
Could the selection of patients have introduced bias?	Low risk	
Applicability Concerns:		
Are there concerns that the included patients and setting do not match the review question?	Unclear	Study did not exclude adolescents with already diagnosed or treated depression, and did not specify the number of such adolescents that were included.
Domain 2: Index Test		
Signaling Questions:		
(1) Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Sample-specific index test threshold based on exploratory data analysis methods.
(2) If a threshold was used, was it pre-specified?	No	
Risk of Bias:		
Could the conduct or interpretation of the index test have introduced bias?	High risk	
Applicability Concerns:		
Are there concerns that the index test, its conduct, or its interpretation differ from the review question?	Low	
Domain 3: Reference Standard		
Signaling Questions:		
(1) Is the reference standard likely to correctly classify the target condition?	Yes	Reference standard interviewers blind to index test results.
(2) Were the reference standard results interpreted without knowledge of the results of the index test?	Yes	
Risk of Bias:		
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk	
Applicability Concerns:		
Are there concerns that the target condition as defined by the reference standard does not match the review question?	Low	Index test compared to reference standard of DSM-IV major depressive disorder.
Domain 4: Flow and Timing		

Signaling Questions:		
(1) Was there an appropriate interval between index test and reference standard?	Yes	All reference standard interviews administered within one week of index test. 18 of 332 participants (5%) were excluded from diagnostic accuracy analysis due to incomplete datasets (missing >20% of items).
(2) Did all patients receive a reference standard?	Yes	
(3) Did all patients receive the same reference standard?	Yes	
(4) Were all patients included in the analysis?	Yes	
Risk of Bias:		
Could the patient flow have introduced bias?	Low risk	

Pietsch, 2013

	Authors' Judgment	Support for Judgment
Domain 1: Patient Selection		
Signaling Questions:		
(1) Was a consecutive or random sample of patients enrolled?	Yes	Consecutive sample. Of 446 eligibles (not counting other reasons for non-participation, such as released from hospital), 332 consented to participate (74%).
(2) Was a case-control design avoided?	Yes	
(3) Did the study avoid inappropriate exclusions?	Yes	
Risk of Bias:		
Could the selection of patients have introduced bias?	Low risk	
Applicability Concerns:		
Are there concerns that the included patients and setting do not match the review question?	Unclear	Study did not exclude adolescents with already diagnosed or treated depression, and did not specify the number of such adolescents that were included.
Domain 2: Index Test		
Signaling Questions:		
(1) Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Sample-specific index test threshold based on exploratory data analysis methods.
(2) If a threshold was used, was it pre-specified?	No	
Risk of Bias:		
Could the conduct or interpretation of the index test have introduced bias?	High risk	
Applicability Concerns:		
Are there concerns that the index test, its conduct, or its interpretation differ from the review question?	Low	
Domain 3: Reference Standard		
Signaling Questions:		
(1) Is the reference standard likely to correctly classify the target condition?	Yes	Reference standard interviewers blind to index test results.
(2) Were the reference standard results interpreted without knowledge of the results of the index test?	Yes	
Risk of Bias:		

Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk

Applicability Concerns:

Are there concerns that the target condition as defined by the reference standard does not match the review question? Low Index test compared to reference standard of DSM-IV major depressive disorder.

Domain 4: Flow and Timing

Signaling Questions:

(1) Was there an appropriate interval between index test and reference standard?	Yes	All reference standard interviews administered within one week of index test. Only 5 of 332 participants excluded from diagnostic accuracy analysis due to incomplete datasets.
(2) Did all patients receive a reference standard?	Yes	
(3) Did all patients receive the same reference standard?	Yes	
(4) Were all patients included in the analysis?	Yes	

Risk of Bias:

Could the patient flow have introduced bias? Low risk

Richardson (2010a, 2010b)

Richardson, 2010a

	Authors' Judgment	Support for Judgment
Domain 1: Patient Selection		
Signaling Questions:		
(1) Was a consecutive or random sample of patients enrolled?	Unclear	Of 3775 eligibles, 2291 consented and completed study measures (61%). Possible bias due to somewhat low response rate in random sample.
(2) Was a case-control design avoided?	Yes	
(3) Did the study avoid inappropriate exclusions?	Yes	
Risk of Bias:		
Could the selection of patients have introduced bias?	Unclear risk	
Applicability Concerns:		
Are there concerns that the included patients and setting do not match the review question?	Unclear	Study did not exclude adolescents with already diagnosed or treated depression, and did not specify the number of such adolescents that were included.
Domain 2: Index Test		
Signaling Questions:		
(1) Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Sample-specific index test threshold based on exploratory data analysis methods.
(2) If a threshold was used, was it pre-specified?	No	
Risk of Bias:		
Could the conduct or interpretation of the index test have introduced bias?	High risk	
Applicability Concerns:		
Are there concerns that the index test, its conduct, or its interpretation differ from	Low	

the review question?

