The Child Outcome Rating Scale: Validating a 4-item measure of psychosocial functioning in community and clinic samples of children aged 10 to 15

Casey, P[1], Patalay, P.[2], Deighton, J.[1], Miller, S.[3] and Wolpert, M.[1]

[1] Evidence Based Practice Unit (EBPU), UCL and Anna Freud National Centre for Children and Families, London, UK

[2] Centre for Longitudinal Studies and MRC Unit for Lifelong Health and Ageing, UCL

[3] International Center for Clinical Excellence, https://www.centerforclinicalexcellence.com/

Corresponding author

Dr Polly Casey, Evidence Based Practice Unit, UCL and the Anna Freud Centre, 12 Maresfield Gardens, London, NW3 5SU; ebpu@annafreud.org

Acknowledgements

The data used in this study was collected as part of HeadStart learning programme and supported by funding from the Big Lottery Fund. The content is solely the responsibility of the authors and it does not reflect the views of the Big Lottery Fund.

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Abstract

Objective: Psychosocial functioning is considered an important and valued outcome in relation to young people's mental health as a construct distinct from psychiatric symptomology, especially in the light of an increasing focus on transdiagnostic approaches. Yet, level of psychosocial functioning is rarely directly asked of young people themselves, despite now widespread recognition that the young person's perspective is valuable and often at odds with those of other reporters, such as parents or professionals. One possible reason for this is that the field lacks a clear agreed tool to capture this information in a non-burdensome way. To begin to address this gap, this paper describes psychometric analysis of the Child Outcome Rating Scale (CORS), a brief and highly accessible self-report measure of young people's psychosocial functioning already used extensively by mental health professionals around the world but with only limited data on psychometric robustness. **Method:** Using large community (n= 7822) and clinic (n= 2604) samples, we explore the factor structure, construct validity, internal consistency, differential item functioning, and sensitivity of the CORS. **Results:** We found that the CORS stands up to psychometric scrutiny, having found satisfactory levels of reliability, validity, and sensitivity in this sample. We also found that the CORS is suitable for use with young people as old as 15 years old. **Conclusion:** That the CORS has been found to be psychometrically robust while being highly feasible (brief, simple, easy to administer) for use in busy clinical settings, combined with the fact that the CORS has already been widely adopted by clinicians and young people, suggests CORS may be an important tool for international use.

Introduction

Children and young people's mental health is increasingly recognised as a global public health challenge (Deighton et al., 2018; Husky et al., 2018; Patel, Flisher, Hetrick, and McGorry, 2007). In the context of increasing numbers of young people in need of mental health support, it is vital that young people who are experiencing mental health problems and those who are seeking to help them have adequate tools with which to communicate or monitor the extent of difficulties, that do not place too great a burden on already over-stretched services. In addition, there is mounting interest in accurately measuring constructs such as psychosocial functioning for young people alongside psychiatric symptomology (Krause, Bear, Edbrooke-Childs, and Wolpert, in press; Sharpe et al., 2016). In this paper we seek to assess the psychometric robustness of a brief measure of young people's psychosocial functioning already widely used by mental health professionals around the world, the Child Outcome Rating Scale (CORS), but lacking extensive psychometric analysis to date. This paper presents the first independent large-scale analysis of this potentially universal tool.

Understanding psychosocial functioning among young people who may be experiencing mental health problems as a distinct construct is important. Psychosocial functioning steps away from psychiatric symptom severity and focus on specific mental health problems and is instead concerned with an individual's functional adaptation in everyday life (e.g. individual functioning, interpersonal (with friends or family) functioning, and functioning in the domains of school or work) regardless of a specific diagnosis category (Patalay et al., 2015). Joseph and Wood (2010), for example, have specifically called for greater consideration to measures of positive functioning, pointing out that psychiatry has for too long adhered to a narrow view of wellbeing as "the absence of distress and dysfunction" (p. 831). Conceptualising psychosocial functioning as distinct from psychopathology is particularly important with regard to child and adolescent mental health where recovery is often partial or out of reach (Costello, Foley, and Angold, 2006).

Two recent reviews have highlighted value placed by service users on outcomes beyond symptom reduction, emphasising psychosocial functioning as a key and valued indicator of improvement (Bradley, Murphy, Fugard, Nolas, and Law, 2013; Childs, Deighton, and Wolpert, 2013; Krause et al., 2018). A recent review by Krause and colleagues (2018) assessed the type of outcomes measured to assess treatment success in young people with adolescent depression, the specific measures used, and whose perspective was sought (clinician-, parent- or child-rated). Of interest was that functioning was the most commonly measured outcome after depressive symptoms (measured in 52% of the studies included), but this was almost always from the perspective of the clinician with only 6% of these studies including a measure of child-reported functioning. It would seem, then, that improved psychosocial functioning is an outcome valued by all concerned but it is seldom asked of young people themselves.

