ARTICLE

Can the Center for Epidemiologic Studies Depression Scale Be Used to Screen for Depression in Children and Adolescents? An Updated Systematic Review

Seong-Hi Park[1] · Young Mi Kwon[2]

Abstract

This study is an updated systematic review verifying whether the Center for Epidemiologic Studies Depression Scale (CES-D) is a valuable screening tool for children and adolescents. Electronic searches were performed on MEDLINE, EMBASE, CINAHL, and PsycArticles, using depression and *CES-D* as keywords. Fourteen studies that included 7,843 children and adolescents were analyzed. In the meta-analysis by CES-D type, the pooled sensitivity and specificity for the long version were 0.81 and 0.72, respectively; they were 0.80 and 0.74 for the short version, respectively. The summary receiver operating characteristic (sROC) curves were 0.83 and 0.86, respectively. Compared to the CES-D and other tools, the pooled sensitivity (0.84 vs. 0.83) and the pooled specificity (0.72 vs. 0.74) were similar, and the sROC curve was the same at 0.83. This review indicates that the CES-D is an available and valuable tool for screening depression in children and adolescents.

Keywords Child · Adolescent · Depression · Sensitivity and specificity · Systematic review

Introduction

Depression in children and adolescents is common but often unrecognized [1]. Some studies report an increasing prevalence of depression in children and adolescents, which reaches or can reach 15–20% [2, 3]. The depressive response of children and adolescents, unlike adults, involves low self-esteem, displays strong guilt, adversely affects essential daily activities such as learning, and is closely related to disruptive behavior [4, 5]. Studies on childhood and adolescent depression confirm that it is a chronic and recurrent condition. Most episodes remit within a year, but there is a high risk of recurrence, and 50–70% have the potential to develop additional episodes within five years [6]. Moreover, as they are still immature and have not fully developed their ability to cope, it is challenging to overcome independently [4,

5, 7]. In 2015, the World Health Organization announced that suicide was the second most common cause of death in the 15–29-year-old age group, calling for attention to the severity of depression in children and adolescents [8, 9]. Therefore, depression in children and adolescents requires early treatment based on early assessment.

Researchers' opinions on screening for depression in children and adolescents are inconsistent [10]. In 2016, the U.S. Preventive Services Task Force (USPSTF) repeated its 2009 recommendations for screening for depression in primary care settings to ensure an accurate diagnosis, effective treatment, and appropriate follow-up [8]. However, some studies' findings were not supported [11]. The USPSTF did not provide a screening interval but concluded with moderate certainty that the net benefit was for children and adolescents between 12–18 years of age [7]. In addition, many depression-screening tools have been developed and used, including the Patient Health Questionnaire (PHQ) for adolescents and the Beck Depression Inventory (BDI) [5, 8].

The Center for Epidemiologic Studies Depression Scale (CES-D), developed by Radloff in 1977, is one of the most widely used self-rating scales to screen for MDD in psychiatric epidemiology [12]. The CES-D can be applied to the general population because its reliability and validity were tested in those under 25, 25–64, and over 64 years

✉ Seong-Hi Park
shpark2015@sch.ac.kr

[1] School of Nursing, Soonchunhyang University, 22 Soonchunhyang-Ro, Sinchang-Myen, Asan-Si 31538, Chungcheongnam-Do, Korea

[2] Department of Nursing, Kyung-in Women's University, Incheon, Korea

of age during its development. In 1991, Radloff reported that the CES-D was verified for use in children and adolescents [13, 14]. However, as this is not a diagnostic tool, the results (scores) should be interpreted only as of the level of symptoms accompanying depression, and the appropriate cut-off scores have not been verified.

The predictive validity of a screening tool is essential for accurate disease detection. Two systematic reviews (SRs) on the accuracy of depression-screening tools in children and adolescents were published in 2015 and 2016 [14, 15]. In both SRs, the CES-D was commonly reviewed in addition to the PHQ-9 and the BDI for children and adolescents. Nevertheless, there was no quantitative analysis such as a meta-analysis. Furthermore, it is difficult to use different screening tools for each patient's age in a primary care setting. Nonetheless, the CES-D is the most commonly used depression-screening tool for epidemiological investigations of adults. Therefore, to apply this tool to children and adolescents, a study is needed to quantitatively verify the screening accuracy of the CES-D.

This study is an updated SR on the screening accuracy of the CES-D in children and adolescents. This SR aims to (i) analyze the difference in screening accuracy of each CES-D type for MDD and (ii) compare the predictive validity of the CES-D with that of other screening tools.

Methods

The methodology and reporting in this SR are consistent with the guidelines of the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy [16] and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 Statement [17].

Search Strategy and Data Sources

The electronic database searches were conducted twice, on August 8, 2020, and March 15, 2022, through MEDLINE, Embase, CINAHL, and PsycArticles. The same search strategy was used for both searches. The keywords were *depression* and *CES-D*. We performed an expanded search for *depression* using MeSH, and *CES-D* was searched for using its full names and abbreviations. Age was not limited for the search process. We also scanned the reference lists and eligible articles that were cited. An example of this search strategy is shown in Supplementary Table 1.

Eligibility Criteria

The inclusion criteria were as follows: (i) Types of studies: Studies (e.g., cohort study and cross-sectional study) that examined the diagnostic accuracy of the CES-D were included, but those providing insufficient information to construct a 2×2 contingency table were excluded. (ii) Types of participants: Only studies in which junior and high school students were included as subjects were considered. (iii) Indexed tests: All the CES-D types were included. (iv) Gold standards: The selected studies were limited to those in which participants were diagnosed with MDD through direct interviews with psychologists that followed the Diagnostic and Statistical Manual of Mental Disorders (DSM) and International Classification of Diseases-Tenth Revision (ICD-10) and/or semi-structured interviews, such as the Structured Clinical Interview for DSM Disorders (SCID), the Mini International Neuropsychiatric Interview (MINI), and the Schedule for Affective Disorders and Schizophrenia for School Age Children (K-SADS). (v) Types of outcomes: Selected studies that could derive true positive (TP), false positive (FP), false negative (FN), and true negative values (TN), and thereby, sensitivity, specificity, positive and negative likelihood ratio, diagnosis odds ratio, and summary receiver operating characteristic (sROC) curves of the CES-D were compared.