Domain 3: Reference Standard

Signaling Questions:

(1) Is the reference standard likely to correctly classify the target condition?	Yes	Telephone interview, in which index test administered followed by reference standard.
(2) Were the reference standard results interpreted without knowledge of the results of the index test?	No	

Risk of Bias:

Could the reference standard, its conduct, or its interpretation have introduced bias?	High risk
--	-----------

Applicability Concerns:

Are there concerns that the target condition as defined by the reference standard does not match the review question?	Low	Index test compared to reference standard of DSM-IV major depressive disorder.
---	-----	--

Domain 4: Flow and Timing

Signaling Questions:

(1) Was there an appropriate interval between index test and reference standard?	Yes	Index test and reference standard administered during same telephone interview. 271 of 291 adolescents with positive screen (PHQ-2 ≥ 3) invited to receive reference standard, versus 228 of 1993 with negative screens (PHQ-2 < 3). Of 499 participants invited to receive reference standard, 442 (89%) completed interview and were included in diagnostic accuracy analysis, including 89% in both positive screen and negative screen groups.
(2) Did all patients receive a reference standard?	No	
(3) Did all patients receive the same reference standard?	Yes	
(4) Were all patients included in the analysis?	Yes	

Risk of Bias:

Could the patient flow have introduced bias?	Unclear risk
--	--------------

Richardson, 2010b

	Authors' Judgment	Support for Judgment
--	--------------------------	-----------------------------

Domain 1: Patient Selection

Signaling Questions:

(1) Was a consecutive or random sample of patients enrolled?	Unclear	Of 3775 eligibles, 2291 consented and completed study measures (61%). Possible bias due to somewhat low response rate in random sample.
(2) Was a case-control design avoided?	Yes	
(3) Did the study avoid inappropriate exclusions?	Yes	

Risk of Bias:

Could the selection of patients have introduced bias?	Unclear risk
---	--------------

Applicability Concerns:

Are there concerns that the included patients and setting do not match the review question?	Unclear	Study did not exclude adolescents with already diagnosed or treated depression, and did not specify the number of such adolescents that were included.
---	---------	--

Domain 2: Index Test

Signaling Questions:

(1) Were the index test results interpreted	Yes	Sample-specific index test
---	-----	----------------------------

without knowledge of the results of the reference standard?			
(2) If a threshold was used, was it pre-specified?	No		threshold based on exploratory data analysis methods.

Risk of Bias:

Could the conduct or interpretation of the index test have introduced bias?	High risk		
---	-----------	--	--

Applicability Concerns:

Are there concerns that the index test, its conduct, or its interpretation differ from the review question?	Low		
---	-----	--	--

Domain 3: Reference Standard

Signaling Questions:

(1) Is the reference standard likely to correctly classify the target condition?	Yes		Telephone interview, in which index test administered followed by reference standard.
(2) Were the reference standard results interpreted without knowledge of the results of the index test?	No		

Risk of Bias:

Could the reference standard, its conduct, or its interpretation have introduced bias?	High risk		
--	-----------	--	--

Applicability Concerns:

Are there concerns that the target condition as defined by the reference standard does not match the review question?	Low		Index test compared to reference standard of DSM-IV major depressive disorder.
---	-----	--	--

Domain 4: Flow and Timing

Signaling Questions:

(1) Was there an appropriate interval between index test and reference standard?	Yes		Index test and reference standard administered during same telephone interview. 271 of 291 adolescents with positive screen (PHQ-2 ≥ 3) invited to receive reference standard, versus 228 of 1993 with negative screens (PHQ-2 <3). Of 499 participants invited to receive reference standard, 442 (89%) completed interview and were included in diagnostic accuracy analysis, including 89% in both positive screen and negative screen groups.
(2) Did all patients receive a reference standard?	No		
(3) Did all patients receive the same reference standard?	Yes		
(4) Were all patients included in the analysis?	Yes		

Risk of Bias:

Could the patient flow have introduced bias?	Unclear risk		
--	--------------	--	--

Roberts, 1991

	Authors' Judgment	Support for Judgment
--	-------------------	----------------------

Domain 1: Patient Selection

Signaling Questions:

(1) Was a consecutive or random sample of patients enrolled?	Unclear		Random selection of students from schools with approximately 61% of those selected participating in study. Possible bias due to somewhat low participation rate in random sample.
(2) Was a case-control design avoided?	Yes		
(3) Did the study avoid inappropriate exclusions?	Yes		

Risk of Bias:

Could the selection of patients have introduced bias?	Unclear risk		
---	--------------	--	--

Applicability Concerns:

Are there concerns that the included patients and setting do not match the review question?	Unclear	Study did not exclude adolescents with already diagnosed or treated depression, and did not specify the number of such adolescents that were included.
---	---------	--

Domain 2: Index Test**Signaling Questions:**

(1) Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Sample-specific index test threshold based on exploratory data analysis methods.
(2) If a threshold was used, was it pre-specified?	No	

Risk of Bias:

Could the conduct or interpretation of the index test have introduced bias?	High risk
---	-----------

Applicability Concerns:

Are there concerns that the index test, its conduct, or its interpretation differ from the review question?	Low
---	-----

Domain 3: Reference Standard**Signaling Questions:**

(1) Is the reference standard likely to correctly classify the target condition?	Yes	Unclear risk due to lack of information on blinding of interviewers to index test results.
(2) Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear	

Risk of Bias:

Could the reference standard, its conduct, or its interpretation have introduced bias?	Unclear risk
--	--------------

Applicability Concerns:

Are there concerns that the target condition as defined by the reference standard does not match the review question?	Low	Index test compared to reference standard of DSM-III major depressive disorder.
---	-----	---

Domain 4: Flow and Timing**Signaling Questions:**

(1) Was there an appropriate interval between index test and reference standard?	Yes	Concurrent administration of index test and reference standard. Thus, significant missing data unlikely
(2) Did all patients receive a reference standard?	Yes	
(3) Did all patients receive the same reference standard?	Yes	
(4) Were all patients included in the analysis?	Yes	

Risk of Bias:

Could the patient flow have introduced bias?	Low risk
--	----------

Tsai, 2014

Authors' Judgment

Support for Judgment

Domain 1: Patient Selection

Signaling Questions:		
(1) Was a consecutive or random sample of patients enrolled?	Unclear	Method of selection of potential participants was not described. 2257 of 3105 agreed to participate in overall cohort (73%).
(2) Was a case-control design avoided?	Yes	
(3) Did the study avoid inappropriate exclusions?	Yes	
Risk of Bias:		
Could the selection of patients have introduced bias?	Unclear risk	
Applicability Concerns:		
Are there concerns that the included patients and setting do not match the review question?	Unclear	Study did not exclude adolescents with already diagnosed or treated depression, and did not specify the number of such adolescents that were included.
Domain 2: Index Test		
Signaling Questions:		
(1) Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Sample-specific index test threshold based on exploratory data analysis methods.
(2) If a threshold was used, was it pre-specified?	No	
Risk of Bias:		
Could the conduct or interpretation of the index test have introduced bias?	High risk	
Applicability Concerns:		
Are there concerns that the index test, its conduct, or its interpretation differ from the review question?	Low	
Domain 3: Reference Standard		
Signaling Questions:		
(1) Is the reference standard likely to correctly classify the target condition?	Yes	Reference standard interviewers blind to index test results.
(2) Were the reference standard results interpreted without knowledge of the results of the index test?	Yes	
Risk of Bias:		
Could the reference standard, its conduct, or its interpretation have introduced bias?	Low risk	
Applicability Concerns:		
Are there concerns that the target condition as defined by the reference standard does not match the review question?	Low	Index test compared to reference standard of DSM-IV major depressive disorder.
Domain 4: Flow and Timing		
Signaling Questions:		
(1) Was there an appropriate interval between index test and reference standard?	Unclear	Interval between administration of index test and reference standard within 2 weeks per author, but not clear if within 1 week. Probability of invitation to receive reference standard based on participant history of self-harm. Only partial information provided on missing
(2) Did all patients receive a reference standard?	No	
(3) Did all patients receive the same reference standard?	Yes	
(4) Were all patients included in the	Unclear	

analysis?