The importance of acquiring the perspectives of both the young person and their parent/carer however, is reinforced by the common finding that reports from parents and children tend to correlate quite poorly with one another (Achenbach, McConaughy and Howell, 1987; Cleridou, Patalay, and Martin, 2017; Sourander, Helstela and Helenius, 1999). Professional guidelines recommend that service outcomes should be sought from the perspective of the service users themselves (UK Department for Health, 2004), as well as their parents and carers (Batty et al., 2013).

The challenge is to offer young people a non-burdensome tool to communicate their level of psychosocial functioning. A measurement tool that is complicated and long is unlikely to taken up and routinely used in mental health settings, nor considered acceptable by young service users. However, increasing feasibility may compromise other psychometric properties. In the present paper, we explore whether the CORS is satisfactorily valid and reliable while at the same time being highly feasible.

The CORS is a brief, child-report measure of psychosocial functioning originally intended for young people aged 6-12 years, developed and only validated in the United States (US) to date (Duncan et al., 2006). Respondents answer the four questions using a visual analogue scale (VAS). While some researchers have highlighted problems with VASs in terms of young respondents' tendency to stick with the anchors of the scale (Pantell and Lewis, 1987), such scales are popular with some because of their ease of administration, while demonstrating reliability and validity comparable to Likert response scales (van Laerhoven, van der Zaag-Loonen and Derkx, 2004).

The CORS was adapted from the adult Outcomes Rating Scale (ORS; Miller, Duncan, Brown, Sparks, and Claud, 2003), and has been typically used as a tool in clinical practice to monitor progress and outcomes in psychological services. We are building on the preliminary work carried out by Duncan and colleagues, which provided initially promising results regarding the psychometric properties of the modified version of the measure for young people, but also had a number of limitations: 1) the sample available to the researchers was relatively small (n= 154 in the non-clinical sample and n= 119 in the clinical sample); 2) concurrent validity was not assessed in relation to the CORS, only the ORS.; 3) the approach to assessing the ability of the CORS to discriminate between the clinical and non-clinical sample was relatively unsophisticated (t-tests to compare mean baseline CORS scores from each sample); and 4) the researchers did not explore the possible presence of bias in items on the CORS according to different respondent subgroups (i.e. does the CORS measure the same latent construct (psychosocial functioning), in the same way, in different groups of respondents?). We attempt to address these shortcomings in the present paper, with UK community and clinic samples of children and adolescents.

The current study uses a large community (schools) sample (n=7822) to explore the factor structure, internal consistency, measurement invariance and structural invariance of the CORS. Using the entire age range of the available sample, we look into whether the CORS is also suitable for use with young people up to the age of 15 years (i.e. beyond the original upper age limit of 12 years). We explore the construct validity of the CORS by examining associations with measures of mental health outcomes (emotional and behavioural difficulties), related factors such as quality of life and resilience and protective factors including perceived support at school, home and in the community. Finally, we also investigate the measure in a clinical sample (n=2604) in order to assess the ability of the CORS to discriminate between clinical and community samples.

Methods

Participants

Community sample

This sample was drawn from a pilot phase of a wider project (HeadStart, funded by the National Lottery Community Fund), which involved assessing mental health and psychosocial functioning in a school-based community sample comprising young people from 90 participating schools, in 12 areas in England during Spring 2016.

The total sample across the 90 schools included 7834 pupils. Of these, 7822 young people completed at least one CORS item and 7609 completed the full CORS. Of the 7609 57.5% were female (n= 4374) and the mean age was 12.82 years (SD= 1.13). 17.0% had an identified special educational need (SEN, includes those with and without a statement of SEN, n= 1278); 19.0% were eligible for free school meals (FSM, n= 1278); 72.8% were from a White ethnic background (n= 5440); and 19.0% of the respondents' first language was not English.

As areas selected for HeadStart are those identified as high-need, the sample is not representative of the rest of the UK schools population[1].

Clinic sample

This sample was drawn from data collected between 2011 and 2015 from 81 publicly funded child and young people's mental health services in the UK. These services were all providing routine care as part of NHS provision and taking part in a national service improvement initiative[2] to ensure services were evidence based and outcomes informed. Most young people had been referred to the service from primary health care (51%)[3].

The sample included here comprised data from 2621 young service users, of whom 2604 completed the full CORS. We restricted the sample for this paper to young people aged 10 to 15 years-old, to be broadly in line with the available data in the community sample, which represented 60.94% of young people who provided CORS data at their first contact with the service. Of the 2604 young people in the clinic sample with full CORS data, 63.29% were female (n= 1648), the mean age was 12.97 years (SD= 1.68), and of the young people for whom there was available ethnicity data (84.91% of those with full CORS data, n= 2211) 50.96% were from a White ethnic background (n= 1327). Unfortunately we don't have information about SEN status, eligibility for FSM (or another index of SES) nor whether or not English was a first language for the clinic sample.