Study Screening and Data Extraction

In the literature selection process, one reviewer first removed duplicate articles, following which two reviewers independently confirmed the potential suitability through titles and abstracts in all articles. Differences at each stage were resolved by consensus. When multiple eligible articles using the same dataset were found, we selected the most recently published articles.

The following data were extracted from the selected studies: year of publication, authors, location, sample size, sample characteristics, the prevalence of MDD, the gold standard for MDD diagnosis, the CES-D type and cut-off scores, comparators, TP, FP, FN, and TN. All processes were first reviewed by one reviewer and then independently by another reviewer. During data extraction, disagreement between the reviewers was resolved by consensus.

Quality Assessment

Two reviewers independently assessed the quality of the selected studies using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) [18]. The QUADAS-2 was determined across four domains: patient selection, index test, reference standard, and flow and timing and was categorized and quantified as either low, unclear, or high to derive a total quality score.

Statistical Analysis

Meta-analysis was performed using the MetaDiSc 1.4 and MetaDTA software programs [19–21]. Based on the TP, FP, FN, and TN described in the 2×2 contingency table, screening accuracy was evaluated by yielding pooled sensitivity, specificity, positive and negative likelihood ratios, and diagnostic odds ratios with 95% confidence intervals (CIs). The area under the curve (AUC) of the sROC curve and the Q* value were analyzed to determine the test accuracy. We used a bivariate random effects model [22]. Heterogeneity among studies was judged using random effect (RE) correlation. The AUC values were interpreted as follows: AUC of 0.5 = non-informative test; AUC of 0.5–0.7 = low accurate; AUC of 0.7–0.9 = moderate accurate; AUC of 0.9–1 = highly accurate; and AUC of 1 = perfect test [23]. The Q* value represents the point at which sensitivity and specificity are equal in the sROC curve, with a value of 1 indicating accuracy of 100% [24].

Results

Selection Process and Assessing Bias Risk

A total of 1,443 articles were identified through two electronic database searches (the first, 1,239 articles; the second, 203 articles) and other sources (1 article). Based on 1,061 articles—excluding 382 duplicated articles—the eligibility criteria were first applied through titles and abstracts. We then located and read the full texts in cases where it was difficult to obtain accurate information only from the title and abstract. Consequently, we excluded 1,047 articles (98.7%), and 14 studies [25–38] met the inclusion criteria. The study selection process is detailed in the PRISMA 2020 flow diagram (Fig. 1).

Of the 14 studies selected, 8 (57.1%) [25, 26, 30–33, 37] were assessed to have a low risk of bias across all domains and items. In terms of the patient domain, all studies that either had consecutive/random samples or included all participants were assessed as having low bias [26, 30], except for three studies [27, 28, 38]. In the index test domain, the
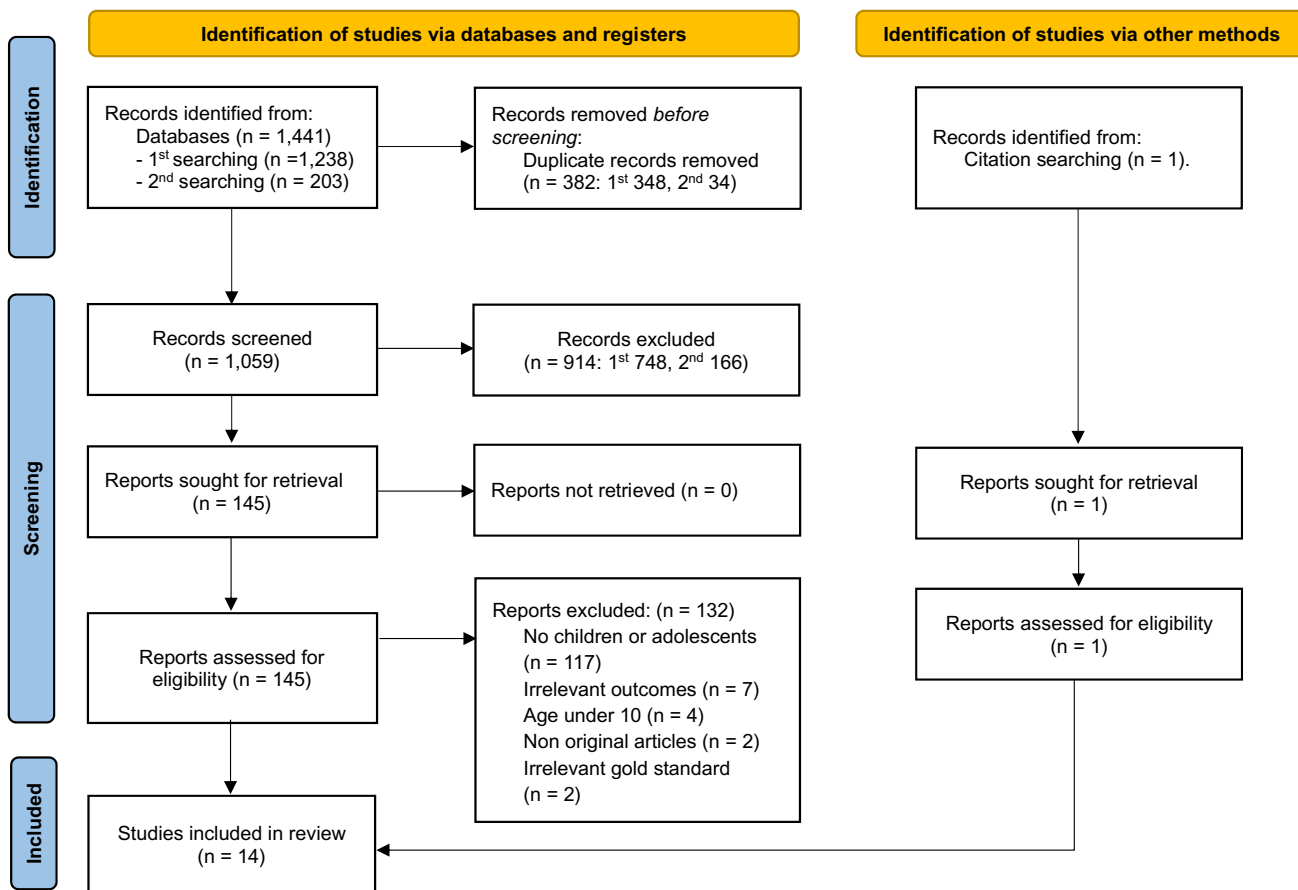

Fig. 1 PRISMA 2020 flow diagram of article selection

risk of bias was low in all the studies. The CES-D is a self-reported measure; therefore, it was assumed that the gold standard for diagnosing depression would not influence its findings. Contrastingly, knowing the CES-D scores in advance may influence the findings of the gold standard. Nine studies [25–28, 30–33, 37] were blinded, and one study [38] used a gold standard before the CES-D. Therefore, the risk of bias was low in these studies. However, four studies were assessed as having an uncertain risk of bias [29, 34–36]. Finally, regarding flow and timing, all studies were assessed as having a low risk of bias (Fig. 2).

