Contents lists available at ScienceDirect

Psychiatry Research

Diagnostic Accuracy of "System for the Assessment of Children and Adolescents" (SENA) for Emotional Disorders in Youth: Insights from the EmoChild Study in Spain

Jose A. Piqueras [a,*] , Raul Castaño [a], Laura Ballester [b,d] , Gemma Vilagut [b,d] ,
Jordi Alonso [b,c,d] , Alexandra Morales [a], Jose P. Espada [a], Mireia Orgiles [a]

[a] Miguel Hernández University, Department of Health Psychology, Avda. de la Universidad s/n, Elche (Alicante) 03202, Spain
[b] Health Services Research Unit, IMIM-Hospital del Mar Medical Research Institute, C/ del Doctor Aiguader, 88 08003, Barcelona, Spain
[c] Department of Medicine and Life Sciences, Pompeu Fabra University (UPF), Carrer del Doctor Aiguader, 88 08003, Barcelona, Spain
[d] CIBER en Epidemiología y Salud Pública (CIBERESP), Carrer del Doctor Aiguader, 88 08003, Barcelona, Spain

ABSTRACT

Objective: To evaluate the diagnostic accuracy of the System for the Assessment of Children and Adolescents (SENA) in identifying emotional disorders and suicidal thoughts and behaviors. SENA is a widely used Spanish screening tool for assessing emotional and behavioral symptoms in youth.
Method: 526 primary and secondary pupils aged 8-16 in several Spanish regions (a subset of the EmoChild Project, n=5,652 completed SENA, and underwent with clinical interviews using KSADS-COMP within 3 months. We screened potential participants by identifying possible positive and negative cases based on a T-score equivalent to the 75th percentile and served as a cut-off point for accessing a diagnostic interview.
Results: SENA's Emotional subscale showed a sensitivity (SN) of 78.5% in children and 80.4% in adolescents for detecting any emotional disorder, with specificities (SP) of 65.9% and 66.1%, and Area Under the ROC curve (AUC) of .74 in children and .73 in adolescents. AUCs were adequate for all subscales (0.72-0.93) other than obsessive compulsive disorder (AUC=.67). Specific subscales performed best: Generalized Anxiety Disorder (SN=100%, SP=88.4%) and Social Anxiety Disorder-SAD (SN = 91.4%, SP = 80.3%) in children, and SAD (SN=88.1%, SP=72.5%), Post-Traumatic Stress Disorder (SN=81.2%, SP=68.2%) and suicidal thoughts in adolescents (SN=84.5%, SP=73.8%).
Conclusions: SENA is a valuable screening tool for educational and clinical settings, facilitating early intervention through a standardized and user-friendly assessment. Nevertheless, there is need for refined thresholds to enhance specificity and clinical alignment.

1. Introduction

Globally, mental problems in children and adolescents range from 11.6% to 15.5%, with emotional disorders (anxiety and depression) contributing the most to disability in this age group (Kieling et al., 2024; Polanczyk et al., 2015; Sacco et al., 2024; WHO, 2024). Furthermore, mostly emotional disorders begin before the age of 14 (Solmi et al., 2022). According to the studies, anxiety and depression prevalence is high (3.4%–7.9%, and 1.9%–2.6%, respectively) and increased by 27% and 24% between 2016 and 2020 (Lebrun-Harris et al., 2022). In Spain,

mental disorders affect 13.6% of children and adolescents (anxiety:8.6%; depression:5%), with prevalence rates rising with age (Subdirección General de Información Sanitaria, 2021). Additionally, suicide has become the second leading cause of death among young people, with 87 cases reported in Spain in 2022 (INE, 2023).

Subclinical symptoms notably increase the reported prevalence of behavioural and emotional issues among children and adolescents. In China, substantial rates of behavioural and emotional problems (17.5%) were identified in school-aged children and adolescents using the Achenbach Child Behaviour Checklist (CBCL) (Cui et al., 2021; Li et al.,

---

2022). In the U.S., the prevalence of externalizing and internalizing disorders ranged from 14.8 and 34.4% using the Behavior Assessment System for Children (BASC-2) tool and DSM-IV criteria (Danielson et al., 2021). French school-aged children reported symptoms consistent with at least one probable mental health disorder (28.7%) (Seconda et al., 2024). Globally, 34% of adolescents exhibit elevated depressive symptoms, including subclinical manifestations (Shorey et al., 2022). In Spain, 11.9% of adolescents reported depressive symptoms (e.g., loss of interest), 19.6% experienced persistent worry or anxiety, and 21.4%, excessive worry on most days (Casares et al., 2024).

The information emphasizes the importance of evidence-based assessment tools that can comprehensively capture emotional problems, including both clinical and subclinical symptoms, through dimensional models. These models, which assess symptoms across a continuum of severity, provide valuable insights into mental health challenges by addressing symptom variability and comorbidity (Haslam et al., 2020; Gore & Widiger, 2015) and are particularly effective in guiding early interventions and personalizing treatments, as they offer transdiagnostic perspectives essential for understanding and addressing a broad range of emotional disorders. Furthermore, their application underscores the need for culturally adapted screening methods to meet the diverse needs of populations and tackle the increasing prevalence of emotional disorders, especially among vulnerable groups like children and adolescents, as highlighted in UNICEF's "Mind Matters" report (Sherr et al., 2021).

Dimensional and hierarchical screening tools such as the Achenbach System for Empirical Assessment (ASEBA; Achenbach & Rescorla, 2001) and dimensional tools such as the BASC (Reynolds & Kamphaus, 1992), the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997), and the Child and Adolescent Assessment System (Sistema de Evaluación de Niños y Adolescentes; SENA; Fernández-Pinto et al., 2015a, 2015b) demonstrate diagnostic utility, with high sensitivity and specificity for effectively identifying emotional disorders at varying levels of severity. For instance, BASC and ASEBA demonstrate high diagnostic validity (sensitivity and specificity around 85% and 90%, respectively) while the SDQ is especially effective for rapid screening despite its limitations in diagnostic precision with 70-80% of sensitivity, as the information provided by parents may be biased when it comes to reporting internalizing problems (Achenbach et al., 2001; Duncan et al., 2024; Piqueras & Carrasco, 2018; Reynolds et al., 2004). However, variations in diagnostic performance exist based on context and population. Armitage et al. (2023) evaluated the SDQ-Emotional subscale (SDQ-E) in young people, finding moderate diagnostic accuracy for Major Depressive Disorder (AUC=.67–.85) but substantial accuracy for Generalized Anxiety Disorder (AUC=.80–.93) and any anxiety disorder (AUC=.74–.83); while a study conducted among in children with epilepsy, reported SDQ sensitivity and specificity of 0.78 and 0.76. Similarly, Kerns et al. (2015) evaluated the use of BASC in identifying anxiety in children with autism spectrum disorder showing greater diagnostic sensitivity (.47) in parent-reported scores than in child self-reports (.13). These findings emphasize the importance of informants and population-specific factors in mental health screening.

One such tool is the SENA, a multidimensional screening instrument specifically developed for and culturally adapted to Spanish-speaking populations. It is widely used in Spain to assess emotional and behavioural problems, contextual factors, and both protective and vulnerability factors. Unlike international tools such as ASEBA or BASC, which require translation and cultural adjustment, the SENA was designed from the outset for the Spanish context, through an iterative process involving Spanish clinicians, educators, and researchers. This ensured that the content was contextually relevant to the educational, familial, and sociocultural environments of Spain, incorporating culturally adapted language, locally meaningful examples, and stressors specific to Spanish youth (e.g., academic pressure, peer conflict, family dynamics).

Additionally, compared to shorter screening tools such as the Strengths and Difficulties Questionnaire (SDQ), which is efficient but has more limited diagnostic precision, the SENA offers a more detailed dimensional framework, separate child and adolescent as well as parent and teacher self-report versions, and comprehensive interpretative indices.

Furthermore, the SENA has been validated using large normative samples of Spanish youth, strengthening its psychometric robustness and ecological validity. Thus, the SENA has demonstrated excellent psychometric properties in prior studies (Fernández-Pinto et al., 2015a, 2015b) and it is ranked as one of the top 25 tools used by Spanish psychologists according to the 2020 survey of psychology professionals (Muñiz et al., 2020). Thus, The SENA has been shown to have high reliability and internal consistency, as coefficient alpha values for the main emotional and behavioural problems were generally above .80 and frequently above .90. Only a few specific scales had values below .70. The test-retest reliability was also above .70. Despite its extensive use and strong psychometric properties, the SENA's diagnostic accuracy (sensitivity and specificity) and for emotional disorders has not been rigorously validated against a clinical gold standard, as test results have not been cross-checked with clinical interviews. This represents a significant gap in its evaluation, particularly given the tool's potential for guiding mental health interventions in educational and clinical settings.

The present study seeks to address this gap by evaluating the diagnostic accuracy of the SENA's emotional disorder scales, including its global index of emotional problems (EMO) and 6 subscale scores (Depression: DEP, Anxiety: ANS, Social Anxiety: ASC, Somatic Complaints: SOM, Post-Traumatic Symptomatology: PTS, and Obsession-Compulsion: OBS), using the Kiddie Schedule for Affective Disorders and Schizophrenia – Computerized Version (KSADS-COMP) as the gold standard (including the diagnoses of Major Depressive Episode: MDE, Generalized Anxiety Disorder: GAD, Social Anxiety Disorder: SAD, Post-Traumatic Stress Disorder: PTSD, and Obsessive-Compulsive Disorder: OCD). This study for the first time aims to provide robust evidence for its utility in screening for emotional disorders and guiding targeted interventions.