Measures

Psychosocial functioning

Psychosocial functioning was measured using the Child Outcome Rating Scale (CORS; Duncan et al., 2006), a four-item self-report measure for young people aged 6-12 years. Respondents answer questions about themselves (How am I doing?), family (How are things in my family?), school (How are things at school?), and their general sense of wellbeing/distress (How is everything going?). Responses are marked on a 10cm line, which has a happy face at the end on the right and an unhappy face at the end on the left. Respondents are asked

[1] According to data from the January 2017 school census based on state-funded primary and secondary schools (Department for Education).
[2] the Children and Young People's Improving Access to Psychological Therapies programme (CYP IAPT).
[3] The most common the most common presenting problems across services were family relationship difficulties (52%), depression/low mood (50%), and/or generalised anxiety (49%) (not mutually exclusive issues). Percentages are based on the full CYP-IAPT dataset, reported by Wolpert et al., (2016; https://www.corc.uk.net/media/1544/0505207_corc-report_for-web.pdf) and so a broad indication of presenting problems in the current sample.

to place a mark somewhere along the line that best represents how they are feeling. Total scores range from 0 to 40, with low scores representing poorer psychosocial functioning.

Mental health difficulties

Mental health difficulties were measured using the Me and My Feelings (MAMF; Deighton et al., 2013; Patalay, Deighton, Fonagy, Vostanis and Wolpert, 2014), a 16-item measure comprising two subscales; emotional and behavioural difficulties. Participants respond to short statements using a 3-point scale of 0 “never”, 1 “sometimes” or 2 “always”. Total scores on the emotional and behavioural subscales can range from 0-20 and 0-12 respectively, with higher scores representing greater levels of difficulties. Internal consistency in the sample was good, with Cronbach’s alphas of .83 for the emotional difficulties subscale, and .79 for the behavioural difficulties subscale.

Health-related quality of life

Health-related quality of life was measured using the EQ 5D-Y (Ravens-Sieberer et al., 2010), a 5-item measure on which respondents are asked to rate how they feel in relation to five dimensions: mobility (‘walking about’), self-care (‘looking after myself’), usual activities (‘doing usual activities’), pain and discomfort (‘having pain or discomfort’) and anxiety and depression (‘feeling worried, sad or unhappy’). Respondents answer questions about how they feel today, using a 3-point scale which equates to 1 “no problems”, 2 “some problems”, to 3 “a lot of problems”. Total scores range from 5 to 15, with higher scores representing poorer quality of life. Internal consistency for this sample was adequate, with a Cronbach’s alpha of .64.

Subjective current state of health

Subjective current state of health is also measured using EQ 5D-Y (Ravens-Sieberer et al, 2010), using a single visual analogue scale which ranges from 0 “the worst health you can imagine” to 100 “the best health you can imagine”. Respondents are asked to place a mark on a vertical line according to how they feel today. Higher scores represent better subjective health, however in the current study scores were reversed so that higher scores represent poorer subjective health.

Resilience

Resilience (risk and protective factors) was measured using the Student Resilience Scale (SRS; Sun and Stewart, 2003; Lereya et al., 2016). The SRS is a 47-item measure comprising 12 subscales measuring respondents’ perceptions of their individual characteristics as well as protective factors/sources of support in their environment. For the purpose of the wider project from which the current data are drawn (HeadStart), 10 of the 12 subscales have been used: family connection, school connection, community connection, participation in home and school life, participation in community life, self-esteem, empathy, problem solving, goals and aspirations and peer support. Respondents are asked to respondent to a series of statements (e.g. “When I need help, I find someone to talk to”) on a five-point Likert scale of 1 “never” to 5 “always”. Total scores vary according to each subscale, with high scores representing higher levels of personal qualities (e.g. self-esteem) or higher levels of external support. Internal consistency in this sample was good, with Cronbach’s alpha ranging from .73 to .93.

Procedure

Young people in the community sample completed the CORS as part of an online battery of questionnaires, in a teacher-facilitated lesson as part of their normal school day. Parental consent was obtained via the school prior to survey completion, and young people’s assent was also obtained via an extra page at the beginning of the online survey (University College London ethics project ID: 1530/006). It was stressed to respondents that neither their teachers nor their parents would ever be shown their individual responses.

Young people in the clinic sample completed the CORS at their first contact with the service (assessment) as part of the services’ routine outcome monitoring. With parental consent, data were then submitted quarterly to a central research team based at the Child Outcomes Research Consortium (CORC) according to the CYP IAPT data specification (CYP IAPT programme, 2013). Since these data were collected for the purpose of an audit no ethical permissions were required.

Analysis

See Table 1. for completeness of CORS data for both the community and clinic sample, plus distribution of scores for individual CORS items and total CORS.