Risk of Bias:

Could the patient flow have introduced bias? Unclear risk

data. It appears from Figure 1 that rate of missing data was low, since 53 of 62 (86%) participants with 1-year incidence of self-harm, of which all were invited to receive reference standard, completed the interview. Rate of missing data not known for group of participants with continued self-harm in past 2 years or control group since only a subset in each group completed the reference standard and it was not reported how many were approached to achieve that subset.

Turner, 2014

	Authors' Judgment	Support for Judgment
Domain 1: Patient Selection		
Signaling Questions:		
(1) Was a consecutive or random sample of patients enrolled?	Unclear	Cohort participants at longitudinal study visit. 10,101 participants invited for this study visit, of which 4503 attended. However, number of eligibles and number approached/recruited for original cohort not reported.
(2) Was a case-control design avoided?	Yes	
(3) Did the study avoid inappropriate exclusions?	Yes	
Risk of Bias:		
Could the selection of patients have introduced bias?	Unclear risk	
Applicability Concerns:		
Are there concerns that the included patients and setting do not match the review question?	Unclear	Study did not exclude adolescents with already diagnosed or treated depression, and did not specify the number of such adolescents that were included.
Domain 2: Index Test		
Signaling Questions:		
(1) Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Unclear to what degree index test threshold was based on exploratory data analysis methods.
(2) If a threshold was used, was it pre-specified?	Unclear	
Risk of Bias:		
Could the conduct or interpretation of the index test have introduced bias?	Unclear risk	
Applicability Concerns:		
Are there concerns that the index test, its conduct, or its interpretation differ from the review question?	Low	
Domain 3: Reference Standard		
Signaling Questions:		
(1) Is the reference standard likely to correctly classify the target condition?	Yes	Self-administered computerized assessment with diagnostic interview.
(2) Were the reference standard results interpreted without knowledge of the results of the index test?	Yes	

Risk of Bias:

Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk

Applicability Concerns:

Are there concerns that the target condition as defined by the reference standard does not match the review question? Low Index test compared to reference standard of ICD-10 major depressive episode.

Domain 4: Flow and Timing**Signaling Questions:**

(1) Was there an appropriate interval between index test and reference standard? Yes Same-day administration of index test and reference standard.
 (2) Did all patients receive a reference standard? Yes Diagnostic accuracy data were reported for 4027 of 4503 participants (91%) who attended clinic research visit.
 (3) Did all patients receive the same reference standard? Yes
 (4) Were all patients included in the analysis? Yes

Risk of Bias:

Could the patient flow have introduced bias? Low risk

Venkatesh, 2014**Authors' Judgment****Support for Judgment****Domain 1: Patient Selection****Signaling Questions:**

(1) Was a consecutive or random sample of patients enrolled? No Study done with adolescent mothers participating in clinical trial. Recruitment rate not reported.
 (2) Was a case-control design avoided? Yes
 (3) Did the study avoid inappropriate exclusions? Yes

Risk of Bias:

Could the selection of patients have introduced bias? Unclear risk

Applicability Concerns:

Are there concerns that the included patients and setting do not match the review question? Low Study excluded adolescents with already diagnosed or treated depression on enrollment.

Domain 2: Index Test**Signaling Questions:**

(1) Were the index test results interpreted without knowledge of the results of the reference standard? Yes Sample-specific index test threshold based on exploratory data analysis methods.
 (2) If a threshold was used, was it pre-specified? No

Risk of Bias:

Could the conduct or interpretation of the index test have introduced bias? High risk

Applicability Concerns:

Are there concerns that the index test, its conduct, or its interpretation differ from the review question? Low