Since the SENA follows a dimensional model, whereas the diagnostic interview adheres to a taxonomic or categorical framework based on the DSM-5, we expected to find acceptable sensitivity and specificity indices (above 70%) in the SENA scores focused on general and specific emotional symptoms.

2. Methods

2.1. Participants

2.1.1. The EMO-CHILD study

The EMO-CHILD project aims to assess the mental health of children and adolescents across Spain, and it is led by the AITANA research group and the Childhood Research Centre of the Universidad Miguel Hernández. From October 2023 to March 2024, children and adolescents aged 8 to 16 years from primary and secondary schools located across the 17 Spanish Autonomous Communities (i.e., regions with devolved administrative authority) were evaluated. The project focuses on internalized mental disorders (emotional and related disorders) and associated variables such as alcohol and substance use, use of new technologies, as well as academic, familial, and social risk and protective factors.

A total of 5,652 students from 108 educational institutions across Spain participated in an online survey via LimeSurvey ©. The sample included 2,343 children (grades 4 to 6 of Primary Education, aged 8–12) and 3,309 adolescents (grades 1 to 4 of Secondary Education, aged 12–16). Participants were excluded if they had missing data (17 children, 55 adolescents) or if their responses were flagged by the assessment tool's validity scales (35 children, 95 adolescents). The use of self-report measures in children under 11 years old is supported by previous studies validating such instruments in this age group (e.g., Achenbach et al., 2001; Piqueras & Carrasco, 2018). Moreover, the child version of the SENA has been psychometrically validated for children aged 6 to 12

(Fernández-Pinto et al., 2015a, 2015b).

The final sample comprised 5,450 participants, with 42% children and 58% adolescents. Gender distribution was similar in both groups: 50.5% of children and 50.7% of adolescents identified as female, 49.1% and 48.4% as male, and 0.3% and 0.8%, respectively, identified as another gender. The mean age was 10.22 years (SD=0.95) for children and 13.65 years (SD=1.22) for adolescents. In terms of academic level, children were distributed across 4th (33.3%), 5th (30.3%), and 6th grade (36.4%), while adolescents were in 1st (27.3%), 2nd (29.6%), 3rd (23.6%), and 4th grade (19.5%) of Secondary Education.

2.1.2. Participant Selection for Clinical Interviews

The sample for the present study included 526 students (266 primary and 260 secondary) belonging to the EMO-CHILD study, who volunteered for clinical interviews. These interviews were regionally distributed as follows: Andalusia (32 primary, 65 secondary), Valencian Community (152 primary, 105 secondary), Community of Madrid (54 primary, 86 secondary), and Region of Murcia (28 primary, 4 secondary).

For the study, schools were selected based on a convenience sampling method, with the following criteria: 1. for feasibility, four Autonomous Communities were randomly selected; 2. Subsequently, all schools within these four selected autonomous communities were contacted and offered the opportunity to participate. Schools that voluntarily agreed to participate in this phase were chosen; 3. Finally, to increase diagnostic yield, we oversampled participants scoring above the 75th percentile on the SENA emotional scores and ensured that the time interval between SENA administration and the KSADS-COMP interview remained below 3 months, this is the the SENA and KSADS administration were separated by three months as average.

In terms of participant selection using SENA cut-points, we followed the recommendations of dimensional models that classify symptoms using percentile-based thresholds: scores above the 90th percentile were classified as 'clinical', scores between the 80th and 90th percentiles as 'at risk' and scores below the 80th percentile as 'non-clinical' (Achenbach et al., 2017; Goodman et al., 2000). Similarly, SENA uses T-scores to identify varying degrees of impairment, with scores of 60-69 indicating caution, 70-79 clinically significant symptoms, and scores above 80 suggesting severe symptoms requiring immediate intervention.

We selected potential participants for the clinical reappraisal subsample by identifying potential positive and negative cases based on a T-score equivalent to the 75th percentile. This threshold was used to identify individuals at higher risk for emotional symptoms (anxiety and/or depression) and served as a cut-off point for accessing a diagnostic interview. According to TEA Editions normative data, this 75th percentile corresponds to a T-score of 63, which roughly aligns with the 90th percentile when the distribution of EMO scores is normal. In our sample, the 75th percentile corresponded to a T-score of around 56 on the Emotional Disturbance Scale (EMO), participants with this score or higher were flagged as potential clinical cases for further clinical evaluation. This threshold was defined to optimise diagnostic sensitivity by identifying individuals with clinically significant symptoms and minimising false positives. Of the 526 students, 104 children and 136 adolescents met this criterion.

A wider table was created (Annex 4) to document T scores for emotional scales (EMO, DEP, ANS, ASC, SOM, PTS, OBS for adolescents) and the suicidal ideation item, following T scores generated from the complete sample of the EmoChild Study.

The following section describes the assessment tools used in the study.

2.2. Measures

2.2.1. Child and Adolescent Assessment System (Sistema de Evaluación de Niños y Adolescentes; SENA; Fernández-Pinto et al., 2015a, 2015b)

The SENA is a comprehensive self-report measure designed to assess mental health in children and adolescents aged 6 to 18 years. The tool differentiates by age group, with versions for children (134 items) and adolescents (188 items). Both evaluate internalizing (e.g., Depression: DEP, Anxiety: ANS, Social Anxiety: ASC, Somatic Complaints: SOM, Post-Traumatic Symptomatology: PTS, and Obsession-Compulsion: OBS) and externalizing (e.g., Inattention, Hyperactivity) problems, as well as contextual and vulnerability factors (e.g., Emotional Regulation, Family Problems). Responses are scored on a Likert scale ranging from "Never or almost never = 0" to "Always or almost always = 4", generating the following indices: the Global Index of Problems, Emotional Problems Index, Behavioral Problems Index, Executive Function Problems Index, Contextual Problems Index, and Resources Index. For the present study, we specifically focused on the Emotional Problems Index and its associated symptom domains. This measure was utilized in this study with the approval of TEA Ediciones.

The ordinal alpha coefficients (α_o) for the SENA subscales among children in this sample were as follows: Depression (DEP) = .85, Anxiety (ANS) = .90, Social Anxiety (ASC) = .86, Somatic Complaints (SOM) = .81, and Post-Traumatic Symptomatology (PTS) = .82. Among adolescents, ordinal alphas were: DEP = .97, ANS = .95, ASC = .92, SOM = .88, PTS = .89, and Obsession-Compulsion (OBS) = .77.

2.2.2. The Clinician-Administered KSADS-COMP (Kiddie Schedule for Affective Disorders and Schizophrenia for school-age children-Computerized; Townsend et al., 2020)

The clinical interview is the gold standard tool for diagnosing mental health disorders in children and adolescents. Therefore, the computerized version of the KSADS-PL-5 (De la Peña et al., 2018) was used, the KSADS-COMP, a web-based version of the K-SADS-PL-5 that retains the content and diagnostic structure of the original interview, aligned with DSM-5 criteria. Its computerized format incorporates automated branching logic (skip patterns) and integrated scoring algorithms, which enhance standardization, reduce interviewer error, and streamline administration without compromising diagnostic validity. This instrument offers a semi-structured format including three main components: an introductory interview, a screening interview, and a supplement interview. The introductory interview gathers essential background information such as life circumstances, adaptive functioning, treatment history, and family history. The screening interview assesses 2 to 4 key symptoms for each psychiatric diagnosis, and the supplement interview provides detailed diagnostic information for disorders that meet the threshold based on the screening.

The KSADS-COMP has demonstrated excellent criterion validity, with strong agreement with diagnoses made by child-adolescent psychiatrists, as well as high sensitivity, specificity, and predictive values (Lee et al., 2023). Its computerized format enhances the efficiency and accuracy of the diagnostic process. Although a validated Spanish version of the KSADS-COMP is not yet available, the Spanish translation is accessible on the platform, and its predecessor, the KSADS-PL-5, has been validated in Spanish-speaking populations (De la Peña et al., 2018). The KSADS-COMP offers several advantages over the original KSADS. As a web-based tool, it streamlines assessment time and provides automated scoring and diagnostic algorithms, enhancing efficiency and accuracy in evaluating psychiatric disorders in children and adolescents. Additionally, the KSADS-COMP includes self-administered versions for both parents and youths, facilitating comprehensive data collection from multiple informants (Towsend et al., 2020).

To maintain high quality and consistency in interviews for the study, we adopted three actions:

i) All interviewers were Master's students in Clinical and Health Psychology (General Health Psychology [Psicología General Sanitaria], the Master's degree qualifying for the practice of private psychology in Spain) or similar programs (such as the Master's degree in Psychological Therapy with Children and Adolescents), with prior training in semi-structured diagnostic interviewing and specific supervision in the administration of the KSADS-COMP.

ii) Two additional training sessions, delivered both in person and online, were held for interviewers. These sessions included instruction and supervised practice in semi-structured diagnostic interviewing, with a specific focus on the administration of the KSADS-COMP. In addition, they received training in risk detection and management procedures to ensure the appropriate handling of cases involving suicidal ideation or other clinical alerts.

iii) To enhance standardization, the wording of the interview items was refined, and a standardized script (Annex 1) was developed to ensure uniformity across interviews. The interviews were conducted online with at least one research technician present to supervise the process. The technician facilitated the setup of devices (tablets or computers), connected the students to the Meet platform, and explained the procedure. Students were provided with a response sheet (Annex 5) and instructed to use headphones during the interview.