Factor structure

We used a combination of exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) to examine the strength of the one factor solution in the community sample. First the sample was split in half by randomising the order of observations and extracting 50% of them. This was to allow us to perform a split-sample model development and subsequent model cross-validation strategy (Cabrera-Nguyen, 2010). EFA was carried out on half of the sample with oblique rotation using Stata 15, a combination of identifying Eigenvalues over 1 and examination of the scree plot was used to determine the number of factors to extract. The resulting solution was tested using multi-level CFA (accounting for clustering by schools) on the other half of the sample using MPlus v7.11 (Muthén & Muthén, 1998-2015), the estimator used was weighted least square mean and variance adjusted (WLSMV). Because the chi-square test for model fit has been shown to be unreliable in large sample sizes (Cheung and Rensvold, 2002; Meade, Johnson and Braddy, 2008), model fit was also assessed using other goodness of fit indices (comparative fit index (CFI), Tucker-Lewis Index (TLI), Root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR)). The criteria used to assess model goodness of fit were a CFI value ≥.90, a TLI value ≥.95, a RMSEA value <.08, and an SRMR value <.08 (Hooper, Coughlan, and Mullen, 2008; Kline, 2005).

Internal consistency

Cronbach's α and McDonald's ω were calculated based on the community sample, using Stata 15 to assess the reliability of the measures. Reliability coefficient values of .70 and above were regarded as acceptable (Nunnally, 1978).

Construct validity

To assess the degree to which the CORS is effectively measuring psychosocial functioning, we examine the association between CORS scores, mental health outcomes and protective factors in the community sample. To do this, correlations were estimated between the CORS total score and scores on the MAMF, EQ 5D-Y, and SRS. We anticipate that the CORS will be associated negatively with symptoms of mental health disorders, and positively with measures that assess protective factors and quality of life.

Measurement invariance

Differential item functioning (DIF) was carried out on the community sample using structural equation modelling (SEM) in MPlus v7.11 (Muthén & Muthén, 1998-2015) to explore the presence of bias in items on the CORS according to different respondent subgroups. Or in other words, does the CORS measure the same latent construct (psychosocial functioning), in the same way, in different groups of respondents? DIF was tested in relation to five characteristics: (1) gender (male n = 3326; female n = 4496), (2) eligibility for free school meals as an indicator of deprivation (eligible for FSM n = 1489; not eligible for FSM n = 6464), (3) presence of special educational needs (young people with SEN n = 1328; young people without SEN n = 6425), (4) whether English is a first or additional language (English as first language n = 1979; English as additional language n = 5843), and (5) age (10-12 years n= 4269; 13-15 years n= 3273) in order to assess whether this measure could also be used with young people beyond the originally specified age range (10-12 years).

For each of the five characteristics listed above, subgroups were tested for invariance first at the configural level then metric, scalar and residual levels if possible.

Robust maximum likelihood (MLR) estimation was used for all analyses because of non-normality in the CORS items (see Table 1); accordingly, nested model comparisons were conducted using a combination of the Satorra-Bentler scaled (mean-adjusted) chi-square test (Satorra and Bentler, 2010), changes in Bayesian information criterion (BIC) and assessment of fluctuation in model fit indices (CFI, RMSEA, and SRMR values). The reason for the combination of these three criteria to assess model fit is because the chi-square statistic is widely known to be unduly influenced by sample size- (Cheung and Rensvold, 2002; Kline, 2005; Meade et al., 2008). There is little consensus as to what constitutes a meaningful change when comparing fit indices of nested models. That said, when comparing nested models we operated by weighing up a combination of change criteria to decide on invariance including: a non-significant changes in chi-square value (bearing in mind the undue influence of sample size here), a change in CFI of -.01 (Cheung and Rensvold, 2002; Chen, 2007), a change in RMSEA of .015 (Chen, 2007), a change in SRMR of .030 (metric invariance) and .015 (scalar invariance; Chen, 2007) and a reduction in the BIC value (a lower value indicates a better trade-off between fit and complexity; there is no rule of thumb in change value because the values depend on actual dataset and the model).

To establish partial invariance, we first investigated the source of non-invariance by sequentially releasing (in a backward approach) item intercept constraints one by one and then retested the model against the metric invariance model until the change in model fit indices were acceptable (Putnick and Bornstein, 2016).

The sample size for the differential item functioning and measurement invariance varied according to the characteristic in question, as the completeness of the young people's characteristic data varied very slightly (sample size ranged from 7542 to 7609).

Sensitivity

To examine the ability of the CORS to distinguish between the community and clinical sample, Receiver Operating Characteristic (ROC) curve analysis was carried out on the data (from both the community and clinic samples). ROC curves test how well a measure can discriminate true positives and negatives, or between individuals with and without a certain health condition and those who don't. ROC analysis is also used to establish where the threshold between to two states should best lie, to balance sensitivity and specificity. Youden's index was used to establish the most appropriate threshold between clinical and non-clinical psychosocial functioning. ROC analysis was carried out using Stata 15, on data collected from the community and clinical samples described above.

Results

Factor structure

EFA carried out on half of the sample (n= 3911) strongly suggested a one factor solution, in fact there was only one factor with an eigenvalue greater than 1 (eigenvalue= 2.60). Goodness of fit indices suggest that the one factor solution fits the data well (CFI= .99; TLI= .97; RMSEA= .08; SRMR= .02) based on widely agreed criteria (Hu and Bentler, 1999). An obliquely rotated solution of the factor produced a solution presented in Table 1, all factor loadings were greater than .60 and were significant at the $p<.001$ level.