Characteristics of Selected Studies

A total of 7,843 participants in 14 studies tested the predictive validity of the CES-D for children and adolescents. The ages of the participants ranged from 10 to 19 years, and the mean age was 14–16 years. The selected studies were published in 10 countries. Three studies were conducted in the United States [30, 36, 38], another three in Colombia [32, 33, 37], and two in Germany [26, 28]. Brazil, China, France, the Netherlands, Rwanda, and Switzerland had one study each. Most of these were large-scale studies with more than 100 participants, except for two studies [37, 38] that included more than 1,000 participants. The 20-item version of the CES-D was used in 13 studies, whereas two studies used the 15-item version [26, 28], and the short 3- and 10-item versions of the CES-D were used in one study [33]. Four studies [29, 30, 34, 38] analyzed the diagnostic accuracy of the CES-D (20-item) using other depression-screening tools. The tools used were as follows: two studies used the BDI [29, 38], one used the Caroll Rating Scale (CRS) [29], one used the Edinburgh Postnatal Depression Scale (EPDS) [30], and one used the Major Depression Inventory (MDI) [34]. The gold standards for diagnosing MDD were direct interviews following the DSM and ICD-10 or semi-structured interviews such as K-SADS, the Diagnostisches Interview bei psychischen Störungen im Kindes- und Jugendalter (Kinder-DIPS), the Mini International Neuropsychiatric Interview for Children and Adolescents (MINI-KID), and the Depression Intensity Scale Circles (DISC). The cut-off scores of the CES-D were 14–30 for the long version and varied without a typical pattern (Table 1).

Predictive Validity of the CES-D in Children and Adolescents

Based on the 14 selected studies, the predictive validity of the CES-D for screening MDD was compared according to the CES-D type and with other depression-screening tools (Table 2).

Predictive validity by the CES-D type

The predictive validity of the long version of the CES-D was assessed for 7,038 children and adolescents in 11 studies (Fig. 3). The prevalence was 23.0%. The sensitivity and specificity ranged from 0.73 to 0.91 and 0.51 to 0.86, respectively. In the meta-analysis, the pooled sensitivity was 0.81 (95% CI, 0.77 to 0.85), pooled specificity was 0.72 (95% CI, 0.68 to 0.76), and RE correlation was 1.000. The sROC AUC was 0.83 (SE = 0.03), and the Q* value was 0.76 (SE = 0.03).

The predictive validity of the short version of the CES-D was conducted on 805 children and adolescents in three studies [26, 28, 33]. The prevalence was 16.0%. Sensitivity ranged from 0.76 to 0.86, and specificity ranged from 0.59 to 0.84. In the meta-analysis, the pooled sensitivity was 0.80 (95% CI, 0.73 to 0.86), pooled
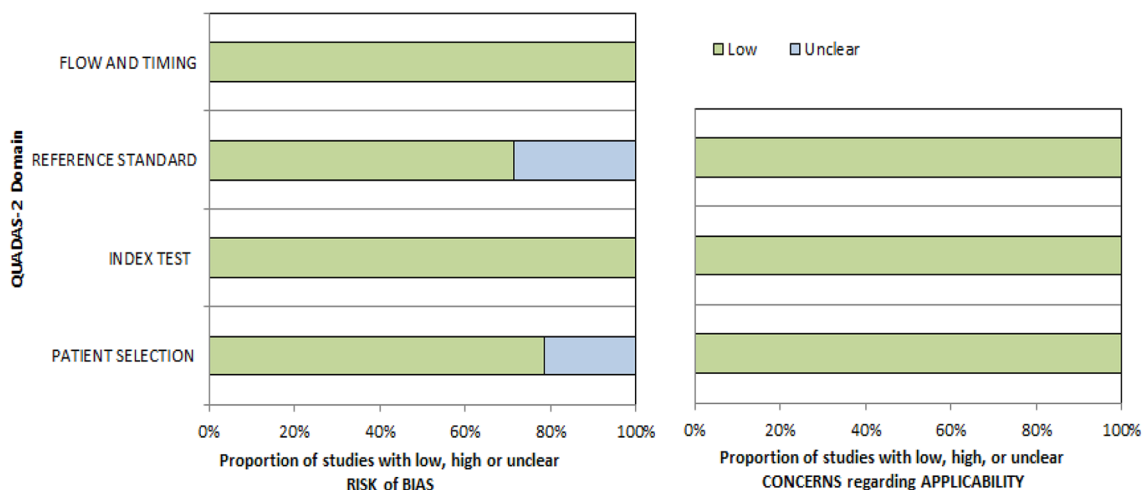

