

# The Montgomery and Åsberg Depression Rating Scale – self-assessment for use in adolescents: an evaluation of psychometric and diagnostic accuracy

I. Ntini, S. Vadlin, S. Olofsdotter, M. Ramklint, K. W. Nilsson, I. Engström & K. Sonnby

To cite this article: I. Ntini, S. Vadlin, S. Olofsdotter, M. Ramklint, K. W. Nilsson, I. Engström & K. Sonnby (2020) The Montgomery and Åsberg Depression Rating Scale – self-assessment for use in adolescents: an evaluation of psychometric and diagnostic accuracy, *Nordic Journal of Psychiatry*, 74:6, 415-422, DOI: [10.1080/08039488.2020.1733077](https://doi.org/10.1080/08039488.2020.1733077)

To link to this article: <https://doi.org/10.1080/08039488.2020.1733077>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 03 Mar 2020.



Submit your article to this journal [↗](#)



Article views: 1118



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

## The Montgomery and Åsberg Depression Rating Scale – self-assessment for use in adolescents: an evaluation of psychometric and diagnostic accuracy

I. Ntini<sup>a</sup>, S. Vadlin<sup>b</sup>, S. Olofsdotter<sup>b</sup>, M. Ramklint<sup>c</sup>, K. W. Nilsson<sup>b</sup>, I. Engström<sup>a</sup> and K. Sonnbj<sup>b</sup>

<sup>a</sup>University Health Care Research Center, Faculty of Medicine and Health, Universitetssjukvårdens forskningscentrum (UFC), Örebro University, Örebro, Sweden; <sup>b</sup>Center for Clinical Research, Uppsala University, County of Västmanland, Hospital of Region Västmanland, Västerås, Sweden; <sup>c</sup>Psychiatry – Department of Neuroscience, Uppsala University, Uppsala, Sweden

### ABSTRACT

**Background:** The Montgomery–Åsberg Depression Rating Scale – Self Assessment (MADRS-S) is used to assess symptom severity in major depressive disorder (MDD) among adolescents, but its psychometric properties and diagnostic accuracy are unclear.

**Aim:** The aim of this study was to explore psychometric properties, including diagnostic accuracy, of the MADRS-S in adolescent psychiatric outpatients.

**Method:** Adolescent psychiatric outpatients ( $N=105$ , mean age 16 years, 46 boys) completed the MADRS-S and were interviewed using the Kiddie Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS).

**Results:** In principal component analysis, two components with eigenvalues of 4.6 and 1.3 explained 51.1% and 14.4% of the variance, respectively. On the first component loaded items assessing Mood, Feelings of unease, Appetite, Initiative, Pessimism, and Zest for life. On the second component loaded items assessing Sleep, Ability to concentrate, and Emotional involvement. Cronbach's alpha (internal consistency) for all items was 0.87. Spearman's rho was 0.68 for concurrent validity (correlation between total MADRS-S-score and K-SADS MDD severity score). In receiver-operating characteristic analysis, the area under the curve was 0.86 (95% confidence interval 0.78–0.93,  $p<.001$ ). For all the participants, the highest combined sensitivity and specificity were reached using cut-offs of 15 and 16 (sensitivity 0.82, specificity 0.86). Optimizing sensitivity for MDD, with specificity still  $\geq 0.5$ , cut off for all was 9, for boys 7 and for girls 10.

**Conclusion:** Psychometric properties of MADRS-S showed good reliability and validity as well as satisfying diagnostic accuracy, indicating good to excellent properties for MDD screening of adolescent psychiatric patients.

### ARTICLE HISTORY

Received 11 January 2019

Revised 16 September 2019

Accepted 12 February 2020

### KEYWORDS

Validation studies;  
sensitivity and specificity;  
MDD; adolescent

### Background

Major depressive disorder (MDD) is one of the most common reasons adolescents seek help from child and adolescent psychiatric services [1]. MDD in adolescence is associated with functional impairment, increased risk for disruptive disorders, substance use disorder, suicide attempts and suicide [2,3]. Therefore, early identification and treatment is crucial [2,3]. For that purpose, self-rating scales are used widely in child and adolescent psychiatry, both as screening instruments and for assessment of symptom severity [4,5]. Self-rating scales are time and cost-effective, and have the additional advantage of involving the patient in both the diagnostic process and evaluation of treatment [6,7]. However, many of the instruments commonly used has not been sufficiently psychometrically evaluated for all age groups and settings [7–9].

The Montgomery and Åsberg Depression Rating Scale – Self Assessment (MADRS-S) is the self-rating version of the original clinician-rated Montgomery and Åsberg Depression Rating Scale (MADRS) [10–12]. The MADRS-S is a widely used rating scale for assessment of depression severity [10–12]. According to previous studies in adult populations, MADRS-S seems to be a reliable and valid instrument with good sensitivity for assessing MDD in primary care and psychiatric settings, as well as for measuring changes during treatment (Table 1) [7,8,13–18].

The MADRS-S is recommended by the Swedish Association for Child and Adolescent Psychiatrists for assessment of MDD severity but not for screening [4,19]. To our knowledge, no previous studies have examined the psychometric properties and diagnostic accuracy of the MADRS-S in adolescents in child and adolescent psychiatric settings. The aim of this study was to fill this knowledge gap. We

Table 1. Previous studies MADRS-S: dimensionality, internal consistency, concurrent validity, and diagnostic accuracy.