Subsequent CFA on the second half of the sample (n= 3911) accounting for clustering by schools confirmed that a one factor solution fits the data well (CFI= .98; TLI= .96; RMSEA= .07; SRMR= .02). See Table 2 for factor loadings for both the EFA and CFA models.

Internal consistency

The internal consistency for overall CORS scale was good; Cronbach's α was 0.81 and McDonald's ω was 0.82 (n = 7822).

Construct validity

Correlation analyses indicated a pattern whereby psychosocial functioning correlated negatively with indicators of mental health problems, and positively with young people's levels of resilience as measured using the various individual subscales of the SRS (i.e. low presence of external risk factors and high presence of protective factors). Moderate negative correlation was found particularly with emotional problems (r= .54), and positive correlation with self-esteem (r= .49). See Table 4.

Measurement invariance

As summarised in Table 5., full or partial scalar invariance was established between subgroups for all five respondent characteristics tested here (gender, eligibility for free school meals, presence of SEN, and whether English is a first or additional language). Of these, full scalar invariance was established between respondents eligible and not eligible for FSM, and between those in the older and younger age groups (which in fact reached residual, or strict invariance). Analyses indicated that for gender and presence of SEN, the item preventing the model from reaching full scalar invariance was CORS item three, “How am I doing at school?”. Female respondents and respondents without SEN were more likely to endorse this item positively. For DIF according to whether English is an additional or first language, the item preventing full scalar invariance was item 2 “How are things in my family?”. Young people whose first language was not English were more likely to endorse this item positively. Once these item intercepts were allowed to be freely estimated in each subgroup, model fit indices indicated a good model fit. Fit indices for all final invariance models are presented in Table 6, path diagrams with final model parameters are shown in Figures 1-5, and comparisons of sequential nested models are presented in Tables 7-11.

Chi-square difference tests indicated that there was a significant decrease in model fit for DIF according to certain subgroups between the metric and scalar (or partial scalar) models. However, chi-square has a high tendency to always be significant, by a more conservative estimate such as BIC or other goodness of fit indices (e.g. SRMR or CFI), all comparisons indicate that the more constrained model fits the data well.

Structural invariance

Once (partial) measurement invariance had been established, tests for structural invariance between subgroups in each of the respondent characteristics were carried out. SEM was used to determine whether latent factor means and variance were equivalent across subgroups.

Young people eligible and ineligible for FSM, in the older and younger age groups, and with and without SEN had equivalent amounts of individual differences (variance) in psychosocial functioning, see Tables 7-11 for details of fit indices. Analysis according to respondent gender and whether English is a first or second language on the other hand, indicated non-equivalence in variance between subgroups.

We then examined subgroup differences in the latent construct of psychosocial functioning for FSM, age group and SEN. After doing so, model fit indices indicated that levels of psychosocial functioning differ between young people eligible and ineligible for FSM, but not meaningfully so between those with and without SEN nor between those in the older and younger age groups.

Sensitivity

The ROC curve analysis indicated that the CORS is reasonably sensitive in detecting the presence or absence of mental health problems as defined as being in the clinic sample or community sample (with the caveat that there will be individuals in the community sample with substantial mental health difficulties). The area under the curve (AUC) statistic is .78 (SE= .01, CI(95%)= .76-.79). The most appropriate threshold between clinical and non-clinical psychosocial functioning was found to be a score of 28 (sensitivity = .73, specificity= .70). On average, young people in the community sample (M= 31.38, SD= 7.69) reported significantly higher total CORS scores (t= 48.28, $p<.001$) than those in the clinic sample (M= 22.67, SD= 8.62).

Figures 6 and 7 show the ROC curve and distribution of total CORS scores for each sample respectively.

Discussion

The current paper builds on preliminary validation work from US on the Child Outcome Rating Scale (CORS), a child self-report measure of psychosocial functioning now widely used internationally but without an extensive base of psychometric analysis. We built on the initial validation work by 1) using a considerably larger sample of children and adolescents; 2) including analysis of concurrent validity by examining the relationship between responses to the CORS and responses on measures of other domains related to mental health; 3) including a more sophisticated analysis to establish a clinical threshold, and; 4) exploring the possible presence of bias in items on the CORS according to different respondent subgroups.

Based on the current analysis it would seem that the CORS is a robust, brief measure of psychosocial functioning, suitable for completion by young people. Factor structure and high internal consistency confirm that the four items that comprise the CORS are reliably tapping into the same construct. The CORS correlates moderately with other domains related to young people's mental health and wellbeing, particularly their ratings of emotional difficulties, health-related quality of life and self-esteem. These associations reflect patterns identified in previous research, whereby functioning and wellbeing outcomes can be influenced by mental health (for example) but need not be dictated by it (Patalay and Fitzsimons, 2018).