The average interview duration was 28:57 minutes for primary school students and 40:11 minutes for secondary students, with standard deviations of 21:75 minutes and 22:31 minutes, respectively. If any risk cases were identified during the interview, they were immediately reported to the technician, who then addressed the issue after the session.

2.3. Data analysis

Analyses were performed separately for children and adolescents. Prevalence estimates for emotional disorders were calculated based on SENA results using the 75th percentile (T-scores ≥ 56) for each emotional scale score and emotional disorder index and also based on KSADS-COMP. McNemar's χ^2 test was used to test prevalence differences between two measures.

To assess the diagnostic accuracy of SENA for diagnosing emotional disorders based on DSM-5 criteria, receiver operating characteristic (ROC) curves were calculated, highlighting how well SENA discriminated between positive and negative cases compared to the reference standard. These curves measured the precision of SENA in relation to clinical interview outcomes, focusing on key metrics: sensitivity (SN), the ability to correctly identify true positive cases; specificity (SP), the ability to correctly identify true negative cases; positive predictive value (PPV), the probability that individuals identified as positive truly have the disorder; negative predictive value (NPV), the probability that individuals identified as negative truly do not have the disorder; likelihood ratio positive (LR+), the ratio indicating how much more likely a positive result is in a person with the disorder than in one without it; and likelihood ratio negative (LR-), the ratio indicating how much less likely a negative result is in a person with the disorder than in one without it.

Sensitivity and specificity with values above 90% are considered excellent, 70-90% fair, and below 70% poor (Levis et al., 2017). However, their clinical utility depends on disease prevalence, which directly influences PPV and NPV; high prevalence increases PPV while decreasing NPV, and vice versa (Bossuyt et al., 2015). Likelihood ratios further refine interpretation, with LR+ values above 10 and LR- values below 0.1 providing strong diagnostic evidence, while values between 2-5 and 0.2-0.5 offer weaker discrimination (Bossuyt et al. 2015). These benchmarks help clinicians and researchers determine whether a diagnostic tool, such as SENA, is suitable for practical application based on the balance of its accuracy measures.

To select the most balanced cut-off point included in Tables 4 and 5, we followed an approach based on the selection of cut-off points that

ensured both high diagnostic performance (SN and SP > 0.70) and external validation while avoiding post-hoc adjustments that could introduce bias (Reitsma et al., 2022). Therefore, cut-off points with acceptable sensitivity (>0.70) and high specificity (>0.70) were used, or, failing this, the best Youden's Index score, which balances SN and SP results.

Subsequently, operating characteristics and McNemar's χ^2 test were calculated for different cut-off points, which can help adjust cut-off points to improve the agreement between SENA and the clinical reference standard prevalence or enhancing diagnostic accuracy and ensuring the tool's effectiveness in clinical settings.

However, according to Levis et al. (2017) or Reitsma et al. (2022), the selective reporting of certain cut-off points may lead to an overestimation of a test's discriminative ability, highlighting the need to report results for all possible values rather than only the optimal ones for each individual study. Consequently, all cut-off scores for primary and secondary school students have been included in the supplementary material (see Supplementary Material).

Sampling weights were calculated and used in the analyses described above. First, to restore the distribution of the eligible sample, post-stratification weights were calculated according to gender, academic year and centre. Then, to adjust for possible imbalances in the clinical reappraisal sample, inverse probability weighting was calculated based on gender, academic year, centre and positive cases according to the SENA threshold. All statistical analyses were conducted using SAS v9.4, R v4.2.2 and SPSS v29.0.

3. Results

3.1. Participants characteristics

As shown in Figure 1, the clinical reappraisal process began with an initial baseline survey of 5,652 participants, followed by the selection of a clinical reappraisal sub-sample. The figure 1 outlines the steps in participant inclusion and exclusion, resulting in a final sample of 266 primary school students and 260 secondary school students for clinical evaluation. Of the 266 primary school children, 97 exceeded the cut-off point for the SENA, and of the 260 secondary school adolescents, 110 exceeded the cut-off point. After the KSADS-COM interview administration, of the 266 primary school children, 97 received a diagnosis, while 110 out of the 260 secondary school adolescents received a positive diagnosis.

As can be seen in Table 1 focused on primary school students, the weighted baseline survey results indicated that the distribution of sociodemographic characteristics between the overall sample and the clinical reappraisal sub-sample is nearly identical. Males represent 51.8% and 51.5%, respectively, with most participants falling between 8 to 12 years old., and Spanish (69.6% in the overall sample and 74.1% in the clinical sub-sample) Regarding emotional disorders, approximately 30% of children in both groups had an emotional disorder, with no notable differences in the prevalence between the samples.

The baseline survey characteristics for Secondary Education students in Table 2 showed similar weighted results, finding no differences between samples, with males comprising 54.0% and 53.3%, and Spanish nationality (88.2%-83%) Grade levels (1st to 4th) show no significant differences between the two groups. In terms of emotional disorders, 38.9% of the overall sample and 41.9% of the clinical reappraisal sub-sample have an emotional disorder, with no significant difference in prevalence. Additionally, responses to the suicide item on the SENA show similar patterns across both groups, with around 30-33% answering "Yes".

3.2. Prevalence estimates of the KSADS-COMP based on the SENA scores

The Table 3 indicates that SENA both for children and adolescents identified emotional disorders in 31.6% of children and 35.0% of
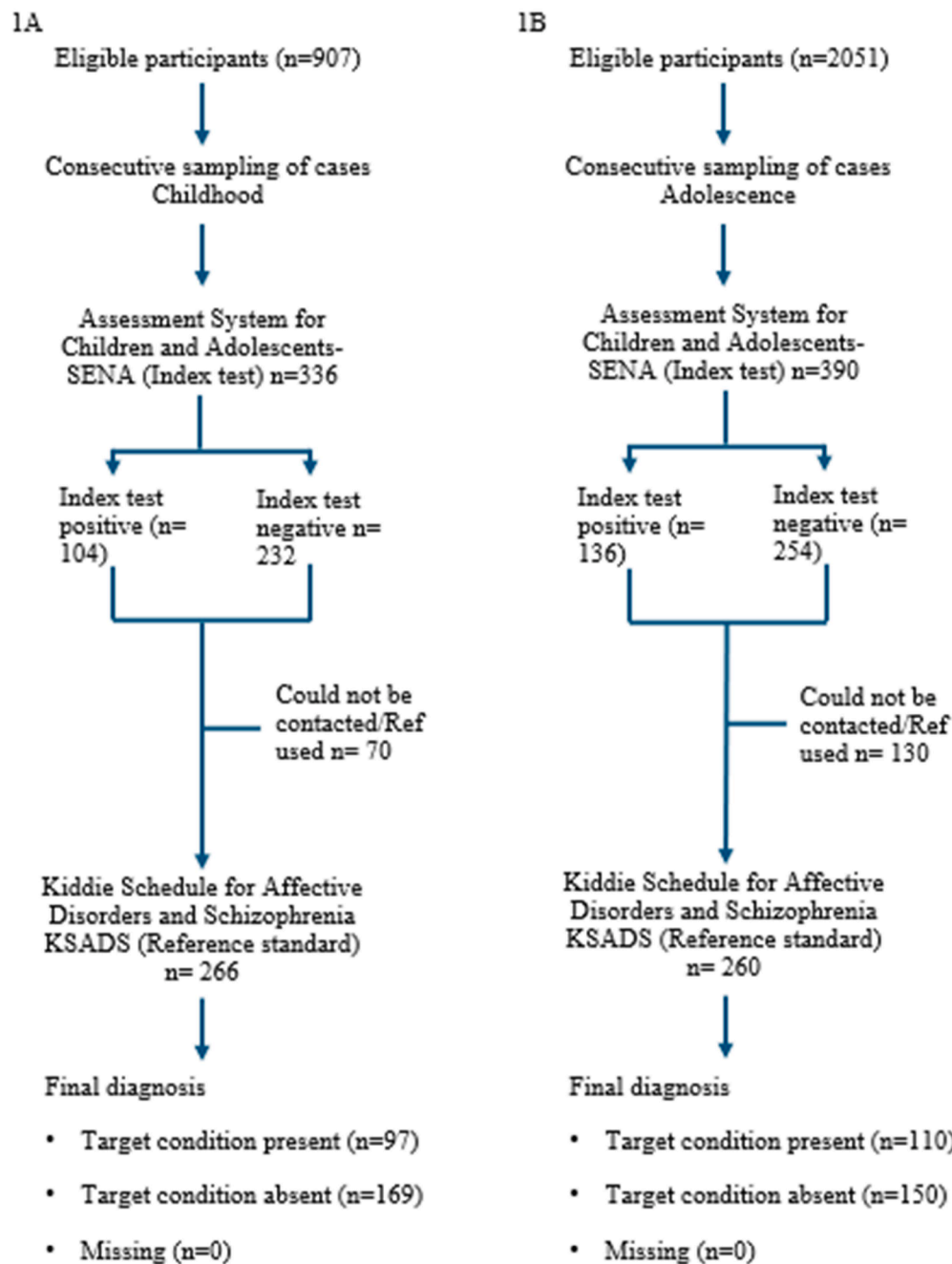
Figure 1. Flow diagram of participants through clinical reappraisal study. Figure 1a shows the clinical reappraisal conducted in children. Figure 1b. shows the clinical reappraisal conducted in adolescents.

adolescents, but KSADS-COMP confirmed a lower but similar prevalence of 24.6% ($p = .028$) for children and a higher prevalence of 51.1% in adolescents ($p < .001$). This suggests that the SENA slightly over-estimates prevalence in younger participants but may underestimate emotional disorders in adolescents. As regards MDE, GAD, SAD, PTSD, and OCD, SENA tool consistently overestimates the prevalence of emotional disorders compared to the clinical assessment conducted using KSADS-COMP, finding statistically significant differences in all the disorders evaluated. Finally, as regards suicidal thoughts and behaviours, SENA identified 20.7% of children and 33.4% of adolescents with suicidal thoughts/behaviours, while KSADS-COMP found 9.8% and 12.4% ($p < .001$), respectively, indicating a substantial overestimation by SENA.