Author	Population	Dimensionality	Internal consistency	Concurrent validity	Diagnostic accuracy
Svanborg and Åsberg, 1994	28 adult psychiatric outpatients with MDD anxiety disorder + 2 inpatients			Pearson's correlation coefficient (MADRS and MADRS-S) = 0.80 and 0.94 at Time 1 and Time 2, respectively	
Mattila-Evenden, 1996	101 adult psychiatric outpatients			Pearson's correlation coefficient (MADRS and MADRS-S) = 0.83	
Svanborg and Åsberg, 2001	86 adults with affective, anxiety, personality disorders, 37 inpatients and 49 psychiatric outpatients			Pearson's correlation coefficient (MADRS-S-BDI) = 0.87	
Fantino and Moore, 2009	278 adults with MDD	One factor (factor analysis) factor loading 0.50 (45% variance)	Cronbach's alpha = 0.84	Pearson's correlation coefficient (MADRS and MADRS-S) = 0.54	Cut-off for remission sensitivity 81.8%, specificity 75.4%, PPV 77.1%, NPV 80.3%
Bondolfi G et al. 2010	63 adult psychiatric outpatients with affective disorder	Time 1: two components, eigenvalues 4.2 and 1.1, 47.0% and 12.5% explained variance, respectively; loadings on the first component > 0.58 for all items Time 2: one component, eigenvalue 6.2, 68.8% explained variance; loadings > 0.76 for all items	Cronbach's alpha = 0.85 and 0.94 for Time 1 and Time 2, respectively	Pearson's correlation coefficients (MADRS and MADRS-S) = 0.81 at Time 1 and 0.91 at Time 2	
Cunningham JL et al. 2011	400 adults with MDD treated in primary care		Cronbach's alpha = 0.84–0.93	Spearman's rank correlation coefficient (MADRS and MADRS-S) = 0.56–0.81	
Wikberg, Carl et al. 2015	146 adults primary care patients with mild to moderate MDD		Cronbach's alpha = 0.76	Pearson's correlation coefficient (MADRS-S and BDI-II) = 0.62 and 0.66	
Yee, Anne et al. 2015	150 adult psychiatric outpatients (50 with MDD, 100 without)	one single factor 61.3% of variance	Cronbach's alpha = 0.78	Spearman's rank correlation coefficient (MADRS-S and BDI-II) = 0.78	AUC = 0.91 (95% CI 0.86–0.96) with/without a MDD cut-off $\geq 4$ , sensitivity 78%, specificity 86% (95% CI 0.86–0.96)

compared the MADRS-S with the semi-structured diagnostic interview Kiddie Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS), which was used as the reference test [20].

## Materials and methods

### Study design

This study is a diagnostic accuracy study that was conducted from September 2011 until June 2013.

### Setting and participants

The study was performed in two child and adolescent psychiatric outpatient clinics, situated in Västerås and Sala, in the County of Västmanland in Sweden. At the time of the study, Västerås and Sala had 130,000 and 20,000 inhabitants, respectively.

### Procedure

The study was approved by the Regional Ethical Review Board of Uppsala (Dnr. 2008/214).

The participants were consecutive adolescent psychiatric outpatients aged 12–17 years, who came for psychiatric evaluation during the study period [21]. Patients in need of an interpreter or patients in need of emergency care were excluded from being invited to the study. The inclusion of participants was interrupted during shorter periods, such as holidays. The total number of eligible patients was 202. Of these, 28 refused to participate, 45 were missed due to clinical and administrative staff workload, and four did not show up for the K-SADS interview. A further 20 patients either had an incomplete K-SADS or MADRS-S, leaving 105 participants with full information for the analyses.

In the letter of invitation to the first visit, the adolescents and their parents obtained written information about the study. At the first visit, the adolescents and parents were then also orally informed about the study before informed and signed consent was collected. In accordance with Swedish legislation, both parents and the adolescent under 15 years of age agreed to participate in the study whereas adolescents above 15 years of age gave their own consent. The first visit included an unstructured clinical assessment together with analogous reports by patients and parents using the Electronic Psychiatric Intake Questionnaire, which assesses psychosocial factors and common psychiatric symptoms [22]. At the second visit, the participants were interviewed using the reference test, K-SADS [20]. The adolescents and parents were interviewed together. Later on the same day, before receiving feedback from the interviewer about the total result of the K-SADS assessment, each participant was asked to complete the MADRS-S. Thereafter they were informed about the K-SADS results. The interviewer providing feedback was unaware of the MADRS-S result because the patients placed their questionnaire in an envelope that was sealed.

### MADRS-S

The clinician-rated MADRS was developed from the Comprehensive Psychopathological Rating Scale [11]. The MADRS includes 10 items rated on a 7-point (0–6) Likert scale, and the total score ranges from 0 to 60 [11]. The version for self-assessment, MADRS-S, includes nine of the 10 items: Mood, Feelings of unease, Sleep, Appetite, Ability to concentrate, Initiative, Emotional involvement, Pessimism, and Zest for life [15]. The MADRS-S exists in two scoring versions. In one scoring version, every item has four statements that are rated on a four-point Likert scale (0–3), with the possibility of marking half steps and with a total score of 27 points [15,18]. In the other version, every item is rated from zero to six (instead of half steps), with a total score of 54 points [8]. The latter was used in this study as the index test under investigation in relation to the reference test [23].