Fig. 2 Quality assessment results of the selected studies by QUADAS-2

Table 1 Characteristics of selected studies

Year	Authors	Location	Participants	Total (n)	Mean age (Range)	Female (%)	Gold standard	Blind	Scale	Cut off	2×2 Table				Value (95% confidence interval)	
											TP	FP	FN	TN	Sensitivity	Specificity
2018	Yang et al. [25]	China	Chinese adolescents	612	14.2 (11–19)	51.3	K-SADS	Yes	CES-D	≥30	37	159	7	409	0.84 (0.71–0.92)	0.72 (0.68–0.76)
2013	Pietsch et al. [26]	Germany	Inpatient and outpatient	327	14.3 (13–16)	59.0	Kinder-DIPS	Yes	CESD-15	≥14	36	46	6	239	0.86 (0.72–0.93)	0.84 (0.79–0.88)
2012	Betancourt et al. [27]	Rwanda	Children and adolescents	367	(10–17)	40.9	MINI-KID	Yes	CES-D	≥30	48	87	10	222	0.83 (0.71–0.90)	0.72 (0.67–0.77)
	Dolle et al. [28]	Germany	Adolescents	88	14.5 (13–16)	58.0	Kinder-DIPS	Yes	CES-D-15	≥12	21	24	4	35	0.84 (0.65–0.94)	0.59 (0.47–0.71)
2011	Salle [29]	Brazil	High school students	442	16.0 (15–17)	59.0	DSM-IV	Unclear	CES-D	≥14	44	92	16	290	0.73 (0.61–0.83)	0.76 (0.71–0.80)
									BDI	≥9	43	99	17	283	0.72 (0.59–0.81)	0.74 (0.69–0.78)
									CRS	≥10	47	99	13	283	0.78 (0.66–0.87)	0.74 (0.69–0.78)
2010	Logsdon et al. [30]	USA	Adolescent mothers	59	16.4 (13–18)	100.0	K-SADS	Yes	CES-D	≥16	7	24	2	25	0.78 (0.45–0.94)	0.51 (0.37–0.64)
									EPDS	≥5	8	20	2	29	0.80 (0.49–0.94)	0.59 (0.45–0.72)
2009	Aebi et al. [31]	Switzerland	Children and adolescents	140	15.5 (11–17)	67.1	ICD-10	Yes	CES-D	≥21	71	8	12	49	0.86 (0.76–0.92)	0.86 (0.75–0.93)
	Camacho et al. [32]	Columbia	Adolescent students	390	14.8 (13–17)	55.6	DSM-IV	Yes	CES-D	≥23	33	91	12	254	0.73 (0.59–0.84)	0.74 (0.69–0.78)
	Rueda-Jaimes et al. [33]	Columbia	Adolescent students	390	14.8 (13–17)	55.6	DSM-IV	Yes	CES-D-10	≥11	35	89	10	256	0.78 (0.64–0.87)	0.74 (0.69–0.79)
									CES-D-3	≥4	34	101	11	244	0.76 (0.61–0.86)	0.71 (0.66–0.75)
2008	Cuijpers et al. [34]	The Netherlands	Adolescent students	243	(14–16)	48.0	MINI (DSM-IV)	Unclear	CES-D	≥22	135	24	14	70	0.91 (0.85–0.94)	0.74 (0.65–0.82)
									MDI	≥19	135	27	14	67	0.91 (0.85–0.95)	0.71 (0.61–0.79)
2002	Chabrol et al. [35]	France	High school students	60	17.0 (15–19)	–	MINI	Unclear	CES-D	≥24	21	9	7	23	0.75 (0.57–0.87)	0.72 (0.55–0.84)
1998	Prescott et al. [36]	USA	High school students	556	16.9 (14–18)	53.8	DISC	Unclear	CES-D	≥16	37	132	10	377	0.79 (0.65–0.88)	0.74 (0.70–0.78)
1991	Garrison et al. [37]	Columbia	High school students	2,465	(11–17)	50.0	K-SADS	Yes	CES-D	≥16	241	861	91	1,272	0.73 (0.68–0.77)	0.60 (0.58–0.62)

Table 1 (continued)

Year	Authors	Location	Participants	Total (n)	Mean age (Range)	Female (%)	Gold standard	Blind	Scale	Cut off	2×2 Table				Value (95% confidence interval)	
											TP	FP	FN	TN	Sensitivity	Specificity
	Roberts et al. [38]	USA	High school students	1,704	16.6 (14–18)	52.9	K-SADS	Yes	CES-D	≥24	36	417	7	1,264	0.84 (0.70–0.92)	0.75 (0.73–0.77)
									BDI	≥11	36	321	7	1,360	0.84 (0.70–0.92)	0.81 (0.79–0.83)

TP true positive, *FP* false positive, *FN* false negative *TN* true negative, *K-SADS* schedule for affective disorders and schizophrenia in school age children, *CES-D* center for epidemiologic studies depression scale, *Kinder-DIPS* diagnostisches interview bei psychischen störungen im kindes- und jugendalter, *MINI-KIDS* mini international neuropsychiatric interview for children and adolescents, *DSM-IV* diagnostic and statistical manual of mental disorders, fourth edition, *BDI* beck depression inventory, *CRS* caroll rating scale, *EPDS* edinburgh postnatal depression scale, *ICD-10* international classification of diseases-tenth revision, *MINI* Mini-international neuropsychiatric interview, *MDI* major depression inventory, *DISC* diagnostic interview schedule for children

Table 2 Summary results of meta-analysis

Subjects	Studies (k)	Subjects (n)	Prevalence (mean)	Pooled diagnostic test accuracy (95% confidence interval)					RE correlation	Summary ROC curve			
				Sensitivity	Specificity	Positive likelihood ratio	Negative likelihood ratio	Diagnostic odds ratio		AUC	SE	Q*	(SE)
Type of the CES-D													
Long version	11	7,038	23.0	0.81 (0.77–0.85)	0.72 (0.68–0.76)	2.91 (2.43–3.49)	0.26 (0.21–0.33)	11.07 (7.39–16.59)	1.000	0.83	0.03	0.76	0.03
Short version	3	805	16.0	0.80 (0.73–0.86)	0.74 (0.65–0.81)	3.06 (2.19–4.27)	0.27 (0.19–0.39)	11.35 (5.85–22.00)	1.000	0.86	0.04	0.79	0.04
Compared with other screening tools													
CES-D	4	2,448	23.0	0.84 (0.72–0.91)	0.72 (0.63–0.79)	2.94 (2.19–3.95)	0.23 (0.13–0.41)	12.85 (5.95–27.74)	0.031	0.83	0.03	0.76	0.02
Other tools				0.83 (0.74–0.89)	0.74 (0.68–0.79)	3.20 (2.56–3.99)	0.24 (0.16–0.36)	13.55 (7.80–23.52)	−0.082	0.83	0.03	0.76	0.02

RE correlation random effect correlation, *ROC curve* receiver operating characteristic curve, *AUC* area under the curve, *SE* standard error