Domain 3: Reference Standard		
Signaling Questions:		
(1) Is the reference standard likely to correctly classify the target condition?	Yes	Unclear risk due to lack of information on blinding of interviewers to index test results.
(2) Were the reference standard results interpreted without knowledge of the results of the index test?	Unclear	
Risk of Bias:		
Could the reference standard, its conduct, or its interpretation have introduced bias?	Unclear risk	
Applicability Concerns:		
Are there concerns that the target condition as defined by the reference standard does not match the review question?	Low	Index test compared to reference standard of DSM-IV major depressive disorder.
Domain 4: Flow and Timing		
Signaling Questions:		
(1) Was there an appropriate interval between index test and reference standard?	Yes	Index test and reference standard administered at same visit. 10 of 106 participants in study (9%) excluded from diagnostic accuracy analysis due to missing data at time point extracted.
(2) Did all patients receive a reference standard?	Yes	
(3) Did all patients receive the same reference standard?	Yes	
(4) Were all patients included in the analysis?	Yes	
Risk of Bias:		
Could the patient flow have introduced bias?	Low risk	
Ventevogel, 2014		
	Authors' Judgment	Support for Judgment
Domain 1: Patient Selection		
Signaling Questions:		
(1) Was a consecutive or random sample of patients enrolled?	Unclear	Children and adolescents randomly sampled from schools, but no data reported on participation rate.
(2) Was a case-control design avoided?	Yes	
(3) Did the study avoid inappropriate exclusions?	Yes	
Risk of Bias:		
Could the selection of patients have introduced bias?	Unclear risk	
Applicability Concerns:		
Are there concerns that the included patients and setting do not match the review question?	High	Sample included 16 of 65 (25%) participants recruited who had already been identified in the psychosocial care program and received individual psychosocial services.
Domain 2: Index Test		
Signaling Questions:		
(1) Were the index test results interpreted without knowledge of the results of the reference standard?	Yes	Sample-specific index test threshold based on exploratory data analysis methods.

(2) If a threshold was used, was it pre-specified? No

Risk of Bias:

Could the conduct or interpretation of the index test have introduced bias? High risk

Applicability Concerns:

Are there concerns that the index test, its conduct, or its interpretation differ from the review question? Low

Domain 3: Reference Standard

Signaling Questions:

(1) Is the reference standard likely to correctly classify the target condition? Yes Reference standard interviewers blind to index test results.

(2) Were the reference standard results interpreted without knowledge of the results of the index test? Yes

Risk of Bias:

Could the reference standard, its conduct, or its interpretation have introduced bias? Low risk

Applicability Concerns:

Are there concerns that the target condition as defined by the reference standard does not match the review question? Unclear

Index test compared to reference standard of DSM-IV depressive disorders. Of 11 cases, either 9 or 10 had MDD and either 1 (9%) or 2 (18%) had adjustment disorder with depressive symptoms.

Domain 4: Flow and Timing

Signaling Questions:

(1) Was there an appropriate interval between index test and reference standard? Yes Same-day administration of index test and reference standard. Only 4 of 65 (6%) in sample not included in diagnostic accuracy analysis.

(2) Did all patients receive a reference standard? Yes

(3) Did all patients receive the same reference standard? Yes

(4) Were all patients included in the analysis? Yes

Risk of Bias:

Could the patient flow have introduced bias? Low risk

SUPPLEMENTARY FILE 5. Excluded Studies from Previous Systematic Reviews

First Author, Journal, Year	Systematic Review	Screening Tool(s) Evaluated	Reason for Exclusion
Aebi, <i>J Affect Dis</i> , 2009 [1]	Stockings (2015)	YSR	No validated diagnostic interview administered. Reference standard was CES-D.
Allgaier, <i>J Psychosom Res</i> , 2012 [2]	Stockings (2015)	CDI, CDI-SF	Compared screening tool to major and minor depression. Number of patients diagnosed with major depression not reported.
Ambrosini, <i>J Am Acad Child Adolesc Psychiatry</i> , 1991 [3]	Stockings (2015)	BDI	Sample comprised of outpatients referred to a psychiatric clinic for depression.
Bennett, <i>J Affect Dis</i> , 1997 [4]	Stockings (2015)	BDI	Sample comprised of children and adolescents referred to a depression clinic.
Betancourt, <i>J Am Acad Child Adolesc Psychiatry</i> , 2012 [5]	Stockings (2015)	CES-DC	Majority of children in study were recruited due to presence of mental health syndromes. Number of patients diagnosed with major depression not reported, so 2x2 table could not be completed.
Blom, <i>J Affect Dis</i> , 2010 [6]	Stockings (2015)	BDI, HADS, SDQ	Clinical group was recruited from psychiatric clinics; number of patients diagnosed with major depression in the non-referred group not reported.
Craighead, <i>Psychol Assess</i> , 1995 [8]	Stockings (2015)	CDI	Sample comprised of psychiatric inpatients.
Dolle, <i>Psychiatry Res</i> , 2012 [9]	Stockings (2015)	BDI-II	Sample comprised of psychiatric patients.
Fendrich, <i>Am J Epidemiol</i> , 1990 [10]	Stockings (2015)	CES-DC	Less than 80% of the sample under age 18; results for children and adolescents under age 18 not reported separately from those over age 18.
Figueras-Masip, <i>Span J Psychol</i> , 2010 [11]	Stockings (2015)	CDI	Diagnosis was not based on a validated diagnostic interview. Only patients recruited from mental health centers and participants from the community sample with positive depression screens received the reference standard interview.