### K-SADS

The K-SADS is a semi-structured diagnostic interview that is appropriate for use in children and adolescents aged 6–18 years [20,24,25]. The interview is designed to assess current and past episodes of psychopathology according to the DSM-III-R and DSM-IV criteria, and is administered by interviewing the parent(s) and the child, and finally achieving summary ratings based on all sources of information (parent, child, school, chart, and other) [20,24,25]. The K-SADS has been shown to have excellent to good validity and reliability for MDD diagnosis in different studies and is one of the most commonly used structured diagnostic interview in child and adolescent psychiatry, both in clinical and research settings [13,18,20,24–32]. In this study, this lifetime version (PL) of the K-SADS in Swedish translation was used to assess the current diagnoses [23].

The K-SADS can also be used as a dimensional measure based on summation of depression scores on each of the 26 items covering MDD and two items covering dysthymia [33]. The options for each item covering MDD are scored 0–3. A score of 0 indicates no available information; score of 1 suggests the symptom is not present; score of 2 indicates sub-threshold levels of symptomatology; and score of 3 represents the threshold criteria. The dysthymia items are rated on a 0–2-point rating scale in which 0 implies no information; 1 implies the symptom is not present; and 2 implies the symptom is present. The K-SADS's total sum for symptoms of depression ranges from 0 to 82, and is henceforth called the K-SADS MDD symptom severity score.

### Interviewers

The interviewers were one specialist in child and adolescent psychiatry, one resident in child and adolescent psychiatry, two clinical psychologists and one clinical social worker; all working in child psychiatric services and trained in using the K-SADS. The specialist in child and adolescent psychiatry (KS) had previously attended a national course for becoming a trainer of other interviewers. She planned the basic training

for the others, which, in accordance with the Swedish standard for K-SADS training, combined four days of theoretical lectures and practice. Each interviewer's training was initially based on assessing individually recorded interviews made by 'master interviewers'. The results of the ratings were compared and discussed together. Thereafter, the interviewers videotaped themselves interviewing patients. The other interviewers assessed the recordings and discussed discrepancies. After the training and before the data collection, inter-rater reliability was calculated based on ratings of five interviews. The specialist's ratings were considered standard [34]. Throughout the study, one joint assessment occurred each month to ensure continuous high levels of agreement.

In total (before and during data collection) each interviewer rated 20, 17, 11 and 5 interviews, respectively. The inter-rater reliability for all interviewer pairs of the diagnoses assessed using Cohen's kappa was 0.86 (range 0.82–1.00). The kappa value for the prevalence- and bias-adjusted kappa (PABAK) was 0.93 (range 0.89–1.00) [35–37]. The results for an MDD diagnosis was 0.89 (range 0.74–1.00), and 0.93 (range 0.82–1.00), respectively.

### Statistical analysis

Statistical analyses were conducted using Statistical Package for the Social Sciences version 24.0. A value of  $p < .05$  was considered to be significant.

Kappa statistics were used to assess the degree of agreement between interviewers. The agreement is considered to be slight if kappa  $\leq 0.2$ , fair if kappa is 0.21–0.40, moderate if kappa is 0.41–0.60, substantial if kappa is 0.61–0.80, and almost perfect if kappa is 0.81–1.0 [38]. When the distribution of reported categories (diagnoses) is unevenly distributed, Cohen's kappa is less reliable [35,36]. Therefore, the PABAK was also used [36].

Univariate analyses were performed to identify differences between groups using a  $t$ -test for continuous and ordinal data [38]. Analyses of not normally distributed ordinal data were then validated with the non-parametric Mann–Whitney U test [38]. The chi-squared test was used for categorical variables [38]. The chi-squared test for proportions was also used to analyze sex differences of sensitivity at each cut-off score [39–41]. Analogous analyses were performed for specificity [39–41].

Dimensionality was explored using principal component analysis and rotation methods using varimax with Kaiser normalization due to limited sample size and not the promax, which can also measure the correlation between factors [42].

The internal consistency of the MADRS-S was assessed by calculating Cronbach's alpha [38,39]. Concurrent validity was assessed by calculating the Spearman rho correlation coefficient for the total MADRS-S score (score: 0–54) with the K-SADS MDD symptom severity score (score 0–82) [38,39].

Receiver-operating characteristic (ROC) curve analysis was used to calculate the area under the curve (AUC) as a measure of diagnostic accuracy and to identify any differences in the AUC between boys and girls [37]. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value

(NPV) for identifying MDD diagnosis were calculated for cut-offs of 7 to 20 [38,39]. All analyses were performed for the total sample and for boys and girls separately [39–41].

## Results

### Description of the sample

The proportions of boys and girls within the total sample did not differ significantly between participants and non-participants ( $\chi^2 = 0.829$ ,  $p = .365$ ). Similarly, the mean age did not differ between participants and non-participants (15.8 years vs., 15.0 years,  $t = 3.466$ ,  $p = .977$ ).