A Sensitivity in the CES-D long version

B Specificity in the CES-D long version

C SROC curve in the CES-D long version

D Sensitivity in the CES-D short version

E Specificity in the CES-D short version

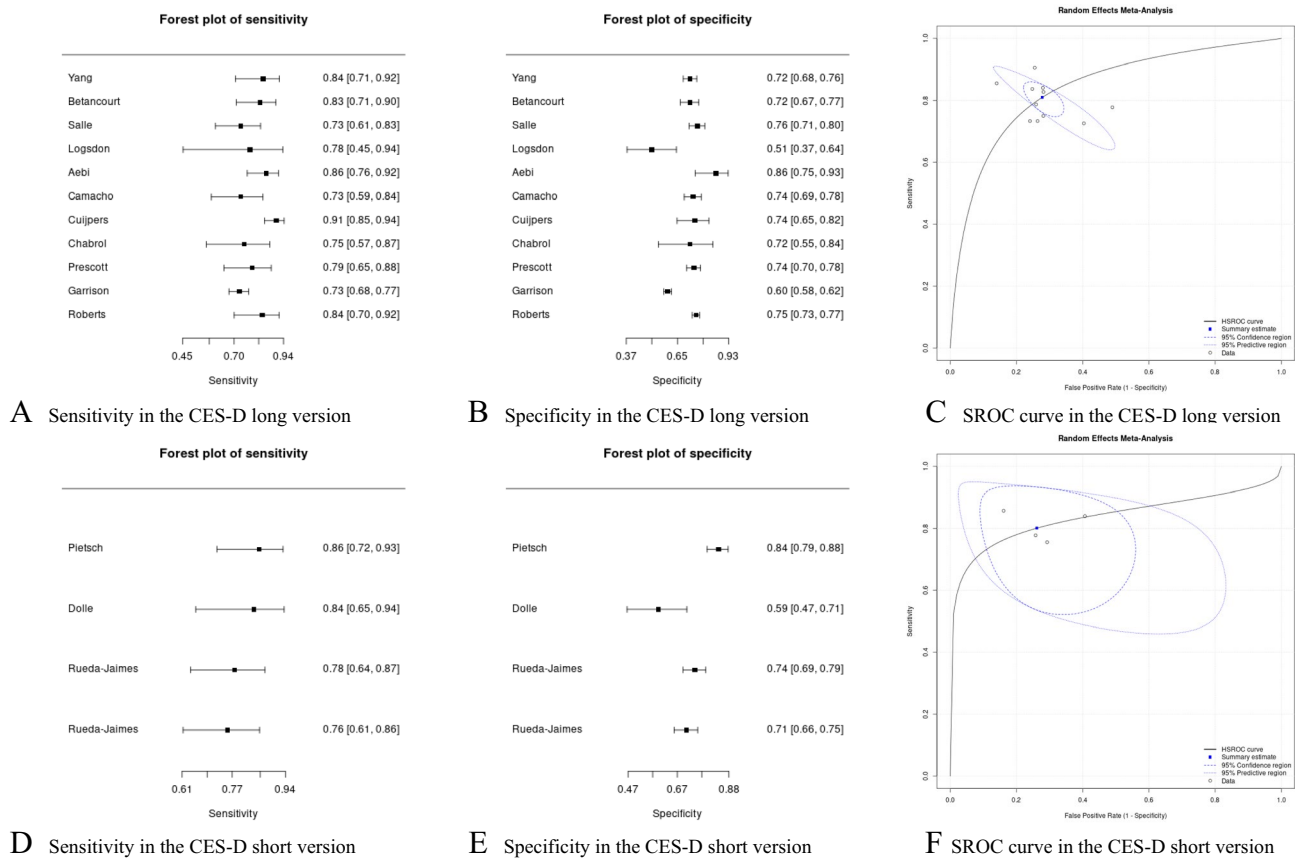**F** SROC curve in the CES-D short version

Fig. 3 Predictive accuracy of the type of the CES-D

specificity was 0.74 (95% CI, 0.65 to 0.81), and RE correlation was 1.000. The sROC AUC was 0.86 (SE = 0.04) and the Q* value was 0.79 (SE = 0.04).

Predictive validity of the CES-D compared to other tools.

The predictive validity of other screening tools compared to the CES-D was assessed for 2,448 children and adolescents in four studies [29, 30, 34, 38] (Fig. 4). The prevalence was 23.0%. The sensitivity and specificity ranged from 0.73 to 0.91 and 0.51 to 0.76, respectively. In a meta-analysis, the pooled sensitivity was 0.84 (95% CI, 0.72 to 0.91), pooled specificity was 0.72 (95% CI, 0.63 to 0.79), and RE correlation was 0.031. The sROC AUC was 0.83 (SE = 0.03), and the Q* value was 0.76 (SE = 0.02).

For the other tools, the sensitivity and specificity ranged from 0.72 to 0.91 and 0.59 to 0.81, respectively. In the meta-analysis, the pooled sensitivity was 0.83 (95% CI, 0.74 to 0.89), pooled specificity was 0.74 (95% CI, 0.68 to 0.79), and RE correlation was -0.082. The sROC AUC was 0.83 (SE = 0.03), and the Q* value was 0.76 (SE = 0.02).

Discussion

Compared to other age groups, one of the best preventive interventions for children and adolescents with a high risk of experiencing crises, such as suicide because of depression, is to quickly and easily screen for MDD using reliable tools, thus, enabling early treatment [39]. This study systematically synthesized and analyzed 14 studies that reported on the diagnostic accuracy of the CES-D, a tool widely used to screen for depression in children and adolescents. This review indicates that the CES-D is a valuable tool for predicting depression in children and adolescents.