3.3. Operating Characteristics of SENA scores

The results in Table 4 summarize the operating characteristics of the SENA scores for estimating the prevalence of emotional and mental disorders based on the KSADS-COMP reference standard, with separate data for children and adolescents.

For emotional disorders, the SENA tool demonstrates moderate sensitivity (SN) and specificity (SP) in both children (SN=78.5%, SP=65.9%) and adolescents (SN=80.4%, SP=66.1%), with area under the curve (AUC) values of .74 and .73, respectively. These results suggest that the tool has moderate accuracy for identifying emotional disorders. When excluding sleep problems and suicide, the sensitivity slightly in-creases for children (SN=87.0%, AUC=.76) but decreases slightly for adolescents (SN=77.8%, AUC=.72). Suicide could be excluded from the

Table 1
Characteristics of the EMO-CHILD baseline survey on the overall sample and clinical reappraisal sub-sample (unweighted observations and weighted percentages). Primary school students subsample.

	Levels	Overall sample n=336[c]		Clinical reappraisal sub-sample n=266[c]		
		n	%	N	%	P-value[*]
Sex	Male	176	51.8	129	51.5	.91
	Female	160	48.2	137	48.5	
Age	8 or 9	89	28.6	68	25.9	.36
	10	95	28.3	87	32.3	
	11	120	34.3	85	31.7	
	12	32	8.8	26	10.0	
Nationality	Spain	233	69.6	195	74.1	.11
	Non-Spanish	103	30.4	71	25.9	
Grade	4th	109	35.2	84	33.0	.19
	5th	98	29.7	94	34.7	
	6th	129	35.2	88	32.3	
Any emotional disorders[a]	No	232	69.0	169	68.4	.83
	Yes	104	31.0	97	31.6	
Suicide item SENA[b]	No	261	77.8	206	79.3	.56
	Yes	75	22.2	60	20.7	
Any emotional disorder (EMO)	mean (SD)	50.6 (10.46)		50.7 (10.51)		
Depression (DEP)	mean (SD)	50.5 (10.57)		50.6 (10.6)		
Anxiety (ANX)	mean (SD)	50.9 (10.51)		51.2 (10.6)		
Social anxiety (ASC)	mean (SD)	50.3 (10.12)		50.6 (10.6)		
Somatic symptom (SOM)	mean (SD)	50.7 (10.79)		50.4 (10.62)		
Post-traumatic Symptomatology (PTS)	mean (SD)	50.1 (9.87)		50.2 (9.86)		

[a] .- Cut-point=> PT 56 in EMO score of SENA (standard scores from EmoChild Project)

[b] .- Death wish extracted from a question in the SENA questionnaire; Yes is any response different to Never (Fernández-Pinto et al., 2015a)

[c] .- The values presented in the table are based on n unweighted and % weighted observations.

[*] P-value statistically significant .05.

Table 2
Characteristics of the EMO-CHILD baseline survey on the overall sample and clinical reappraisal sub-sample (unweighted observations and weighted percentages). Secondary school students subsample.

Variable	Levels	Overall sample n=390[c]		Clinical reappraisal sub-sample n=260[c]		
		n	%	n	%	p-value [*]
Sex	Male	181	54.0	107	53.3	.81
	Female	209	46.0	153	46.7	
Age	12	82	19.6	55	20.4	.96
	13	93	24.2	71	25.2	
	14	93	25.2	60	25.6	
	15	81	20.9	47	19.1	
	16	41	10.1	27	9.8	
Nationality	Spain	346	88.2	218	83.0	.01
	Non-Spanish	44	11.8	42	17.0	
Grade	1st	117	29.6	82	32.1	.59
	2nd	106	26.9	82	27.5	
	3rd	78	21.2	42	19.5	
	4th	89	22.3	54	21.0	
Any emotional disorder[a]	No	254	66.5	150	65.0	
	Yes	136	33.5	110	35.0	.62
Suicide item SENA[b]	No	274	69.1	168	66.6	
	Yes	116	30.9	92	33.4	.38
Any emotional disorder (EMO)	mean (SD)	52.3 (9.85)		53.0 (9.97)		
Depression (DEP)	mean (SD)	52.1 (10.23)		53.1 (10.36)		
Generalized Anxiety (ANS)	mean (SD)	52.2 (9.63)		52.8 (9.54)		
Social anxiety (ASC)	mean (SD)	52.1 (10.62)		52.7 (10.82)		
Somatic symptom (SOM)	mean (SD)	51.3 (9.73)		52.0 (10.20)		
Post-Traumatic Symptoms (PTS)	mean (SD)	51.9 (9.65)		52.6 (10.24)		
Obsession-Compulsion (OBS)	mean (SD)	51.5 (9.81)		51.9 (9.43)		

[a] .- Cut-point=>PT 56 in EMO score of SENA (standard scores from EmoChild Project)

[b] .- Death wish extracted from a question in the SENA questionnaire; Yes is any response different to Never (Fernández-Pinto et al., 2015a)

[c] .- The values presented in the table are based on n unweighted and % weighted observations.

[*] P-value statistically significant .05.

primary SENA score analysis to avoid inflating the emotional index and because suicide-related interventions often fall outside the scope of school-based screenings. Sleep items should be excluded due to low variance and limited discriminant validity in pilot analyses.

For specific disorders, sensitivities ranged from 79 to 100, but a lower SN was found for OCD (69.9); and specificities were high (64 to 88.4). PPVs ranged from 15.3 to 40.8 while NPVs ranged from 91.8 to 100. SENA achieves moderate to substantial accuracy for all disorders (AUCs=.72-.93), except for OCD (AUC=.67).

3.4. Improving diagnostic capacity through cut-off point changes

Table 5 illustrates the effects of adjusting diagnostic cut-off points to improve comparability between SENA prevalence estimates and the KSADS-COMP reference standard. For most diagnoses, increasing cut-off points reduced the overestimation of prevalence, aligning estimates closer to the reference standard, particularly under the "Optimal for Prevalence" approach. However, these adjustments often came at the expense of lower sensitivity (SN) and area under the curve (AUC), as seen in emotional disorders without sleep problems and suicide (SN=57.2%, AUC=.73) and MDE (SN=53.9%, AUC=.72). Positive predictive values (PPVs) generally improved with optimal cut-off points, as observed in SAD and GAD, with increases ranging from 30.6% to 45.9%, while negative predictive values (NPVs) remained high, often exceeding 86%. Notably, the "High Sensitivity" cut-off points

maintained strong sensitivity but overestimated prevalence significantly, especially for PTSD and suicidal thoughts and behaviours, highlighting the trade-offs in optimizing diagnostic accuracy versus sensitivity.

In Table 5 also is presented the operating characteristics of SENA scores for secondary school students when diagnostic cut-off points are adjusted to maximize sensitivity (High Sn) or align prevalence estimates with KSADS-COMP (Optimal for Prevalence). Adjusting the cut-offs significantly impacts the balance between sensitivity and specificity. For "High Sensitivity" approaches, SENA generally overestimates prevalence compared to KSADS-COM, as observed in emotional disorders (57.6% vs. 51.1%, $p = .041$) and PTSD (38.2% vs. 12.8%, $p < .001$), with AUC values ranging between .67 and .76. The "Optimal for Prevalence" cut-offs improve alignment between SENA and KSADS-COMP prevalence estimates, such as for emotional disorders (53.7% vs. 51.1%, p = 0.424) and OCD (27.9% vs. 22.6%, $p = .103$), but typically reduce sensitivity (e.g., 71.1% for emotional disorders without sleep problems) while maintaining moderate-to-high specificity (e.g., 89.8% for GAD). Positive predictive values (PPVs) show modest improvements in most cases under optimal cut-off points, while negative predictive values (NPVs) remain consistently high, often exceeding 86%. These results highlight the trade-offs between enhancing diagnostic capacity through sensitivity maximization and achieving prevalence alignment.

Table 3

Current prevalence estimates of common emotional mental disorders and suicidal thoughts and behaviours in the clinical reappraisal sample, according to the SENA scores and the KSADS-COMP (unweighted observations and weighted observations).