Among the 105 participants with full information included in the analyses, 46 were boys (44%). Of the total 105 participants, 49 (47%) had an MDD diagnosis: 16 (35%) boys and 33 (56%) girls ( $\chi^2 = 4.645$ ,  $p = .003$ ). Seventy-one participants (68%) had more than one diagnosis: 26 boys (57%) and 45 girls (75%) ( $\chi^2 = 4.604$ ,  $p = .032$ ).

### Psychometric properties of MADRS-S

#### Dimensionality

Two components were found with eigenvalues of 4.6 and 1.3, which explained 51.1% and 14.4% of the variance, respectively; these values were confirmed by the scree plot. Items loading on the first component showed the following factor loadings: Mood 0.87, Feelings of unease 0.82, Appetite 0.64, Emotional involvement 0.78, Pessimism 0.83, and Zest for life 0.83. Items loading on the second component showed the following factor loadings: Sleep 0.41, Ability to concentrate 0.89, and Initiative 0.84.

#### Internal consistency

Internal consistency of the nine items in the MADRS-S showed Cronbach's alpha of 0.87 (0.82 for boys and 0.86 for girls). Internal consistency for six items of the first component according to factor analysis was 0.90 (0.88 for boys and 0.89 for girls). For the three items of the second component, 0.65 (0.71 for boys and 0.51 for girls).

#### Concurrent validity

Concurrent validity, measured as the correlation between the total score on the MADRS-S and the K-SADS MDD symptoms severity score, showed a Spearman rho of 0.71 for all participants (0.66 for boys and 0.67 for girls). Likewise, Spearman rho for the items of the first component was 0.70 (0.66 for boys and 0.60 for girls). For the second component, a result of 0.44 for all (0.22, but not statistically significant, for boys, and 0. significant for girls).

#### Diagnostic accuracy of the MADRS-S

In the ROC analysis, the AUC for all participants was 0.86 (95% confidence interval (CI) 0.78–0.93,  $p < .001$ ); the respective values were 0.84 (95% CI 0.71–0.96,  $p < .001$ ) for boys and 0.86 (95% CI 0.75–0.96,  $p < .001$ ) for girls. The AUC



values did not differ significantly between boys and girls ( $p = .803$ ) (Figure 1). The AUC for the first component was 0.88 (95% CI 0.81-0.95,  $p < .001$ ); 0.90 for boys (95% CI 0.82-0.99,  $p < .001$ ), and 0.85 for girls (95% CI 0.74-0.96,  $p < .001$ ). For the second component 0.70 (95% CI 0.60-0.80,  $p = .001$ ); 0.66 for boys (95% CI 0.51-0.81,  $p = .076$ ) and 0.72 for girls (95% CI 0.60-0.85,  $p = .004$ ).

Sensitivity, specificity, PPVs and NPVs for cut-off scores 7–20 on the MADRS-S are displayed in Table 2. Sex differences regarding sensitivity as well as specificity were found at all cut-off scores ( $p < .05$ ). For all participants, the highest combined sensitivity and specificity were at cut-offs of 15 and 16 (sensitivity 0.82, specificity 0.86). Highest combined sensitivity for boys was at cut-off 11 and for girls 15 and 16. Optimizing sensitivity for MDD, with specificity still  $\geq 0.5$ , cut-off was 9 for the whole sample, for boys 7 and girls 10, respectively [43].

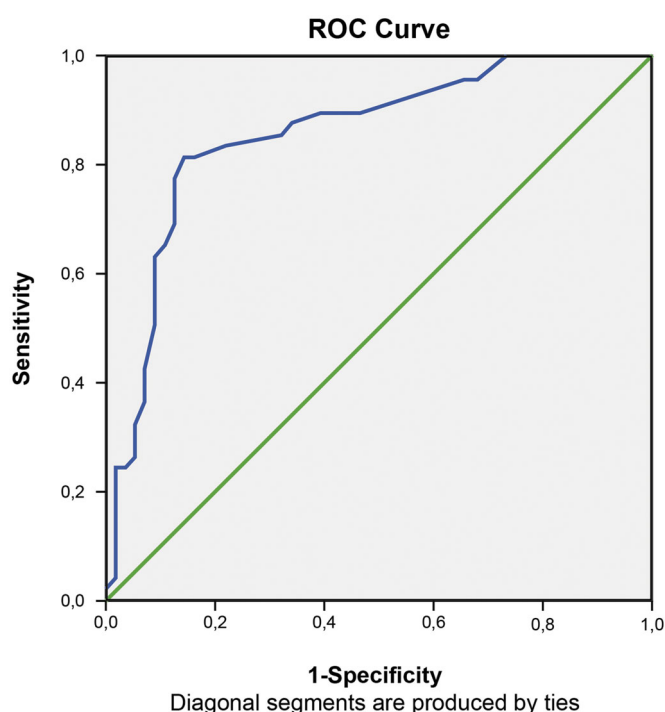


Figure 1. Receiver-operating characteristic (ROC) analysis of the MADRS-S in the sample of 105 adolescent psychiatric outpatients.

## Discussion

The unidimensionality of the MADRS-S could not be confirmed. However, the values for internal consistency, concurrent validity, and diagnostic accuracy were satisfactory for use in diagnosing adolescents in psychiatric settings, including the sensitivity for screening of MDD. Diagnostic accuracy for the whole scale, as measured by AUC, did not differ between boys and girls. Nevertheless, at each cut off score, there were sex differences for sensitivity and specificity, indicating that MADRS-S shows higher sensitivity and lower specificity for MDD in girls than in boys.