Differential item functioning (DIF) was carried out to determine whether the underlying construct assessed by the CORS is conceptually similar when measured in different subgroups of young people (gender, eligibility for free school meals, presence of SEN, whether English is a first or additional language and age group). If not (if we detected bias in any of the items) then comparisons between different respondent subgroups (male and female respondents, for example) would not be meaningful. Broadly speaking the DIF analysis indicated that there is little bias in the CORS items that would make group comparison difficult. However, best practice would recommend that one should account for the slight invariance between subgroups using appropriate statistical methods (e.g. using factor scores rather than mean scores) if comparing CORS scores from different subgroups. That the overarching concept of psychosocial functioning as measured by these items appears to be understood in the same way according to younger and older respondents is of particular interest, and indicates that the CORS is acceptable for use with young people up to the age of 15 years (beyond the initial upper recommendation of 12 years). This makes it a valuable tool for longitudinal measurement of psychosocial functioning, and limits the need to change measures between different age groups in a wider age range, making possible easily comparable research and outcome monitoring across these ages.

Based on the current analysis, scores of 28 and under are likely to indicate levels of psychosocial functioning typically found in clinical, help-seeking populations. This is slightly lower than the cut-off score of 32 recommended for the CORS in the original preliminary validation study by Duncan and colleagues (2006), but in line with that recommended for young people aged 13-17 completing the ORS (Miller et al., 2003). This is perhaps not surprising given that average age is 12.8 years in the community sample and 13.0 years in the clinic sample (ranging from 10-15 years), and so is more closely aligned with the age range of the ORS than the CORS (6-12 years). Based on this we would endorse the threshold of 28 originally recommended for young people aged 13-17 years, and tentatively suggest that clinicians continue to adhere to the threshold of 32 for young people aged 6-12 until more data is available to confirm this too. We also advise that thresholds are established in populations with specific diagnoses, so that expected levels of psychosocial functioning more accurately reflect optimal outcomes.

This analysis comes at an important time when prevalence rates of mental health problems in young people are rising globally (Husky et al., 2018; Patel, Flisher, Hetrick, and McGorry, 2007), and there is increasing emphasis on measuring outcomes beyond psychopathology such as psychosocial functioning (Bradley et al., 2013; Leamy, Bird, Le Boutillier, Williams and Slade, 2011). This is especially so with regard to young people where mental health problems and associated symptoms tend to be enduring (Costello, Foley, and Angold, 2006). Most importantly, there is recognition that young people’s perspectives on their own mental health are to be valued and steps taken to ensure they are captured and taken into account (Deighton et al., 2014). Government guidelines from the UK and US (among others) firmly advocate that service outcomes should be sought from the perspective of the young service users themselves (UK Department for Health, 2004).

The CORS itself is already in widespread use around the world in clinical settings, and further psychometric analysis is overdue to provide professionals with the reassurance that the measure is valid and reliable. In the UK for example, the government has already invested money to make the CORS freely available to those working in the field. This makes this analysis extremely timely.

Besides psychometric robustness, that the CORS is just four items long should make this measure feasible and particularly attractive to busy clinical services. In judging the feasibility of a given measure, Slade et al. (1999) recommend assessment according to six criteria: brevity, simplicity, relevance (to young services users and clinical staff), acceptability (to young services users and clinical staff), availability, and value. It is our opinion that the CORS meets the criteria for feasibility as well as for psychometric acceptability.

Limitations to the current paper include the following: First, the CORS has been created for use with young people aged 6-to-12 years-old and the sample included here ranges in age from 10 to 15 years across both the community and clinical samples. In this paper we do not re-test the lower limit of the CORS age range, but seek to extend the upper age range. Second, while the current analysis can provide reassurance with regard to the robustness of the CORS as a static measure of psychosocial functioning, it does not allow any comment on the suitability of the CORS as a measure to track young people’s progress over time or in response to clinical intervention.

Further limitations of this analysis concern the assessment of the sensitivity of the CORS to distinguish between clinical and non-clinical levels of psychosocial functioning. The ‘non-clinical’ or community sample is imperfect in that it is likely that this sample, drawn from mainstream schools around the UK, also includes some clinical cases. It was not possible to obtain information about receipt of clinical intervention for this schools sample. Nonetheless, that this sensitivity analysis is in agreement with the original analysis for adolescents (a cut-off of 28; Miller et al., 2003) lends weight to this finding.

Despite these limitations, on the basis of our findings we feel that that the CORS has potential for use as a broad measure of psychosocial functioning, and as a rapid and free (in the UK) measure to consider progress and impact of services for young people seeking mental health support.

References

Achenbach. T.M., McConaughy, S.H. & Howell, C.T. (1987). Child/adolescent behavioural and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101, 213-232*.