		SENA Emotional scores			KSADS-COMP			McNemar	
		n	%	CI 90%	n	%	CI 90%	x2	p-value*
Emotional disorders[a]	Children	97	31.6	26.0-37.2	70	24.6	19.5-29.8	4.83	.028
	Adolescents	110	35.0	29.2-40.8	140	51.1	45-57.2	20.14	<.001
MDE	Children	76	25.6	20.3-30.8	30	9.8	6.3-13.4	31.19	<.001
	Adolescents	113	37.7	31.8-43.6	61	20.2	15.3-25.1	30.83	<.001
GAD	Children	103	35.6	29.9-41.4	7	2.2	0.4-3.9	88.92	<.001
	Adolescents	120	39.3	33.4-45.3	28	8.3	4.9-11.7	72.7	<.001
SAD	Children	87	29.3	23.8-34.8	11	3.9	1.6-6.2	65.91	<.001
	Adolescents	109	40.4	34.4-46.3	21	5.8	3-8.6	88.05	<.001
PTSD	Children	96	32.5	26.8-38.1	27	8.9	5.5-12.3	51.87	<.001
	Adolescents	96	32.5	26.8-38.1	44	12.8	8.8-16.9	37.49	<.001
OCD	Children	-	-	-	-	-	-	-	-
	Adolescents	96	33.7	27.9-39.4	60	22.6	17.5-27.7	10.48	.001
Suicidal thoughts	Children	60	20.7	15.8-25.6	25	9.8	6.2-13.3	19.34	<.001
and behaviors	Adolescents	92	33.4	27.7-39.2	36	12.4	8.4-16.4	46,32	<.001

[a] .- Emotional disorders: MDE: Major Depressive Episode; GAD: Generalized Anxiety Disorder; SAD: Social Anxiety Disorder; PTSD: Post-Traumatic Stress Disorder; OCD: Obsessive Compulsive Disorder [KSADS-COMP]
* P-value statistically significant .05. CI 90%: Confidence interval
Note. Analyses were performed separately for children and adolescents. Prevalence estimates for emotional disorders were calculated based on SENA results using the 75th percentile (T-scores ≥56).

Table 4

SENA scores operating characteristics for estimating reference standard (KSADS-COM) prevalence (n=526; 266 for each primary and 260 secondary education level) (weighted values).

			Positive operating characteristics					Negative operating characteristics					AUC
		Cut-point	SN	SE (SN)	PPV	SE (PPV)	LR+	SP	SE(SP)	NPV	SE(NPV)	LR-	
Emotional disorders[a]	Children	(51)	78.5	5.1	40.8	5.1	2.3	65.9	5.1	91.8	3.1	0.3	.74
	Adolescents	(51)	80.4	5.1	71.4	5.1	2.4	66.1	5.1	76.5	3.1	0.3	.73
Emotional disorders[b]	Children	(53)	76.5	5.1	35.7	5.1	2.5	68.8	5.1	91.8	3.1	0.3	.73
	Adolescents	(52)	76.1	5.1	66.3	5.1	2.2	64.9	5.1	76.5	3.1	0.4	.72
Emotional disorders[c]	Children	(53)	87.0	5.1	30.6	5.1	2.8	68.4	5.1	96.9	0	0.2	.76
	Adolescents	(52)	77.8	5.1	61.2	5.1	2.2	64.6	5.1	81.6	3.1	0.3	.72
MDE	Children	(55)	79.7	10.2	25.5	5.1	3.2	75.3	5.1	96.9	0	0.3	.77
	Adolescents	(54)	80.8	5.1	40.8	5.1	2.8	71.3	5.1	91.8	3.1	0.3	.75
GAD	Children	(64)	100	0	15.3	5.1	8.6	88.4	0	102	0	0	.93
	Adolescents	(56)	79.4	10.2	15.3	5.1	2.2	64.3	5.1	96.9	0	0.3	.74
SAD	Children	(59)	91.4	10.2	15.3	5.1	4.6	80.3	0	102	0	0.1	.84
	Adolescents	(58)	88.1	10.2	15.3	5.1	3.2	72.5	5.1	96.9	0	0.2	.76
PTSD	Children	(53)	81.5	10.2	20.4	5.1	2.5	67.1	5.1	96.9	0	0.3	.75
	Adolescents	(55)	81.2	5.1	25.5	5.1	2.6	68.2	5.1	96.9	0	0.3	.73
OCD	Children	-	-	-	-	-	-	-	-	-	-	-	-
	Adolescents	(54)	69.9	5.1	35.7	5.1	2	64.9	5.1	86.7	3.1	0.5	.67
Suicide[d]	Children	(1)	72.0	10.2	35.7	5.1	4.8	84.9	0	96.99	0	0.3	.82
	Adolescents	(1)	84.5	5.1	30.6	5.1	3.2	73.8	5.1	96.9	0	0.2	.76

[a] - Emotional disorders: MDE: Major Depressive Episode; GAD: Generalized Anxiety Disorder; SAD: Social Anxiety Disorder; PTSD:Post-Traumatic Stress Disorder; OCD: Obsessive Compulsive Disorder [KSADS-COM]
[b] .-Emotional disorders without sleep problems
[c] .- Emotional disorders without sleep problems and suicide
[d] .- Suicide: Including Suicidal thoughts and behaviours

4. Discussion

This study evaluated the diagnostic concordance of the SENA with the DSM-5 diagnostic interview KSADS-COMP for detecting and diagnosing emotional disorders in a sample drawn from a broad representative cohort of Spanish children and adolescents. Overall, the results indicate that the SENA demonstrates adequate concordance, with area under the curve (AUC) values ranging from .67 to .93 and sensitivity values between 70% and 100% for identifying disorders such as major depressive episode (MDE), generalized anxiety disorder (GAD), and post-traumatic stress disorder (PTSD), among others, with specificity values between 64.6% and 88.4%. These results suggest that the SENA is an effective screening tool for detecting emotional disorders in children and adolescents, with an adequate discriminative ability, as well as a high sensitivity that facilitates the identification of a high proportion of

positive cases. It also has a moderate specificity indicating the need for follow-up clinical assessments to confirm diagnoses and minimise false positives.

In terms of diagnostic accuracy, concerning the presence of any emotional disorder, the SENA demonstrates adequate sensitivity, fair to poor specificity and fair discrimination in both children (SN=78.5%, SP=65.9%, AUC=.74) and adolescents (SN=80.4%, SP=66.1%, AUC=.73), reflecting moderate accuracy in identifying emotional disorders. Excluding sleep problems and suicide, sensitivity slightly improves for children (SN=87.0%, AUC=.76) but decreases for adolescents (SN=77.8%, AUC=.72). The inclusion/exclusion of sleep problems and suicide diagnoses is controversial, but it could be justified due to their close relationship with disorders such as anxiety and depression, where they often appear as relevant or comorbid symptoms. However, it is important to recognize that, strictly speaking, they are distinct

Table 5

SENA scores operating characteristics for estimating reference standard (KSADS-COM) current prevalence when diagnostic cut-off point are changed to maximize sensitivity or have an optimal prevalence (n=526; 266 for each primary and 260 secondary education level) (weighted valued). Primary and Secondary school students.

		Cut-point	SENA prevalence estimate		KSADS-COMP prevalence		Positive operating characteristics					Negative operating characteristics					McNemar		AUC
			%	SE (%)	%	SE (%)	SN	SE (SN)	PPV	SE (PPV)	LR+	SP	SE (SP)	NPV	SE (NPV)	LR-	x2	p-value*	
Any Emotional disorders[a]	High Sn	(51)	45.1	3.1	24.6	2.6	78.5	5.1	40.8	5.1	2.3	65.9	5.1	91.8	3.1	0.3	35.77	<.001	.74
	Optimal for prevalence	(57)	28.1	2.8	24.6	2.6	56.1	5.1	51	5.1	3	81.1	5.1	86.7	3.1	0.5	1.24	.266	.71
Any Emotional disorders[b]	High Sn	(51)	57.6	3.1	51.1	3.1	80.4	5.1	71.4	5.1	2.4	66.1	5.1	76.5	3.1	0.3	4.19	.041	.73
	Optimal for prevalence	(52)	53.7	3.1	51.1	3.1	76.3	5.1	71.4	5.1	2.5	69.9	5.1	71.4	3.1	0.3	0.64	.424	.74

[a] .- Primary: Any emotional disorder: MDE: Major Depressive Episode; GAD: Generalized Anxiety Disorder; SAD: Social Anxiety Disorder; PTSD: Post-Traumatic Stress Disorder.

[b] .- Secondary: Any emotional disorder: MDE: Major Depressive Episode; GAD: Generalized Anxiety Disorder; SAD: Social Anxiety Disorder; PTSD: Post-Traumatic Stress Disorder; OCD: Obsessive Compulsive Disorder [KSADS-COMP]

SE: Standard error.

* P-value statistically significant .05.

phenomena. Sleep problems, for instance, may be present as a primary indicator in sleep disorders or as a secondary symptom related to emotional distress. Similarly, suicidal behaviours reflect a severe manifestation of hopelessness, which can be characteristic of depression but does not necessarily define the entire clinical picture. This approach allows for a nuanced analysis, evaluating both their contribution within the emotional spectrum and their independent value as specific phenomena.

Regarding specific disorders, SENA demonstrates moderate to high sensitivity across most disorders, particularly in children, with notable performance for detecting generalized anxiety disorder (GAD) and social anxiety disorder (SAD). However, specificity and positive predictive values vary, with lower values observed for GAD and suicidal thoughts and behaviours. The tool performs moderately well for post-traumatic stress disorder (PTSD) but shows limitations in detecting obsessive-compulsive disorder (OCD) in adolescents. These findings highlight the utility of SENA as a screening tool while emphasizing the need for complementary assessments to address variability in predictive accuracy across disorders.