The results of the factor analysis showed two components for dimensionality. Results for internal consistency, concurrent validity and AUC, especially for the first component, were similar to the results for the whole scale. For boys analyses of AUC and concurrent validity showed non-significant results for the second component in, probably due to weak power. Previous studies of adults have reported unidimensionality for the MADRS-S (Table 1). Fantino and Moore [13] found that all items correlated with the first factor, although it was unclear whether the factor analysis detected more than one factor. Yee et al. [18] extracted a single component in the study population of 150 people, 100 of whom were healthy medical workers in the institute where the study was conducted. The only study whose results are consistent with ours is the study by Bondolfi et al. [7], who found two components at the first assessment but only one component at the second assessment after an intervention. Direct comparisons between our results and those of earlier studies of the MADRS-S are difficult because of differences in study design, methods, and study populations. Physical and neurobiological changes associated with maturation in adolescents might interfere with the symptoms assessed in the MADRS-S; for example, sleep, which may contribute to the two components found in the factor analysis in this study [44,45].

As shown in Table 1, the internal consistency in our study was within the range reported in earlier studies of adult populations, despite the differences in study populations, ages, methods, and settings. We investigated concurrent validity and found a strong correlation between the MADRS-S score and the K-SADS MDD severity score. Previous studies of adult populations calculated concurrent validity using other

Table 2. Sensitivity, specificity, positive predictive values (PPV) and negative predictive values (NPV) for cut-off scores 7–20.

Cut-off score	All				Boys $n = 46$				Girls $n = 59$			
	Sens	Spec	PPV	NPV	Sens	Spec	PPV	NPV	Sens	Spec	PPV	NPV
7	0.96	0.34	0.56	0.90	0.88	0.50	0.48	0.88	1.00	0.15	0.60	1.00
8	0.91	0.46	0.60	0.87	0.81	0.63	0.54	0.86	0.97	0.27	0.63	0.88
9	0.90	0.54	0.63	0.86	0.75	0.67	0.55	0.83	0.97	0.38	0.67	0.91
10	0.90	0.61	0.67	0.87	0.75	0.70	0.57	0.84	0.97	0.50	0.71	0.93
11	0.88	0.66	0.69	0.86	0.75	0.77	0.63	0.85	0.94	0.54	0.72	0.88
12	0.86	0.68	0.70	0.84	0.69	0.77	0.61	0.82	0.94	0.58	0.74	0.88
13	0.84	0.79	0.77	0.85	0.63	0.90	0.77	0.82	0.94	0.65	0.78	0.89
14	0.82	0.84	0.82	0.84	0.63	0.93	0.83	0.82	0.91	0.73	0.81	0.86
15	0.82	0.86	0.83	0.84	0.63	0.93	0.83	0.82	0.91	0.77	0.83	0.87
16	0.82	0.86	0.83	0.84	0.63	0.93	0.83	0.82	0.91	0.77	0.83	0.87
17	0.78	0.88	0.84	0.82	0.56	0.97	0.90	0.81	0.88	0.77	0.83	0.83
18	0.69	0.88	0.83	0.77	0.44	0.97	0.88	0.76	0.82	0.77	0.82	0.77
19	0.65	0.89	0.84	0.75	0.44	0.97	0.88	0.76	0.76	0.81	0.83	0.72
20	0.63	0.91	0.86	0.86	0.44	0.97	0.88	0.76	0.73	0.85	0.86	0.71

reference tests of depressive symptoms (Table 1). These results showed moderate to strong correlations with the clinically rated MADRS, very strong correlations with the Beck Depression Rating Scale (BDI) and strong correlations with the second version of the Beck Depression Rating Scale (BDI-II) [7,8,13–18]. No previous study has, to our knowledge, used a diagnostic interview as the reference test for assessing concurrent validity. Because of the use of different methods and different populations (e.g. healthy participants, adult patients), the results are not directly comparable between studies, although they trend in the same direction.

The ROC analyses of the MADRS-S supported very good diagnostic accuracy, which is consistent with previous studies by Fantino and Moore [13] and Yee et al. [18], despite the difficulties with making direct comparisons due to differences in study design, methods, and study populations. As shown in Table 2, the highest combined sensitivity and specificity was found at cut-off scores of 15 and 16 for all participants. When applying the cut-off scores, in addition to the differences mentioned above, our results contrast with those from previous studies because different MADRS-S versions were used [13,18].

Even with high combined sensitivity and specificity, a questionnaire alone is insufficient basis for making an MDD diagnosis since factors such as symptom duration, clinical impairment and differential diagnostic considerations also needs to be assessed [46,47]. In clinical work, a diagnostic tool with high sensitivity for identifying patients in need of further diagnostic assessment is important [46]. Therefore, high sensitivity for detecting MDD may be the most important trait [43]. In that context, a diagnostic tool with a sensitivity of 0.80 and specificity of 0.50 has been regarded as sufficient [43]. For the MADRS-S, a screening cut-off score of 9 improves sensitivity further, so that nine of 10 patients with MDD would be identified as having MDD with specificity  $\geq 0.50$ .