Batty, M.J., Moldavsky, M., Foroushani, P.S., Pass, S., Marriott, M., Sayal, K. & Hollis, C. (2013). Implementing routine outcome measures in child and adolescent mental health services: from present to future practice. *Child and Adolescent Mental Health, 18(2), 82-87.*

Bradley, J., Murphy, S., Fugard, A.J.B., Nolas, S-M. and Law, D. (2013). What kinds of goals do children and young people set for themselves in therapy? Developing a goals framework using CORC data. *Clinical Child and Family Psychology Review,1(1), 8-18.*

Cabrera-Nguyen, P. (2010). Author Guidelines for Reporting Scale Development and Validation Results in the Journal of the Society for Social Work and Research. Journal of the Society for Social Work and Research *Journal of the Society for Social Work and Research, 1(2)*, 99-103.

Chen, F.F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modelling, 14(3)*, 464-504.

Cheung, G.W. & Rensvold, R.B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9(2), 233-255*.

Children and Young People's Improving Access to Psychological Therapies Programme. (2013). National curriculum for core, cognitive behavioural therapy, parenting training (3–10-year olds), systemic family practice, interpersonal psychotherapy for adolescents, supervision, and transformational service leadership. London: IAPT.

Childs J, Deighton J, Wolpert M. Defining and Measuring Mental Health and Wellbeing: A Response Mode Report Requested by the Department of Health for the Policy Research Unit in the Health of Children, Young People and Families. London; 2013. https://www.ucl.ac.uk/evidence-based-practice-unit/sites/evidence-based-practice-unit/files/pub_and_resources_project_reports_defining_and_measuring_mental_health.pdf.

Cleridou, K., Patalay, P. & Martin, P. (2017). Does parent–child agreement vary based on presenting problems? Results from a UK clinical sample. *Child and adolescent Psychiatry and Mental Health, 11, 22*.

Costello, E.J., Foley, D.L., & Angold, A. (2006). 10-year research update review: The epidemiology of child and adolescent psychiatric disorders: II. Developmental epidemiology. *Journal of the American Academy of Child and Adolescent Psychiatry, 45(1), 8–25*.

Day, C., Carey, M., & Surgenor, T. (2006). Children's key concerns: Piloting a qualitative approach to understanding their experience of mental health care. *Journal of Clinical Child Psychology and Psychiatry, 11 (1), 139-155.*

Deighton., J., Croudace, T., Fonagy, P., Brown, J., Patalya, P., & Wolpert, M. (2014). Measuring mental health and wellbeing outcomes for children and adolescents to inform practice and policy: A review of child self-report measures. *Child and Adolescent Psychiatry and Mental Health, 8, 14.*

Deighton, J., Lereya, T., Patalay, P., Casey, P., Humphrey, N. & Wolpert, W. (2018). *Mental health problems in young people, aged 11 to 14: Results from the first HeadStart annual survey of 30,000 children*. London: CAMHS Press.

Deighton, J., Tymms, P., Vostanis, P., Belsky, J., Fonagy, P., Brown, A., Martin, A., Patalay, P., Wolpert, M. (2013). The development of a school-based measure of child mental health. *Journal of Psychoeducational Assessment, 31, 247–57.*

Duncan, B.L., sparks, J.A., Miller, S.D., Bohanske, R.Y., & Claud, D.A. (2006). Giving Youth a Voice: A preliminary study of the reliability and validity of a brief outcome measure for children, adolescents, and caretakers. *Journal of Brief Therapy, 5(2), 71-88*.

Hooper, D., Coughlan, J. & Mullen, M.R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal on Business Research Methods, 6(1),* 53-60.

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6(1), 1-55*.

Husky, M.M., Boyd, A., Bitfoi, A., Carta, M.G., Chan-Chee, C., Goelitz, D., Koç, C., Lesinskiene, S., Mihova, Z., Otten, R., Pez, O., Shojaei, T., Kovess-Masfety, V. (2018). Self-reported mental health in children ages 6–12 years across eight European countries. *European Child & Adolescent Psychiatry, 27(6), 785-795*.

Joseph, S. & Wood, A. (2010). Assessment of positive functioning in clinical psychology: Theoretical and practical issues. *Clinical Psychology Review, 30, 830-838*.

Kline, R.B. (2005). *Principles and Practice of Structural Equation Modeling (2nd Edition)*. New York: Guilford Press

Krause, K. R., Bear, H. A., Edbrooke-Childs, J. and Wolpert, M. (2018). What Outcomes Count? A Review of Outcomes Measured for Adolescent Depression between 2007 and 2017. *Journal of the American Academy of Child and Adolescent Psychiatry*.

van Laerhoven, H., van der Zaag-Loonen, H.J. & Derkx, B.H.F. (2004). A comparison of Likert scale and visual analogue scales as response options in children's questionnaires. *Acta Paediatrica, 93, 830-835*.

Leamy, M., Bird, V., Le Boutillier, C., Williams, J., & Slade, M. (2011). Conceptual framework for personal recovery in mental health: systematic review and narrative synthesis. *British Journal of Psychiatry, 199(6), 445-452*.