Fristad, <i>J Affect Disord</i> , 1988 [12]	Stockings (2015)	CDI	Study included only children in psychiatric care and children who were pre-screened to rule out psychiatric diagnoses. Thus, no children with possible depression who were not already in psychiatric treatment were included.
Fundudis, <i>Br J Psychiatry</i> , 1991 [13]	Stockings (2015)	CDI, DSRS	Sample recruited at a child psychiatry department.
Garrison, <i>J Am Acad Child Adolesc Psychiatry</i> , 1991 [14]	AHRQ(2009), AHRQ (2016) and Stockings (2015)	CES-D	Diagnostic interview consistently administered more than 2 weeks after the screening instrument (per author report).
Goodman, <i>Int Rev Psychiatry</i> , 2003 [15]	AHRQ (2009)	SDQ	Compared screening tool to any depressive disorder. Number of patients diagnosed with major depression not reported.
Johnson, <i>J Adolesc Health</i> , 2002 [16]	AHRQ (2009), AHRQ (2016)	PHQ-A	Diagnostic interview administered was a composite of items from the SCID and PRIME-MD, but not a validated diagnostic interview.
Kashani, <i>J Am Acad Child Adolesc Psychiatry</i> , 1990 [17]	Stockings (2015)	BDI	Sample comprised of adolescents attending an outpatient counselling centre.
Krefetz, <i>J Pers Assess</i> , 2002 [18]	Stockings (2015)	BDI-II, RADS	Sample comprised of psychiatric inpatients.
Kumar, <i>Assessment</i> , 2002 [19]	Stockings (2015)	BDI-II	Sample comprised of psychiatric inpatients.
Marton, <i>Can J Psychiat</i> , 1991 [20]	Stockings (2015)	BDI	Sample comprised of psychiatric inpatients and outpatients.
Osman, <i>J Clin Psychol</i> , 2010 [21]	Stockings (2015)	RADS-2	Sample comprised of psychiatric inpatients.
Patton, <i>Soc Psych Psych Epid</i> , 1999 [22]	AHRQ (2009), AHRQ (2016)	N/A	No screening instrument administered. Index test was the CIS-R, which is a diagnostic interview.
Prescott, <i>J Am Acad Child Adolesc Psychiatry</i> , 1998 [23]	Stockings (2015)	CES-D	Diagnostic interview administered more than 2 weeks after the screening instrument.
Roelofs, <i>Psychol Assess</i> , 2010 [24]	Stockings (2015)	CDI, YSR	Diagnostic interview administered only to children in treatment for psychopathology.
Russell, <i>Indian J Pediatrics</i> , 2012 [25]	Stockings (2015)	BDI	Of those diagnosed with “depressive disorders”, less than 80% were diagnosed with major depression.

Shemesh, <i>J Am Acad Child Adolesc Psychiatry</i> , 2005 [26]	Stockings (2015)	CDI	Sample comprised of children referred for evaluation for emotional problems.
Sørensen, <i>Eur Child Adolesc Psychiatry</i> , 2005 [27]	Stockings (2015)	CDI	Sample comprised of psychiatric patients.
Strober, <i>J Consult Clin Psychol</i> , 1981 [28]	Stockings (2015)	BDI	Sample comprised of patients admitted to an adolescent inpatient psychiatric service.
Timbremont, <i>J Clin Child Adolesc</i> , 2004 [29]	Stockings (2015)	CDI	Sample comprised of psychiatric inpatients and outpatients.
Whitaker, <i>Arch Gen Psychiatry</i> , 1990 [30]	AHRQ (2009) and Stockings (2015)	BDI	No validated diagnostic interview administered; reference standard was a composite of sections from different interviews based on DSM-III.
Winter, <i>J Adolesc Health</i> , 1999 [31]	AHRQ (2009)	BDI-PC	No validated diagnostic interview administered. Reference standard was the PRIME-MD.
Yang, <i>J Affect Disord</i> , 2004 [32]	Stockings (2015)	CES-D	Study did not use a set cutoff to classify children as positive screens. Rather, they identified those above a percentile in the study sample.