Overall findings suggest that the SENA is a useful screening tool, performing well in identifying individuals at risk for emotional and mental disorders, particularly for conditions like GAD and SAD in children and, SAD and PTSD in adolescents. These results support its use in educational and clinical settings as an initial measure to identify cases requiring further evaluation. However, its precision varies across disorders and age groups, underscoring the need to complement it with clinical diagnostic tools to improve specificity and reduce false positives.

5. Comparison with the diagnostic accuracy metrics of other dimensional measures

When compared with dimensional measures such as the Strengths and Difficulties Questionnaire (SDQ), BASC or Achenbach System of Empirically Based Assessment (ASEBA), the SENA performs similarly in terms of sensitivity but shows lower specificity and PPV values. Studies on the SDQ (e.g., Reilly et al., 2014) and ASEBA (Achenbach et al., 2001) report moderate to high diagnostic accuracy, with sensitivity and specificity values around 70%-80% for emotional subscales, depending on the population and context. Therefore, our findings align with previous research evaluating mental health screening tools for children. A meta-analysis by Stockings et al. (2015) evaluating screening tools for depression reported a pooled sensitivity of 80% and specificity of 78%, with AUC values like those observed for SENA in this study (e.g., AUC = .74 for emotional disorders). Similarly, studies such as Armitage et al. (2023) evaluating the SDQ-emotional subscale found moderate diagnostic accuracy for MDE (AUC = .67–.85) and high accuracy for GAD (AUC = .80–.93), comparable to SENA's performance for these disorders (AUC = .75-.93 in children).

However, in our study, the SENA tends to overestimate the prevalence of emotional disorders, a trend that could be attributed to differences in populations across studies, particularly the inclusion of a non-clinical sample in contrast to previous research conducted in clinical settings. Thus, previous studies, such as those conducted with the SDQ and CBCL, report more accurate concordance between screening tools and clinical diagnoses in controlled clinical settings (Armitage et al., 2023). In our study, this discrepancy could be due to SENA's use in a non-clinical setting, such as the educational context, underscoring the need to validate these tools in diverse population settings, such as clinical populations. SENA's tendency to overestimate prevalence suggests the need for refining diagnostic thresholds and case definitions.

Although tools like ASEBA and BASC often achieve higher specificity due to their structured item design, SENA's strong sensitivity and cultural adaptation for Spanish-speaking populations make it particularly valuable in contexts where early identification and broad screening are prioritized. However, as with other dimensional measures, the findings underscore the importance of integrating clinical interviews or

additional diagnostic tools to confirm cases and mitigate overestimation tendencies.

The results presented highlight the impact of adjusting cut-off points on SENA's diagnostic capacity, showing that these changes can improve the alignment of prevalence estimates with the KSADS-COMP reference standard, albeit often at the cost of sensitivity (SN) and overall accuracy (AUC). For example, applying the "Optimal for Prevalence" approach significantly aligned SENA's prevalence estimates with those of KSADS-COMP for most diagnoses. For instance, prevalence rates for emotional disorders excluding sleep problems and suicide adjusted to 53.7% (SENA) versus 51.1% (KSADS-COM, $p = .424$), while for OCD, SENA's estimates (27.9%) were similar to KSADS-COM's (22.6%, $p = .103$). However, these improvements in specificity and PPVs came with a notable reduction in sensitivity, as observed for emotional disorders (SN = 57.2%, AUC = .73) and MDE (SN = 53.9%, AUC = .72).

The adjustments to prioritize sensitivity ("High Sensitivity") allowed for higher detection rates but at the cost of significant overestimation of prevalence. This was observed in disorders such as PTSD, where SENA's prevalence rates (38.2%) were significantly higher than KSADS-COM's (12.8%, $p < .001$), and in suicidal thoughts and behaviours. Although sensitivity remained high under this approach (SN $\geq$ 70% for most cases), PPVs were lower compared to the "Optimal for Prevalence" approach. This indicates that prioritizing sensitivity may be useful in clinical or educational contexts where minimizing false negatives is critical but could result in higher rates of false positives.

Regardless of the approach, SENA's negative predictive values (NPVs) remained consistently high, exceeding 86% for most diagnoses. This reinforces its utility as a screening tool for confidently ruling out negative cases. On the other hand, PPVs showed moderate improvements under optimal cut-off points, with increases of up to 45.9% for diagnoses such as GAD and SAD, suggesting that adjustments can reduce false positives and enhance overall diagnostic accuracy.

Adjusting SENA's cut-off points to maximize sensitivity or align prevalence estimates with KSADS-COMP outcomes significantly impacted its diagnostic performance, aligning with findings from other studies such as Ballester et al. (2019). For example, optimizing prevalence estimates for emotional disorders reduced sensitivity (e.g., SN = 56.1% for primary school students) but improved specificity and alignment with KSADS-COMP estimates. These adjustments highlight the inherent trade-offs between sensitivity and specificity, as well as the challenges of adapting screening thresholds to specific populations and contexts.

These findings underscore the importance of tailoring cut-off points to the specific objectives of the evaluation. In research or population monitoring contexts, "Optimal for Prevalence" thresholds may be preferable to ensure estimates aligned with clinical standards. However, in settings where early detection is prioritized, such as schools or mental health centres, thresholds designed to maximize sensitivity may be more appropriate, although they should be complemented by additional clinical evaluations to confirm positive cases.

Although this study did not primarily aimed to analyse the prevalence of emotional disorders, it identified significant discrepancies between SENA's and KSADS-COM's prevalence estimates. The results show that SENA consistently overestimates prevalence rates compared to clinical evaluations conducted using KSADS-COM, particularly among children. While SENA identified emotional disorders in 31.6% of children and 35.0% of adolescents, KSADS-COMP confirmed significantly lower rates of 24.6% in children ($p < .001$) and higher rates in adolescents (51.1%, $p = .090$). This disparity highlights SENA's tendency to generate slightly inflated prevalence rates, particularly among younger populations. Interestingly, the gap between SENA and KSADS-COMP prevalence estimates was greater for adolescents, with significant differences observed. This suggests that SENA may perform better in younger groups for detecting general emotional disorders, possibly reflecting developmental differences in symptom presentation or alignment of its scales with adolescent-specific symptoms. This pattern

underscores SENA's sensitivity as a screening tool but also its limitations in specificity, particularly in adolescent populations. A tentative explanation for these differences is that, in younger children, parent-report forms such as the SENA often capture a broader spectrum of emotional symptoms, including transient or subclinical manifestations. These questionnaires are designed to integrate observations over time, which increases their sensitivity to mood and behavioral fluctuations that may not meet clinical diagnostic thresholds (Achenbach et al., 2017). In contrast, structured clinical interviews like the KSADS-COMP apply stricter diagnostic criteria based on DSM definitions, thereby reducing the likelihood of identifying subthreshold symptoms (Townsend et al., 2020). Among adolescents, increased self-awareness and greater capacity to articulate internal states may enhance the detection of emotional disorders during interviews. Cognitive and emotional maturation in this developmental stage supports adolescents' ability to identify and report internalizing symptoms—such as anxiety or depression—that are often invisible to parents (Lee et al., 2023).

For specific conditions, SENA consistently overestimated prevalence compared to KSADS-COM. This aligns with previous studies showing that screening tools tend to report higher prevalence rates than structured diagnostic interviews. For instance, Polanczyk et al. (2015) and Sherr et al. (2021) reported global prevalence rates for anxiety and depression in children and adolescents of 6.5%-7.9% and 1.7%-2.6%, respectively, significantly lower than SENA's rates, particularly for GAD (34.9% in adolescents) and depression (37.9% in adolescents). Studies such as those conducted by Spain's National Health System (2020) and Kusters et al. (2023) also documented lower rates (10% for anxiety and 4% for mood disorders), aligning more closely with KSADS-COMP findings. SENA's overestimation tendencies emphasize the importance of interpreting SENA's estimates cautiously. More clearly, the findings indicate that while SENA is a valuable tool for identifying at-risk individuals, it should not be used as a standalone diagnostic measure. Significant discrepancies in prevalence rates for specific conditions suggest that SENA's dimensional design captures subclinical symptoms, which may explain its high sensitivity but low specificity.

This study has several limitations that must be considered. First, the use of online interviews and the lack of inter-rater reliability evaluation may introduce biases related to self-reporting and participant engagement. Additionally, this methodology may not fully capture the complexity of emotional disorders, particularly in cases of comorbidity, where overlapping symptoms could lead to misclassification or underestimation of specific conditions. Furthermore, in more severe cases, the fact that participants shared the classroom with their peers—especially in smaller classrooms—sometimes could have led to feelings of inhibition, which may have affected their willingness to disclose personal experiences. This limitation highlights the importance of considering contextual and environmental factors in mental health assessments, as these can influence the accuracy of self-reported symptoms. Furthermore, social desirability bias may have influenced participants' responses, potentially affecting the accuracy of self-reported emotional experiences.

Second, the time gaps between the initial SENA assessment and the clinical interview may have introduced biases related to symptom progression. Although we set a time frame of three months between the initial evaluation and the clinical interview, this does not invalidate the results. A shorter interval could have increased the possibility of conscious or unconscious recall of responses, potentially biasing the results. On the other hand, the SENA is designed to measure traits or aspects that, in the absence of significant interventions or events, remain relatively stable over time, making it reasonable to assume that the characteristics assessed do not undergo drastic changes within this period.