The sex differences regarding the optimal cut-off score could be explained by several potential factors. Biological aspects like the neurological maturation as well as psychosocial factors for boys and girls of the same age can have an impact on group level [48]. Higher levels of maturation among girls may affect the ability to identify own emotions, ability of self-observation and consequently also to report symptoms in questionnaires [49]. In fact, girls are known to report higher levels of symptoms than boys on self-reports [49]. Moreover, males are more reluctant help-seekers than females, which may also affect the willingness to report more severe symptoms compared to females [50,51]. Another reason for sex differences could be that the MADRS-S questionnaire provides no information about symptoms of irritability, which may occur more often in depressed males than females [52]. In DSM-IV and in the K-SADS-interview, the symptom of irritable mood is one of the obligate symptoms to identify MDD in adolescents [24,47]. Altogether these aspects may contribute to why MADRS-S as assessment tool has lower sensitivity for boys. Irritability as a potential item could theoretically improve content validity and diagnostic accuracy for MDD for both sexes.

Our results should be interpreted with respect to some limitations. Even though there were no significant differences in sex and age distribution between the drop-outs and participants, the external drop-out rate was high. Severely depressed patients might have dropped out more frequently, and those who needed emergency psychiatric care were also excluded from the study. The study sample may therefore represent a less impaired sample than the true patient population. This might have affected the PPV and NPV values, who are susceptible to prevalence. However, one-third of the participants had an MDD diagnosis and more than two-thirds showed comorbidity according to the K-SADS.

Another limitation was that the MADRS-S was administered after the K-SADS interview. It is not possible to say to what extent the K-SADS interview may have altered the MADRS-S report. Neither does it mimic a screening procedure before the comprehensive diagnostic evaluation. On the other hand, the MADRS-S report occurred after the K-SADS interview and, therefore, the MADRS-S could not introduce bias into the MDD diagnosis according to the K-SADS, which was the reference.

The varimax method was used for factor analyses due to the limited power of the study sample. Other oblique transformation methods such as the promax method could provide further information about the correlation between the factors found. However, as the power is limited, results may lack stability and in addition to that, complementary analyses with the promax method provide similar results to those shown.

The strengths of this study were that the participants came from two sites of different size and they both comprised consecutively referred patients. Moreover, the result of the MADRS-S assessment was compared with the diagnosis of MDD based on a diagnostic interview, the K-SADS, as the reference test, and not with a clinical diagnosis or another rating scale. Furthermore, the diagnostic interviews were made by trained interviewers with good to excellent inter-rater reliability during the data collection. Another strength is that the time frame between the index test, the MADRS-S, and the reference test (K-SADS) was less than one day, which inhibits time bias or fluctuation of symptoms between measures. Therefore, despite these limitations, our results may be generalizable to other psychiatric outpatient settings for adolescents.

In conclusion, this study shows that the MADRS-S has sufficient psychometric properties and appropriate diagnostic accuracy for use as an assessment tool for MDD in adolescents in general psychiatric outpatient care in Sweden. The unidimensionality of the scale could not be proven and further investigation is needed. As for other self-report instruments, the MADRS-S is not sufficient for diagnosing depression, but it can be part of a best-estimate diagnostic procedure and help to support the accuracy of clinician-diagnosed MDD [27].

## Acknowledgements

We would like to thank Professor Marie Åsberg for allowing us to validate the MADRS-S, all patients and parents for their participation, as well as all clinical staff for help with the research logistics in the psychiatric

clinics of Sala and Västerås, which made this study possible. We are also very grateful to Mattias Rehn for excellent data management.

## Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

## Funding

This work was supported by the County Council of Västmanland [grant number LTV-473811], [grant number LTV-466871], [grant number LTV-398781], [grant number LTV-379991], [grant number LTV-375811], [grant number LTV369201], [grant number LTV-353451] and the Uppsala and Örebro Regional Research Council [grant number RFR-475881], [grant number RFR-376361].

## Notes on contributors

**Iordana Ntini**, Doctoral Student, Örebro University, Universitetssjukvårdens forskningscentrum (UFC) S-huset vän. 2 Universitetssjukhuset Örebro 701 85 Örebro.

**Sofia Vadlin**, post doctoral at Centre for Clinical Research, County of Västmanland, Västmanlands sjukhus Västerås 721 89 Västerås.

**Susanne Olofsdotter**, post doctoral at Centre for Clinical Research, County of Västmanland. Västmanlands sjukhus Västerås, 721 89 Västerås.

**Ramklint Mia**, Professor at Department of Neuroscience, Child and Adolescent Psychiatry, Akademiska sjukhuset 751 85 UPPSALA

**Kent Nilsson**, Adjunct professor at Centre for Clinical Research, County of Västmanland, Västmanlands sjukhus Västerås, 721 89 Västerås

**Ingemar Engström**, Adjunct professor at Örebro University, Universitetssjukvårdens forskningscentrum (UFC) S-huset vän. 2 Universitetssjukhuset Örebro, 701 85 Örebro

**Karin Sonnbj**, post doctoral at Centre for Clinical Research, County of Västmanland, Västmanlands sjukhus Västerås, 721 89 Västerås