Abbreviations: BDI = Beck Depression Inventory; BDI-II = Beck Depression Inventory, Second Edition; BDI-PC = Beck Depression Inventory for Primary Care; CDI = Children's Depression Inventory; CDI-SF = Children's Depression Inventory, Short Form; CDRS-R = Children's Depression Rating Scale-Revised; CES-D = Center for Epidemiological Studies Depression Scale; CES-DC = Center for Epidemiological Studies Depression Scale for Children; CIS-R = Clinical Interview Schedule-Revised; DSM-III = Diagnostic and Statistical Manual of Mental Disorders, Third Edition; DSRS = Depression Self-Rating Scale; HADS = Hospital Anxiety and Depression Scale; PHQ-A = Patient Health Questionnaire for Adolescents; PRIME-MD = Primary Care Evaluation of Mental Disorders; RADS = Reynolds Adolescent Depression Scale; RADS-2 = Reynolds Adolescent Depression Scale, Second Edition; SCID = Structured Clinical Interview for DSM Disorders; SDQ = Strength and Difficulties Questionnaire; USPSTF = United States Preventive Services Task Force; YSR = Youth Self-Report

References for Excluded Studies Table

- [1] Aebi M, Winkler Metzke C, Steinhausen H-C. Prediction of major affective disorders in adolescents by self-report measures. *J Affect Dis.* 2009;115:140-9.
- [2] Allgaier A-K, Frühe B, Pietsch K, Saravo B, Baethmann M, Schulte-Körne G. Is the Children's Depression Inventory Short version a valid screening tool in pediatric care? A comparison to its full-length version. *J Psychosom Res.* 2012;73:369-74.
- [3] Ambrosini PJ, Metz C, Bianchi MD, Rabinovich H, Undie A. Concurrent validity and psychometric properties of the Beck Depression Inventory in outpatient adolescents. *J Am Acad Child Adolesc Psychiatry.* 1991;30:51-7.
- [4] Bennett DS, Ambrosini PJ, Bianchi M, Barnett D, Metz C, Rabinovich H. Relationship of Beck Depression Inventory factors to depression among adolescents. *J Affect Dis.* 1997;45:127-34.
- [5] Betancourt T, Scorza P, Meyers-Ohki S, Mushashi C, Kayiteshonga Y, Binagwaho A, et al. Validating the Center for Epidemiological Studies Depression Scale for children in Rwanda. *J Am Acad Child Adolesc Psychiatry.* 2012;51:1284-92.
- [6] Blom EH, Larsson J-O, Serlachius E, Ingvar M. The differentiation between depressive and anxious adolescent females and controls by behavioural self-rating scales. *J Affect Dis.* 2010;122:232-40.
- [7] Canals J, Marti-Henneberg C, Fernández-Ballart J, Domènech E. A longitudinal study of depression in an urban spanish pubertal population. *Eur Child Adolesc Psychiatry.* 1995;4:102-11.
- [8] Craighead WE, Curry JF, Ilardi SS. Relationship of Children's Depression Inventory factors to major depression among adolescents. *Psychol Assess.* 1995;7:171-6.
- [9] Dolle K, Schulte-Körne G, O'Leary AM, von Hofacker N, Izat Y, Allgaier A-K. The Beck Depression Inventory-II in adolescent mental health patients: cut-off scores for detecting depression and rating severity. *Psychiatry Res.* 2012;200:843-8.
- [10] Fendrich M, Weissman MM, Warner V. Screening for depressive disorder in children and adolescents: validating the Center for Epidemiologic Studies Depression Scale for children. *Am J Epidemiol.* 1990;131:538-51.