Additionally, the small sample sizes for certain specific disorders, such as OCD, limit the reliability of subgroup analyses. Future studies should aim to increase sample sizes for these conditions to enhance statistical power and generalizability.

Another limitation was the choice of an instrument that has not been fully validated in Spain (KSADS-COMP) as the gold standard. While the KSADS-COMP can be administered in Spanish, it lacks a complete validation study within the Spanish population, which could affect cross-study comparisons. However, the Spanish version was reviewed and adjusted to ensure proper usage while maintaining the structure of the interview's diagnostic criteria. Additionally, although the sample was drawn from multiple regions across Spain, generalizing the findings to a broader population requires further exploration to ensure applicability to diverse demographic and clinical groups. Lastly, longitudinal research is needed to assess SENA's predictive validity over time and its effectiveness in identifying long-term outcomes.

Future research should focus on expanding the samples and disorders assessed to refine SENA's diagnostic thresholds, further clarifying its relative strengths and weaknesses. Additional studies should also examine the predictive validity of SENA in identifying individuals at risk who later develop clinically significant disorders. Efforts to integrate SENA into multi-level intervention systems could leverage its high sensitivity to support early identification while ensuring that clinical follow-up guarantees accurate diagnoses.

Another limitation concerns the length and comprehensiveness of the SENA, which, while advantageous for detailed assessment, may reduce efficiency when used solely as a screening tool. In particular, some subscales with lower specificity may contribute to increased false positives or capture subclinical symptoms. Future versions of the SENA could consider item reduction strategies aimed at improving specificity without compromising sensitivity. Moreover, its clinical application could be strengthened by combining the tool with brief follow-up protocols administered by trained professionals to confirm risk and guide appropriate intervention.

Adjusting SENA's cut-off points highlights the trade-offs inherent between sensitivity, specificity, and prevalence alignment. While the "High Sensitivity" approach identifies more potential cases, the "Optimal for Prevalence" strategy enhances diagnostic accuracy by reducing overestimation. These results reinforce SENA's utility as a flexible tool that can adapt to various contexts and objectives but stress the need to combine it with clinical evaluations to ensure precise and effective detection of emotional and mental disorders. For example, in clinical applications, prioritizing higher sensitivity could help reduce false negatives, while for population-based research, optimizing prevalence estimates is crucial for generating more representative data.

In conclusion, given the high sensitivity observed across several subscales, the SENA appears to be a valuable tool for initial screening in both educational and clinical settings, particularly for identifying individuals at risk of developing emotional disorders. However, to ensure diagnostic accuracy—especially in younger populations and in conditions with serious implications such as suicidal behaviors—it is important to complement screening results with additional diagnostic interviews. While the KSADS-COMP provides a rigorous reference standard for research contexts, in routine clinical and educational practice, structured interviews conducted by qualified mental health professionals are generally considered sufficient to confirm positive cases.

This study represents an important step in addressing the growing burden of emotional disorders among Spanish children and adolescents by validating a widely used, culturally adapted assessment tool. The results will contribute to the development of more precise and effective mental health screening practices, ultimately improving outcomes for this vulnerable population.

The EmoChild study highlights the critical role of SENA as a culturally sensitive screening instrument for emotional disorders and suicidal behaviors in children and adolescents. Developed specifically for Spanish-speaking populations, SENA's strong psychometric properties and comprehensive design make it an invaluable resource for addressing mental health challenges in Spain. While it demonstrates moderate diagnostic accuracy and high sensitivity, its overestimation of prevalence rates and moderate PPVs emphasize the importance of complementary diagnostic evaluations. The findings underline the necessity for culturally adapted, evidence-based mental health tools to address the growing burden of emotional disorders among youth in Spain while also identifying opportunities to enhance SENA's precision and effectiveness in clinical and educational settings.

Data Availability Statement

Data associated with this study are available at University Miguel Hernandez, Prof. J.A. Piqueras. They can be accessed for data analysis upon request by mail to the first author.

CRediT authorship contribution statement

Jose A. Piqueras: Writing – original draft, Supervision, Investigation, Data curation, Conceptualization. **Raul Castaño:** Writing – original draft, Investigation. **Laura Ballester:** Writing – review & editing, Methodology, Formal analysis, Conceptualization. **Gemma Vilagut:** Writing – review & editing, Methodology, Formal analysis, Conceptualization. **Jordi Alonso:** Writing – review & editing, Methodology, Formal analysis, Conceptualization. **Alexandra Morales:** Writing – review & editing, Data curation, Conceptualization. **Jose P. Espada:** Writing – review & editing, Project administration, Conceptualization. **Mireia Orgiles:** Writing – review & editing, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

All authors disclose that they do not have got any financial and personal relationships with other people or organizations that could inappropriately influence (bias) their work.

Acknowledgements

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.psychres.2026.116994.

References

Achenbach, T.M., 2001. What are norms and why do we need valid ones? Clinical Psychology: Science and Practice 8 (4), 446–450. https://doi.org/10.1093/clipsy.8.4.446.

Achenbach, T.M., Rescorla, L.A., 2001. Manual for the ASEBA school-age forms & profiles. University of Vermont, Research Center for Children, Youth & Families.

Achenbach, T.M., Ivanova, M.Y., Rescorla, L.A., 2017. Empirically based assessment and taxonomy of psychopathology for ages 1½-90+ years: Developmental, multi-informant, and multicultural findings. Comprehensive Psychiatry 79, 4–18. https://doi.org/10.1016/j.comppsych.2017.03.006.

Armitage, J.M., Tseliou, F., Riglin, L., Dennison, C., Eyre, O., Bevan Jones, R., et al., 2023. Validation of the Strengths and Difficulties Questionnaire (SDQ) emotional subscale in assessing depression and anxiety across development. PLoS ONE 18 (7), e0288882. https://doi.org/10.1371/journal.pone.0288882.

Ballester, L., Alayo, I., Vilagut, G., Almenara, J., Cebrià, A.I., Echeburúa, E., Gabilondo, A., Gili, M., Lagares, C., Piqueras, J.A., Roca, M., Soto-Sanz, V., Blasco, M.J., Castellví, P., Forero, C.G., Bruffaerts, R., Mortier, P., Auerbach, R.P., Nock, M.K., Sampson, N., UNIVERSAL study group, 2019. Accuracy of online survey assessment of mental disorders and suicidal thoughts and behaviors in Spanish university students. Results of the WHO World Mental Health- International College

Student initiative. PloS one 14 (9), e0221529. https://doi.org/10.1371/journal. pone.0221529.

Bossuyt, P.M., Reitsma, J.B., Bruns, D.E., Gatsonis, C.A., Glasziou, P.P., Irwig, L., Lijmer, J.G., Moher, D., Rennie, D., de Vet, H.C., Kressel, H.Y., Rifai, N., Golub, R.M., Altman, D.G., Hooft, L., Korevaar, D.A., Cohen, J.F., Group, STARD, 2015. STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies. Radiology 277 (3), 826–832. https://doi.org/10.1148/radiol.2015151516.

Casares, M.Á., Díez-Gómez, A., Pérez-Albéniz, A., Fonseca-Pedrero, E., 2024. Anxiety and depression in educational settings: Prevalence, assessment, and impact on psychological adjustment. Revista de Psicodidáctica (English ed.) 29 (1), 1–8. https://doi.org/10.1016/j.psicoe.2023.12.003.

Cui, Y., Li, F., Leckman, J.F., Guo, L., Ke, X., Liu, J., Zheng, Y., Li, Y., 2021. The prevalence of behavioral and emotional problems among Chinese school children and adolescents aged 6–16: A national survey. European Child & Adolescent Psychiatry 30, 233–241. https://doi.org/10.1007/s00787-020-01507-6.

Danielson, M.L., Bitsko, R.H., Holbrook, J.R., Charania, S.N., Claussen, A.H., McKeown, R.E., Cuffe, S.P., Owens, J.S., Evans, S.W., Kubicek, L., Flory, K., 2021. Community-Based Prevalence of Externalizing and Internalizing Disorders among School-Aged Children and Adolescents in Four Geographically Dispersed School Districts in the United States. Child Psychiatry and Human Development 52 (3), 500–514. https://doi.org/10.1007/s10578-020-01027-z.

De la Peña, F.R., Villavicencio, L.R., Palacio, J.D., Félix, F.J., Larraguibel, M., Viola, L., Ortiz, S., Rosetti, M., Abadi, A., Montiel, C., Mayer, P.A., Fernández, S., Jaimes, A., Feria, M., Sosa, L., Rodríguez, A., Zavaleta, P., Uribe, D., Galicia, F., Botero, D., Ulloa, R.E., 2018. Validity and reliability of the kiddie schedule for affective disorders and schizophrenia present and lifetime version DSM-5 (K-SADS-PL-5) Spanish version. BMC Psychiatry 18 (1), 193. https://doi.org/10.1186/s12888-018-1773-0.

Duncan, L., Wang, L., Boyle, M.H., 2024. Comparing the reliability and validity of youth-reported checklists and standardized interviews for categorical measurement of emotional and behavioral problems. Journal of Adolescence 96 (2), 360–369. https://doi.org/10.1002/jad.12280.

Fernández-Pinto, I., Santamaría, P., Sánchez-Sánchez, F., Carrasco, M.A., del Barrio, V., 2015a. SENA: Sistema de evaluación de niños y adolescentes. Manual de aplicación, corrección e interpretación. TEA Ediciones.