## References

- [1] Staller JA. Diagnostic profiles in outpatient child psychiatry. *Am J Orthopsychiatry*. 2006;76(1):98–102.
- [2] Lewis M. Child and adolescent psychiatry: a comprehensive textbook. Fourth edition, Martin A., Volkman F.R, editors. Philadelphia (PA): Lippincott Williams & Wilkins; 2002: 503–513
- [3] Lundervold AJ, Breivik K, Posserud MB, et al. Symptoms of depression as reported by Norwegian adolescents on the Short Mood and Feelings Questionnaire. *Front Psychol*. 2013;4: 613–24062708.
- [4] Barn- och ungdomspsykiatrins metoder En nationell inventering. [Child and adolescent psychiatry's methods. National Inventory] Västerås: Socialstyrelsen; 2009. Available from: <https://docplayer.se/994912-Barn-och-ungdomspsykiatrins-metoder-en-nationell-inventering.html>.
- [5] Diagnostik och uppföljning av förstämningssyndrom En systemisk litteratur översikt. [Diagnostics and follow-up of affective disorders: a systemic literature overview]. Stockholm: SBU Swedish Council on Health Technology Assessment; 2012.
- [6] Beck AT. Cognitive therapy of depression (The Guilford clinical psychology and psychotherapy series, 99-0267920-X). New York: Guilford; 1979.
- [7] Bondolfi G, Jermann F, Rouget BW, et al. Self- and clinician-rated Montgomery-Asberg Depression Rating Scale: evaluation in clinical practice. *J Affect Disord*. 2010;121(3):268–272.
- [8] Cunningham JL, Wernroth L, von Knorring L, et al. Agreement between physicians' and patients' ratings on the Montgomery-Asberg Depression Rating Scale. *J Affect Disord*. 2011;135(1–3): 148–153.
- [9] Moller HJ. Rating depressed patients: observer- vs self-assessment. *Eur Psychiatry*. 2000;15(3):160–172.
- [10] Carmody TJ, Rush AJ, Bernstein I, et al. The Montgomery Asberg and the Hamilton ratings of depression: a comparison of measures. *Eur Neuropsychopharmacol*. 2006;16(8):601–611.
- [11] Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979;134(4):382–389.
- [12] Williams JB, Kobak KA. Development and reliability of a structured interview guide for the Montgomery Asberg Depression Rating Scale (SIGMA). *Br J Psychiatry*. 2008;192(1):52–58.
- [13] Fantino B, Moore N. The self-reported Montgomery-Asberg Depression Rating Scale is a useful evaluative tool in Major Depressive Disorder. *BMC Psychiatry*. 2009;9(1):26.
- [14] Mattila-Evenden M, Svanborg P, Gustavsson P, et al. Determinants of self-rating and expert rating concordance in psychiatric out-patients, using the affective subscales of the CPRS. *Acta Psychiatr Scand*. 1996;94(6):386–396.
- [15] Svanborg P, Asberg M. A new self-rating scale for depression and anxiety states based on the Comprehensive Psychopathological Rating Scale. *Acta Psychiatr Scand*. 1994;89(1):21–28.
- [16] Svanborg P, Asberg M. A comparison between the Beck Depression Inventory (BDI) and the self-rating version of the Montgomery Asberg Depression Rating Scale (MADRS). *J Affect Disord*. 2001;64(2–3):203–216.
- [17] Wikberg C, Nejati S, Larsson ME, et al. Comparison between the Montgomery-Asberg Depression Rating Scale-Self and the Beck Depression Inventory II in primary care. *Prim Care Companion CNS Disord*. 2015; 17(3): 10.4088/PCC.14m01758. Published online 2015 Jun 25. doi: 10.4088/PCC.14m01758
- [18] Yee A, Yassim AR, Loh HS, et al. Psychometric evaluation of the Malay version of the Montgomery-Asberg Depression Rating Scale (MADRS-BM). *BMC Psychiatry*. 2015;15(1):19.
- [19] Jarbin H, Von Knorring A-L, Zetterqvist M. Riktlinje Depression 2014 [Guidelines depression 2014]. Svenska föreningen för barn- och ungdomspsykiatri [Swedish Association for Child and Adolescent Psychiatry]; 2014. Available from: <http://www.sfbup.se/wp-content/uploads/2017/01/SFBUPRiktlinjeDepression2014.pdf>.
- [20] Ambrosini PJ. Historical development and present status of the schedule for affective disorders and schizophrenia for school-age children (K-SADS). *J Am Acad Child Adolesc Psychiatry*. 2000; 39(1):49–58.
- [21] Torres Soler C, Olofsdotter S, Vadlin S, et al. Diagnostic accuracy of the Montgomery-Asberg Depression Rating Scale parent report among adolescent psychiatric outpatients. *Nordic Journal of Psychiatry*. 2018;72(3):184–190.
- [22] Lövenhag S. Substance use in Swedish adolescents: the importance of co-occurring psychiatric symptoms and psychosocial risk. Uppsala, Uppsala University; 2015. [cited 2015]. Available from: <https://www.diva-portal.org/smash/get/diva2:806551/FULLTEXT01.pdf>.
- [23] Cohen JF, Korevaar DA, Gatsonis CA, et al. STARD for Abstracts: essential items for reporting diagnostic accuracy studies in journal or conference abstracts. *BMJ*. 2017;358:j3751.
- [24] Kaufman J, Birmaher B, Brent D, et al. Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *J Am Acad Child Adolesc Psychiatry*. 1997;36(7):980–988.
- [25] Kim YS, Cheon KA, Kim BN, et al. The reliability and validity of Kiddie-Schedule for Affective Disorders and Schizophrenia-Present and Lifetime Version-Korean version (K-SADS-PL-K). *Yonsei Med J*. 2004;45(1):81–89.
- [26] Brasil HH, Bordin IA. Convergent validity of K-SADS-PL by comparison with CBCL in a Portuguese speaking outpatient population. *BMC Psychiatry*. 2010;10(1):83.