- [11] Figueras-Masip A, Amador-Campos JA, Gómez-Benito J, Gándara VDB. Psychometric properties of the Children's Depression Inventory in community and clinical sample. *Span J Psychol*. 2010;13:990-9.
- [12] Fristad MA, Weller EB, Weller RA, Teare M, Preskorn SH. Self-report vs. biological markers in assessment of childhood depression. *J Affect Dis*. 1988;15:339-45.
- [13] Fundudis T, Berney T, Kolvin I, Famuyiwa O, Barrett L, Bhate S, et al. Reliability and validity of two self-rating scales in the assessment of childhood depression. *Br J Psychiatry*. 1991;159:36-40.
- [14] Garrison CZ, Addy CL, Jackson KL, McKeown RE, Waller JL. The CES-D as a screen for depression and other psychiatric disorders in adolescents. *J Am Acad Child Adolesc Psychiatry*. 1991;30:636-41.
- [15] Goodman R, Ford T, Simmons H, Gatward R, Meltzer H. Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *Int Rev Psychiatry*. 2003;15:166-72.
- [16] Johnson JG, Harris ES, Spitzer RL, Williams JBW. The patient health questionnaire for adolescents: validation of an instrument for the assessment of mental disorders among adolescent primary care patients. *J Adolesc Health*. 2002;30:196-204.
- [17] Kashani JH, Sherman DD, Parker DR, Reid JC. Utility of the Beck Depression Inventory with clinic-referred adolescents. *J Am Acad Child Adolesc Psychiatry*. 1990;29:278-82.
- [18] Krefetz DG, Steer RA, Gulab NA, Beck AT. Convergent validity of the Beck Depression Inventory-II with the Reynolds Adolescent Depression Scale in psychiatric inpatients. *J Pers Assess*. 2002;78:451-60.
- [19] Kumar G, Steer RA, Teitelman KB, Villacis L. Effectiveness of Beck Depression Inventory-II subscales in screening for major depressive disorders in adolescent psychiatric inpatients. *Assessment*. 2002;9:164-70.
- [20] Marton P, Churchard M, Kutcher S, Korenblum M. Diagnostic utility of the Beck Depression Inventory with adolescent psychiatric outpatients and inpatients. *Can J Psychiat*. 1991;36:428-31.
- [21] Osman A, Gutierrez PM, Bagge CL, Fang Q, Emmerich A. Reynolds adolescent depression scale-second edition: a reliable and useful instrument. *J Clin Psychol*. 2010;66:1324-45.
- [22] Patton GC, Coffey C, Posterino M, Carlin JB, Wolfe R, Bowes G. A computerised screening instrument for adolescent depression: population-based validation and application to a two-phase case-control study. *Soc Psych Psych Epidemiol*. 1999;34:166-72.

- [23] Prescott CA, McArdle JJ, Hishinuma ES, Johnson RC, Miyamoto RH, Andrade NN, et al. Prediction of major depression and dysthymia from CES-D scores among ethnic minority adolescents. *J Am Acad Child Adolesc Psychiatry*. 1998;37:495-503.
- [24] Roelofs J, Braet C, Rood L, Timbremont B, van Vlierberghe L, Goossens L, et al. Norms and screening utility of the Dutch version of the Children's Depression Inventory in clinical and nonclinical youths. *Psychol Assess*. 2010;22:866-77.
- [25] Russell PS, Basker M, Russell S, Moses PD, C. Nair MK, Minju KA. Comparison of a self-rated and a clinician-rated measure for identifying depression among adolescents in a primary-care setting. *Indian J Pediatrics*. 2012;79:S45-S51.
- [26] Shemesh E, Yehuda R, Rockmore L, Shneider BL, Emre S, Bartell AS, et al. Assessment of depression in medically ill children presenting to pediatric specialty clinics. *J Am Acad Child Adolesc Psychiatry*. 2005;44:1249-57.
- [27] Sørensen M, Frydenberg M, Thastum M, Thomsen P. The Children's Depression Inventory and classification of major depressive disorder. *Eur Child Adolesc Psychiatry*. 2005;14:328-34.
- [28] Strober M, Green J, Carlson G. Utility of the Beck Depression Inventory with psychiatrically hospitalized adolescents. *J Consult Clin Psychol*. 1981;49:482-3.
- [29] Timbremont B, Braet C, Dreesen L. Assessing depression in youth: relation between the Children's Depression Inventory and a structured interview. *J Clin Child Adoles*. 2004;33:149-57.
- [30] Whitaker A, Johnson J, Shaffer D, et al. Uncommon troubles in young people: prevalence estimates of selected psychiatric disorders in a nonreferred adolescent population. *Arch Gen Psychiatry*. 1990;47:487-96.
- [31] Winter LB, Steer RA, Jones-Hicks L, Beck AT. Screening for major depression disorders in adolescent medical outpatients with the Beck Depression Inventory for primary care. *J Adolesc Health*. 1999;24:389-94.
- [32] Yang H-J, Soong W-T, Kuo P-H, Chang H-L, Chen WJ. Using the CES-D in a two-phase survey for depressive disorders among nonreferred adolescents in Taipei: a stratum-specific likelihood ratio analysis. *J Affect Dis*. 2004;82:419-30.