In this study, the literature search was conducted twice. The same search terms were used in both searches. A total of 203 articles was identified in the second search, but no new articles met the eligibility criteria. The selected 14 studies were published from 1991 to 2018, but 9 studies were concentrated between 2008 and 2013. Due to the coronavirus disease (COVID-19) pandemic, children and adolescents worldwide have not been able to go to school for the second consecutive year and are studying online without interacting with their peers. Although depression
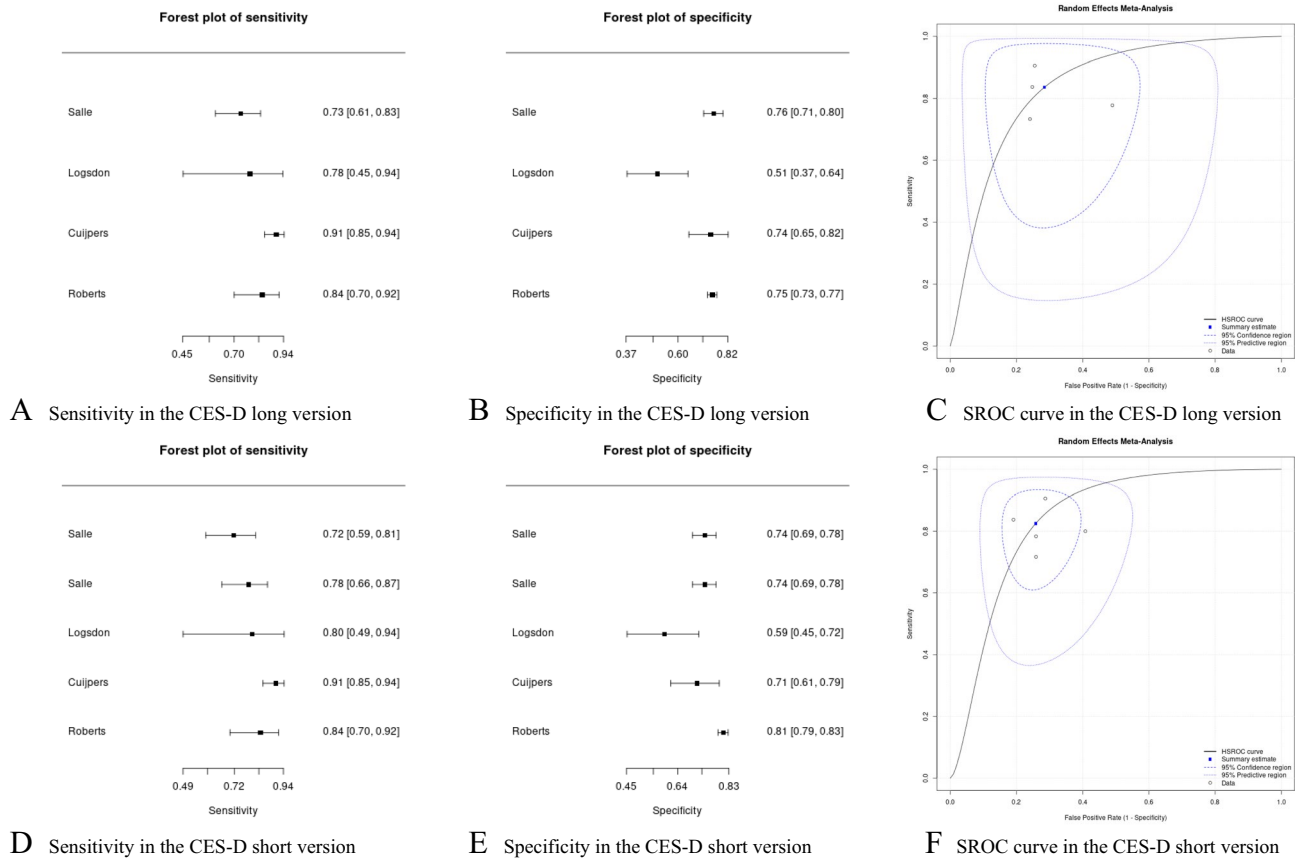
A Sensitivity in the CES-D long version

B Specificity in the CES-D long version

C SROC curve in the CES-D long version

D Sensitivity in the CES-D short version

E Specificity in the CES-D short version

F SROC curve in the CES-D short version

Fig. 4 Predictive accuracy of the CES-D versus other tools

in children and adolescents has not emerged as a severe social issue, it has increased in recent years [40]. Therefore, the demand for screening tests for depression in children and adolescents is relatively high.

Of the 14 studies selected for this review, 11 presented the CES-D long version, and 3 presented the short version. The sensitivity of the CES-D in individual studies was in the range of 0.73–0.91. Nonetheless, the specificity was 0.51–0.86, with some deviations. The meta-analysis showed that the pooled sensitivity and specificity of the CES-D by type were similar: long version (0.81, 0.72) and short version (0.80, 0.74). The sROC AUC was similar in both the long (0.83) and short versions (0.86). In the short version of the CES-D, two 15-item cases and one case each of 10-item and 3-item were analyzed. The sensitivity of the 15-item scale was in the range 0.84–0.86, but the specificity range was 0.59–0.84, showing a difference. The sensitivity of the 10-item (0.78) and 3-item (0.76) scales was somewhat lower than that of the 15-item CES-D. However, there was insufficient evidence to draw any conclusions, as there were only three studies. In the case of the Geriatric Depression Scale used for older adults, the short version tends to be more actively used because of the tool's convenience [41].

Notwithstanding this, the advantages of the short version do not seem to be significant because there is no time constraint for children and adolescents to complete the questionnaire. It is important to note that the 15-item CES-D used in the two studies referred to a tool developed by Pietsch [26]. However, even though the number of items is the same in the short version, the items may differ and must therefore be confirmed [42].

In four studies, the CES-D and other depression-screening tools were compared for their predictive validity. Although the number of studies was insufficient, the CES-D scored similarly with other depression-screening tools regarding pooled sensitivity (0.84 vs. 0.83) and pooled specificity (0.72 vs. 0.74). The sROC AUC and Q* values were also the same at 0.83 and 0.76, respectively. As for the other depression tools, the BDI was used in two cases and the rest in only one case, so a comparative advantage could not be established. The sensitivity and specificity of the CES-D were 0.73–0.86 and 0.75–0.76, respectively, similar to those of the relatively widely used BDI at 0.72–0.84 and 0.74–0.81, respectively. This is in line with the results of existing SRs [14, 15]. The meta-analysis results for the CES-D alone were not reported, but the sensitivity of the four tools (BDI, CES-D, etc.) as

reported by Stockings et al. was 0.80, and their specificity was 0.78 [14]. Therefore, the CES-D is similar to the BDI proposed by USPSTF regarding screening accuracy and can be considered for use in children and adolescents [8].

This review had some limitations. None of the 14 selected studies showed high risk, but all were observational and not experimental. RE correlations were all positive, except for other scales, and there was heterogeneity among the studies. The specificity of the CES-D was lower than its sensitivity. However, because the CES-D is a screening tool, a higher sensitivity may be more valuable than specificity. In Radloff's version of the CES-D, the recommended cut-off score for distinguishing the general population from psychiatric patients was 16; however, the sensitivity and specificity were not presented in that study [12]. Among the studies included in this review, there were no similarities found in the cut-off scores, with two studies [25, 27] indicating a cut-off score of 30 or higher, five studies [31, 32, 34, 35, 38] showing a score of 21–24, three studies [30, 36, 37] indicating a score of 16, and one study [29] indicating a score of 14. Therefore, this review did not include a subgroup analysis according to cut-off scores.