- [27] Jarbin H, Andersson M, Rastam M, et al. Predictive validity of the K-SADS-PL 2009 version in school-aged and adolescent outpatients. *Nord J Psychiatry*. 2017;71(4):270–276.
- [28] Lauth B, Arnkelsson GB, Magnusson P, et al. Validity of K-SADS-PL (Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Lifetime Version) depression diagnoses in an adolescent clinical population. *Nord J Psychiatry*. 2010; 64(6):409–420.
- [29] Polanczyk GV, Eizirik M, Aranovich V, et al. Interrater agreement for the schedule for affective disorders and schizophrenia epidemiological version for school-age children (K-SADS-E). *Rev Bras Psiquiatr*. 2003;25(2):87–90.
- [30] Shanee N, Apter A, Weizman A. Psychometric properties of the K-SADS-PL in an Israeli adolescent clinical population. *Isr J Psychiatry Relat Sci*. 1997;34(3):179–186.
- [31] Ulloa RE, Ortiz S, Higuera F, et al. [Interrater reliability of the Spanish version of Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Lifetime version (K-SADS-PL)] . *Actas Esp Psiquiatr*. 2006;34(1):36–40.
- [32] Kragh K, Husby M, Melin K, et al. Convergent and divergent validity of the schedule for affective disorders and schizophrenia for school-age children – present and lifetime version diagnoses in a sample of children and adolescents with obsessive-compulsive disorder. *Nord J Psychiatry*. 2019;73(2):111–117.
- [33] Kaufman J, Birmaher B, Brent D. Kiddie-Sads – Present and Lifetime Version (K-SADS-PL). Diagnostic Interview Version 1.0 of October 1996. [cited 2020 Feb 24]. Available from: <http://www.icctc.org/PMM%20Handouts/Kiddie-SADS.pdf>.
- [34] Sonnby K, Skordas K, Olofsdotter S, et al. Validation of the World Health Organization Adult ADHD Self-Report Scale for adolescents. *Nord J Psychiatry*. 2015;69(3):216–223.
- [35] Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol*. 1993;46(5):423–429.
- [36] Chen G, Faris P, Hemmelgarn B, et al. Measuring agreement of administrative data with chart data using prevalence unadjusted and adjusted kappa. *BMC Med Res Methodol*. 2009;9(1):5.
- [37] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982; 143(1):29–36.
- [38] Bland M. An introduction to medical statistics. 4th ed. Oxford, Oxford University Press; 2015.
- [39] Altman A, Douglas G. Statistics with confidence: confidence intervals and statistical guidelines. 2nd ed. London: BMJ Books; 2000.
- [40] Campbell I. Chi-squared and Fisher-Irwin tests of two-by-two tables with small sample recommendations. *Statist Med*. 2007; 26(19):3661–3675.
- [41] Richardson JT. The analysis of 2 x 2 contingency tables–yet again. *Statist Med*. 2011;30(8):890–891.
- [42] Sullivan JJ, Pett MA, Lackey NR. Making sense of factor analysis: The use of factor analysis for instrument development in health care research. Thousand Oaks, Calif.: Sage; 2003.
- [43] Runeson B. Interview instrument provides no reliable assessment of suicide risk. Scientific support is lacking according to report from the Swedish Council on Health Technology Assessment (SBU). *Lakartidningen*. 2015;112:26646963.
- [44] Colrain IM, Baker FC. Changes in sleep as a function of adolescent development. *Neuropsychol Rev*. 2011;21(1):5–21.
- [45] Dahl RE, Lewin DS. Pathways to adolescent health sleep regulation and behavior. *J Adolesc Health*. 2002;31(6):175–184.
- [46] DSM-IV-TR diagnostic and statistical manual of mental disorders. 4th ed. Text Revision. Washington, DC; American Psychiatric Association; 2000.
- [47] Diagnostic and statistical manual of mental disorders (DSM-5 (R)). 5th ed. American Psychiatric Association; Arlington, VA; 2013.
- [48] Mahone EM, Wodka EL. The neurobiological profile of girls with ADHD. *Dev Disabil Res Revs*. 2008;14(4):276–284.
- [49] Biederman J, Ball SW, Mick E, et al. Informativeness of maternal reports on the diagnosis of ADHD: an analysis of mother and youth reports. *J Atten Disord*. 2007;10(4):410–417.
- [50] Yap MB, Reavley N, Jorm AF. Where would young people seek help for mental disorders and what stops them? Findings from an Australian national survey. *J Affect Disord*. 2013;147(1–3): 255–261.
- [51] Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. *BMC Psychiatry*. 2010;10(1):113.
- [52] Rutz W, Walinder J, Rhimer Z, et al. Male depression–stress reaction combined with serotonin deficiency?. *Lakartidningen*. 1999; 96(10):1177–1178.