Fernández-Pinto, I., Santamaría, P., Sánchez-Sánchez, F., Carrasco, M.A., del Barrio, V., 2015b. SENA. Sistema de Evaluación de Niños y Adolescentes. Manual técnico. TEA Ediciones.

Goodman, R., 1997. The Strengths and Difficulties Questionnaire: A research note. Journal of Child Psychology and Psychiatry 38, 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x.

Goodman, R., Ford, T., Simmons, H., Gatward, R., Meltzer, H., 2000. Using the Strengths and Difficulties Questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. The British Journal of Psychiatry 177, 534–539. https://doi.org/10.1192/bjp.177.6.534.

Gore, W.L., Widiger, T.A., 2015. Assessment of dependency by the FFDI: Comparisons to the PID-5 and maladaptive agreeableness. Personality and Mental Health 9 (4), 258–276. https://doi.org/10.1002/pmh.1308.

Haslam, N., McGrath, M.J., Viechtbauer, W., Kuppens, P., 2020. Dimensions over categories: A meta-analysis of taxometric research. Psychological Medicine 50 (9), 1418–1432. https://doi.org/10.1017/s003329172000183x.

Instituto Nacional de Estadística, 2023. Defunciones según la Causa de Muerte. *Notas de prensa* [PDF]. https://www.ine.es/prensa/edcm_2022_d.pdf.

Kerns, C.M., Maddox, B.B., Kendall, P.C., Rump, K., Berry, L., Schultz, R.T., Souders, M. C., Bennett, A., Herrington, J., Miller, J., 2015. Brief measures of anxiety in non-treatment-seeking youth with autism spectrum disorder. A*utism: the International Journal of Research and Practice* 19 (8), 969–979. https://doi.org/10.1177/1362361314558465.

Kieling, C., Buchweitz, C., Caye, A., et al., 2024. Worldwide Prevalence and Disability From Mental Disorders Across Childhood and Adolescence: Evidence From the Global Burden of Disease Study. JAMA Psychiatry 81 (4), 347–356. https://doi.org/10.1001/jamapsychiatry.2023.505.

Kusters, M.S.W., Pérez-Crespo, L., Canals, J., Guxens, M., 2023. Lifetime prevalence and temporal trends of incidence of child's mental disorder diagnoses in Catalonia, Spain. Spanish Journal of Psychiatry and Mental Health 16 (1), 24–31. https://doi.org/10.1016/j.rpsm.2021.02.005.

Lebrun-Harris, L.A., Ghandour, R.M., Kogan, M.D., Warren, M.D., 2022. Five-Year Trends in US Children's Health and Well-being, 2016-2020. J*AMA Pediatrics* 176 (7), e220056. https://doi.org/10.1001/jamapediatrics.2022.0056.

Lee, J., Kim, S.J., Kim, D., Yang, S.J., Lee, S., Lee, A., Shin, Y., 2023. Computerized Version of the Kiddie Schedule for Affective Disorders and Schizophrenia for School-Age Children (KSADS-COMP): Development and Validation of the Korean Clinician-Administered Version. Psychiatry Investigation 20 (2), 174–179. https://doi.org/10.30773/pi.2022.0273.

Levis, B., Benedetti, A., Levis, A.W., Ioannidis, J.P.A., Shrier, I., Cuijpers, P., Gilbody, S., Kloda, L.A., McMillan, D., Patten, S.B., Steele, R.J., Ziegelstein, R.C., Bombardier, C. H., de Lima Osório, F., Fann, J.R., Gjerdingen, D., Lamers, F., Lotrakul, M., Loureiro, S.R., Löwe, B., Thombs, B.D., 2017. Selective Cutoff Reporting in Studies of Diagnostic Test Accuracy: A Comparison of Conventional and Individual-Patient-Data Meta-Analyses of the Patient Health Questionnaire-9 Depression Screening Tool. American Journal of Epidemiology 185 (10), 954–964. https://doi.org/10.1093/aje/kww191.

Li, F., Cui, Y., Li, Y., Guo, L., Ke, X., Liu, J., Luo, X., Zheng, Y., Leckman, J.F., 2022. Prevalence of mental disorders in school children and adolescents in China: diagnostic data from detailed clinical assessments of 17,524 individuals. Journal of Child Psychology and Psychiatry, and Allied Disciplines 63 (1), 34–46. https://doi.org/10.1111/jcpp.13445.

Muñiz, J., Hernández, A., Fernández-Hermida, J.R., 2020. Utilización de los test en España: El punto de vista de los psicólogos [Test use in Spain: The psychologists' viewpoint]. Papeles del Psicólogo 41 (1), 1–15. https://doi.org/10.23923/pap.psicol2020.2921.

Piqueras, J.A., Carrasco, M.A., 2018. Evaluación de problemas emocionales y nuevas tecnologías en los jóvenes. Revista de Estudios de Juventud (121), 91–115.

Polanczyk, G.V., Salum, G.A., Sugaya, L.S., Caye, A., Rohde, L.A., 2015. Annual research review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. Journal of Child Psychology and Psychiatry, and Allied Disciplines 56 (3), 345–365. https://doi.org/10.1111/jcpp.12381.

Reilly, C., Atkinson, P., Das, K.B., Chin, R.F., Aylett, S.E., Burch, V., Gillberg, C., Scott, R. C., Neville, B.G., 2014. Screening for mental health disorders in active childhood epilepsy: population-based data. Epilepsy Research 108 (10), 1917–1926. https://doi.org/10.1016/j.eplepsyres.2014.09.028.

Reitsma, J.B., Rutjes, A., Whiting, P., Yang, B., Leeflang, M.M., Bossuyt, P.M., Deeks, J.J., 2022. Chapter 8: Assessing risk of bias and applicability. Draft version (4 October 2022). In: Deeks, J.J., Bossuyt, P.M., Leeflang, M.M., Takwoingi, Y. (Eds.), Cochrane handbook for systematic reviews of diagnostic test accuracy version 2. Cochrane. https://training.cochrane.org/handbook-diagnostic-test-accuracy/current.

Reynolds, C.R., Kamphaus, R.W., 1992. Behavior assessment system for children (BASC). American Guidance Service.

Reynolds, C.R., Kamphaus, R.W., 2004. Behavior Assessment System for Children, 2nd ed. AGS Publishing.

Sacco, R., Camilleri, N., Eberhardt, J., Umla-Runge, K., Newbury-Birch, D., 2024. A systematic review and meta-analysis on the prevalence of mental disorders among children and adolescents in Europe. European Child & Adolescent Psychiatry 33 (9), 2877–2894. https://doi.org/10.1007/s00787-022-02131-2.

Seconda, L., Kovess, V., Bessonneau, P., Pognon, D., Decio, V., Perrine, A.L., Regnault, N., 2024. Self-reported mental health in school-aged children using the Dominic interactive questionnaire in the French Enabee 2022 study. Preprint. https://doi.org/10.21203/rs.3.rs-4300080/v1.

Sherr, L., Cluver, L., Tomlinson, M., Idele, P., Banati, P., Anthony, D., Roberts, K., Haag, K., Hunt, X., 2021. Mind Matters: Lessons from past crises for child and adolescent mental health during COVID-19. UNICEF Office of Research - Innocenti.

Shorey, S., Ng, E.D., Wong, C.H., 2022. Global prevalence of depression and elevated depressive symptoms among adolescents: A systematic review and meta-analysis. British Journal of Clinical Psychology 61 (2), 287–305. https://doi.org/10.1111/bjc.12333.

Solmi, M., Radua, J., Olivola, M., Croce, E., Soardo, L., Salazar de Pablo, G., Il Shin, J., Kirkbride, J.B., Jones, P., Kim, J.H., Kim, J.Y., Carvalho, A.F., Seeman, M.V., Correll, C.U., Fusar-Poli, P., 2022. Age at onset of mental disorders worldwide: Large-scale meta-analysis of 192 epidemiological studies. Molecular Psychiatry 27 (1), 281–295. https://doi.org/10.1038/s41380-021-01161-7.

Stockings, E., Degenhardt, L., Lee, Y.Y., 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. Journal of Affective Disorders 174, 447–463. https://doi.org/10.1016/j.jad.2014.11.061.

Sanitaria, Subdirección General de Información, 2021. Salud mental en datos: prevalencia de los problemas de salud y consumo de psicofármacos y fármacos relacionados a partir de registros clínicos de atención primaria. BDCAP Series 2, 76. Available online. https://www.sanidad.gob.es/estadEstudios/estadisticas/estadisticas/estMinisterio/SIAP/Salud_mental_datos.pdf (accessed on 20 May 2023).

Townsend, L., Kobak, K., Kearney, C., Milham, M., Andreotti, C., Escalera, J., Alexander, L., Gill, M.K., Birmaher, B., Sylvester, R., Rice, D., Deep, A., Kaufman, J., 2020. Development of Three Web-Based Computerized Versions of the Kiddie Schedule for Affective Disorders and Schizophrenia Child Psychiatric Diagnostic Interview: Preliminary Validity Data. Journal of the American Academy of Child and Adolescent Psychiatry 59 (2), 309–325. https://doi.org/10.1016/j.jaac.2019.05.009.

World Health Organization, 2024. Depression. https://www.who.int/news-room/fact-sheets/detail/depression.