Summary

The prevention of depression in children and adolescents during physical and mental development is a significant public health issue. This study aimed to explore the predictive validity of each CES-D type and the screening utility of the CES-D in comparison with other tools. Fourteen studies met the eligibility criteria for inclusion. In 11 studies, the CES-D long version was measured, and in 3 studies, the predictive validity of the CES-D short version was measured. As a result of the meta-analysis, the pooled sensitivity (0.81 and 0.80) and pooled specificity (0.72 and 0.74) of the CES-D by type were similar. The sROC AUC was 0.83 for the long version and 0.86 for the short version. Four studies compared the CES-D to other tools. The pooled sensitivity (0.84 and 0.83) and specificity (0.72 and 0.74), including the sROC AUC, were similar at 0.83. The CES-D has shown consistent moderate predictive validity, with moderate or high accuracy in identifying MDD in children and adolescents. The CES-D can be used, like other depression-screening tools, such as the BDI, to screen for depression in children and adolescents. Depression during childhood and adolescence is strongly associated with recurrent depression in adulthood. These findings may help in the early detection and rapid recognition of depression in children and adolescents and help them maintain a healthy life through appropriate treatment.

Declarations

References

1. Ghandour RM, Sherman LJ, Vladutiu CJ, Ali MM, Lynch SE, Bitsko RH, Blumberg SJ (2019) Prevalence and treatment of depression, anxiety, and conduct problems in US children. J Pediatr 206:256–267. https://doi.org/10.1016/j.jpeds.2018.09.021

2. Bansal V, Goyal S, Srivastava K (2009) Study of prevalence of depression in adolescent students of a public school. Ind Psychiatry J 18:43–46. https://doi.org/10.4103/0972-6748.57859

3. Bitsko RH, Claussen AH, Lichstein J, Black LI, Jones SE, Danielson ML, Hoenig JM, Davis Jack SP, Brody DJ, Gyawali S, Maenner MJ, Warner M, Holland KM, Perou R, Crosby AE, Blumberg SJ, Avenevoli S, Kaminski JW, Ghandour RM (2022) Mental health surveillance among children - United States, 2013–2019. MMWR. https://doi.org/10.15585/mmwr.su7102a1

4. Thapar A, Collishaw S, Pine DS, Thapar AK (2012) Depression in adolescence. Lancet 379:1056–1067. https://doi.org/10.1016/S0140-6736(11)60871-4

5. Yu R, Aaltonen M, Branje S, Ristikari T, Meeus W, Salmela-Aro K, Goodwin GM, Fazel S (2017) Depression and violence in adolescence and young adults: findings from three longitudinal cohorts. J Am Acad Child Adolesc Psychiatry 56:652-658.e1. https://doi.org/10.1016/j.jaac.2017.05.016

6. Maughan B, Collishaw S, Stringaris A (2013) Depression in childhood and adolescence. J Can Acad Child Adolesc Psychiatry 22:35–40

7. Birmaher B, Williamson DE, Dahl RE, Axelson DA, Kaufman J, Dorn LD, Ryan ND (2004) Clinical presentation and course of depression in youth: does onset in childhood differ from onset in adolescence? J Am Acad Child Adolesc Psychiatry 43:63–70. https://doi.org/10.1097/00004583-200401000-00015

8. Siu AL, US Preventive Services Task Force (2016) Screening for depression in children and adolescents: US Preventive Services Task Force recommendation statement. Pediatrics 137:e20154467. https://doi.org/10.1542/peds.2015-4467

9. World Health Organization. (2017) Depression and other common mental disorders: global health estimates. https://www.who.int/mental_health/management/depression/prevalence_global_health_estimates/en/. Accessed 20 January 2022

10. Bhatia SK, Bhatia SC (2007) Childhood and adolescent depression. Am Fam Physician 75:73–80

11. Roseman M, Saadat N, Riehm KE, Kloda LA, Boruff J, Ickowicz A, Baltzer F, Katz LY, Patten SB, Rousseau C, Thombs BD (2017) Depression screening and health outcomes in children and adolescents: a systematic review. Can J Psychiatry 62:813–817. https://doi.org/10.1177/0706743717727243

12. Radloff LS (1977) The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1:385–401

13. Carleton RN, Thibodeau MA, Teale MJ, Welch PG, Abrams MP, Robinson T, Asmundson GJ (2013) The center for epidemiologic studies depression scale: a review with a theoretical and empirical examination of item content and factor structure. PLoS ONE 8:e58067. https://doi.org/10.1371/journal.pone.0058067

14. Stockings E, Degenhardt L, Lee YY, Mihalopoulos C, Liu A, Hobbs M, Patton G (2015) Symptom screening scales for detecting major depressive disorder in children and adolescents: a systematic review and meta-analysis of reliability, validity and diagnostic utility. J Affect Disord 174:447–463. https://doi.org/10.1016/j.jad.2014.11.061

15. Roseman M, Kloda LA, Saadat N, Riehm KE, Ickowicz A, Baltzer F, Katz LY, Patten SB, Rousseau C, Thombs BD (2016) Accuracy of depression screening tools to detect major depression in children and adolescents: a systematic review. Can J Psychiatry 61:746–757. https://doi.org/10.1177/0706743716651833. (**Epub 2016 Jul 9**)

16. Bossuyt PM, Davenport C, Deeks JJ, Gatsonis C, Wisniewski S (2013) Cochrane handbook for systematic reviews of diagnostic test accuracy. The Cochrane Collaboration. https://methods.cochrane.org/sdt/handbook-dta-reviews. Accessed on December 15, 2021

17. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. https://doi.org/10.1136/bmj.n71

18. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM, QUADAS-2 Group (2011) QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 155:529–536. https://doi.org/10.7326/0003-4819-155-8-201110180-00009

19. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A (2006) Meta-Disc: a software for meta-analysis of test accuracy data. BMC Med Res Methodol 6:31. https://doi.org/10.1186/1471-2288-6-31

20. Patel A, Cooper N, Freeman S, Sutton A (2021) Graphical enhancements to summary receiver operating characteristic plots to facilitate the analysis and reporting of meta-analysis of diagnostic test accuracy data. Res Synth Methods 12:34–44. https://doi.org/10.1002/jrsm.1439

21. Freeman SC, Kerby CR, Patel A, Cooper NJ, Quinn T, Sutton AJ (2019) Development of an interactive web-based tool to conduct and interrogate meta-analysis of diagnostic test accuracy studies: MetaDTA. BMC Med Res Methodol 19:81. https://doi.org/10.1186/s12874-019-0724-x

22. Negeri ZF, Shaikh M, Beyene J (2018) Bivariate random-effects meta-analysis models for diagnostic test accuracy studies using arcsine-based transformations. Biom J 60:827–844. https://doi.org/10.1002/bimj.201700101

23. Greiner M, Pfeiffer D, Smith RD (2000) Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. Prev Vet Med 45:23–41. https://doi.org/10.1016/s0167-5877(00)00115-x

24. Walter SD (2002) Properties of the summary receiver operating characteristic (SROC) curve for diagnostic test data. Stat Med 21:1237–1256. https://doi.org/10.1002/sim.1099

25. Yang W, Xiong G, Garrido LE, Zhang JX, Wang MC, Wang C (2018) Factor structure and criterion validity across the full scale and ten short forms of the CES-D among Chinese adolescents. Psychol Assess 30:1186–1198. https://doi.org/10.1037/pas0000559

26. Pietsch K, Allgaier AK, Frühe B, Hoyler A, Rohde S, Schulte-Körne G (2013) Screening for adolescent depression in paediatric care: validity of a new brief version of the Center for Epidemiological Studies Depression Scale. Child Adolesc Ment Health 18:76–81. https://doi.org/10.1111/j.1475-3588.2012.00650.x

27. Betancourt T, Scorza P, Meyers-Ohki S, Mushashi C, Kayiteshonga Y, Binagwaho A, Stulac S, Beardslee WR (2012) Validating the center for epidemiological studies depression scale for children in Rwanda. J Am Acad Child Adolesc Psychiatry 51:1284–1292. https://doi.org/10.1016/j.jaac.2012.09.003

28. Dolle K, Schulte-Körne G, von Hofacker N, Izat Y, Allgaier AK (2012) Agreement of clinical diagnosis, structured interviews, and self-report questionnaires for depression in children and adolescents. Z Kinder Jugendpsychiatr Psychother 40:405–414. https://doi.org/10.1024/1422-4917/a000200

29. Salle E, Rocha NS, Rocha TS, Nunes C, Chaves MLF (2012) Depression rating scales as screening tools for depression in high school students. Revista de Psiquiatria Clinica 39:24–27. https://doi.org/10.1590/S0101-60832012000100005

30. Logsdon MC, Myers JA (2010) Comparative performance of two depression screening instruments in adolescent mothers. J Womens Health 9:1123–1128. https://doi.org/10.1089/jwh.2009

31. Aebi M, Metzke CW, Steinhausen HC (2009) Prediction of major affective disorders in adolescents by self-report measures. J Affect Disord 115:140–149. https://doi.org/10.1016/j.jad.2008.09.017

32. Camacho PA, Rueda-Jaimes GE, Latorre JF, Navarro-Mancilla AA, Escobar M, Franco JA (2009) Validity and reliability of the center for epidemiologic studies-depression scale in Colombian adolescent students. Biomedica 29:260–269

33. Rueda-Jaimes GE, Camacho López PA, Rangel-Martínez-Villalba AM (2009) Validation of two short versions of the Centre for epidemiological studies depression scale in Colombian adolescents. Aten Primaria 41:255–261. https://doi.org/10.1016/j.aprim.2008.09.005

34. Cuijpers P, Boluijt P, van Straten A (2008) Screening of depression in adolescents through the Internet: sensitivity and specificity of two screening questionnaires. Eur Child Adolesc Psychiatry 17:32–38. https://doi.org/10.1007/s00787-007-0631-2

35. Chabrol H, Montovany A, Chouicha K, Duconge E (2002) Study of the CES-D on a sample of 1,953 adolescent students. Encephale 28:429–432

36. Prescott CA, McArdle JJ, Hishinuma ES, Johnson RC, Miyamoto RH, Andrade NN, Edman JL, Makini GK Jr, Nahulu LB, Yuen NY, Carlton BS (1998) Prediction of major depression and dysthymia from CES-D scores among ethnic minority adolescents. J Am Acad Child Adolesc Psychiatry 37:495–503. https://doi.org/10.1097/00004583-199805000-00012

37. Garrison CZ, Addy CL, Jackson KL, McKeown RE, Waller JL (1991) The CES-D as a screen for depression and other psychiatric disorders in adolescents. J Am Acad Child Adolesc Psychiatry 30:636–641. https://doi.org/10.1097/00004583-199107000-00017

38. Roberts RE, Lewinsohn PM, Seeley JR (1991) Screening for adolescent depression: a comparison of depression scales. J Am Acad Child Adolesc Psychiatry 30:58–66. https://doi.org/10.1097/00004583-199101000-00009

39. Kieling C, Adewuya A, Fisher HL, Karmacharya R, Kohrt BA, Swartz JR, Mondelli V (2019) Identifying depression early in adolescence. Lancet Child Adolesc Healthr 3:211–213. https://doi.org/10.1016/S2352-4642(19)30059-8

40. Racine N, McArthur BA, Cooke JE, Eirich R, Zhu J, Madigan S (2021) Global prevalence of depressive and anxiety symptoms in children and adolescents during COVID-19: a meta-analysis. JAMA Pediatr 175:1142–1150. https://doi.org/10.1001/jamapediatrics.2021.2482

41. Durmaz B, Soysal P, Ellidokuz H, Isik AT (2018) Validity and reliability of geriatric depression scale-15 (short form) in Turkish older adults. North Clin Istanbal 5:216–220. https://doi.org/10.14744/nci.2017.85047

42. González P, Nuñez A, Merz E, Brintz C, Weitzman O, Navas EL, Camacho A, Buelna C, Penedo FJ, Wassertheil-Smoller S et al (2017) Measurement properties of the center for epidemiologic studies depression scale (CES-D 10): findings from HCHS/SOL. Psychol Assess 29:372–381. https://doi.org/10.1037/pas0000330