Lereya, S.T., Humphrey, N., Patalay, P., Wolpert, M., Böhnke, J.R., Macdougall, A. & Deighton, J. (2016). The student resilience survey: Psychometric validation and associations with mental health. *Child and Adolescent Psychiatry and Mental Health, 10:44*.

Meade, A.W., Johnson, E.C. & Braddy, P.W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93(3), 568-592*.

Miller, S. Duncan, B.L., Brown, J., Sparks, J.A., Claud, D.A. (2003). The Outcome Rating Scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *Journal of Brief Therapy, 2(2), 91-100.*

Muthén LK, Muthén BO. (1998-2015). *Mplus user's guide*. Seventh Edition. Los Angeles, CA: Muthén & Muthén.

Nunnally, J. C. (1978). *Psychometric theory*. New York, NY: McGrawHill

Pantell, R.H. & Lewis, C.C. (1987). Measuring the impact of medical care on children. *Journal of Chronic Diseases, 40(1), 99–108*.

Patalay, P., Deighton, J., Fonagy, P., Vostanis, P., & Wolpert, M. (2014). Clinical validity of the Me and My School questionnaire: A. self-report mental health measure for children and adolescents. *Child and Adolescent Psychiatry and Mental Health*, 8:17.

Patalay, P., & Fitzsimons, E. (2018). Development and predictors of mental ill-health and wellbeing from childhood to adolescence. *Social psychiatry and psychiatric epidemiology, 53(12)*, 1311-1323..

Patalay, P., Fonagy, P., Deighton, J., Belsky, J., Vostanis, P., & Wolpert, M. (2015). A general psychopathology factor in early adolescence. *The British Journal of Psychiatry, 207, 15-22*.

Patel, V., Flisher, A.J., Hetrick, S. & McGorry, P. (2007). Mental health of young people: A global public-health challenge. *Lancet, 369, 1302-1313*.

Putnick, D.L. & Bornstein, M.H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review, 41*, 71-90.

Ravens-Sieberer, U., Wille, N., Badia, X., Bonsel, G., Burström, K., Cavrini, G., Devlin, N., Egmar, A-C, Gusi, N., Herdman, M., et al. (2010). Feasibility, reliability, and validity of the EQ-5D-Y: results from a multinational study. *Quality of Life Research, 19(6), 887–97.*

Satorra, A. & Bentler, P.M. (2010) Ensuring positiveness of the scaled difference chi square test statistic. *Psychometrika, 75, 243-248*.

Sharpe, H., Patalya, P., Fink, E., Vostanis, P., Deighton, J. and Wolpert, M. (2016). Exploring the relationship between quality of life and mental health problems in children: Implications for measurement and practice. *European Child and Adolescent Psychiatry, 25, 659-667*.

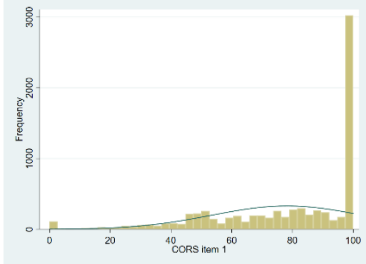
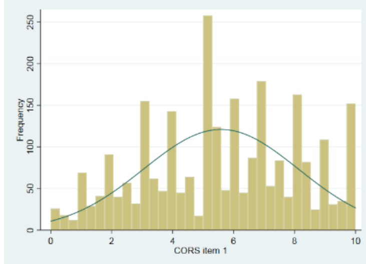
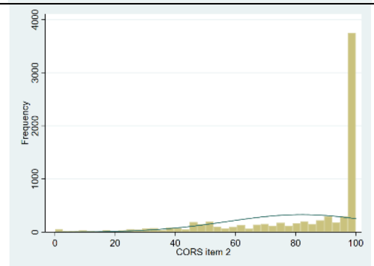
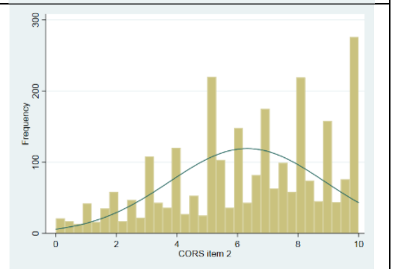
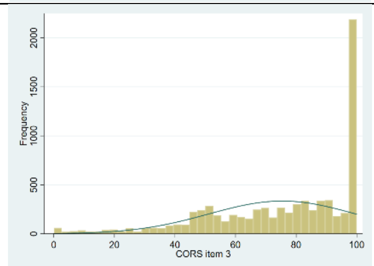
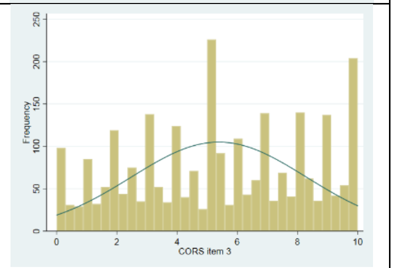
Slade, M., Thornicroft, G. & Glover, G. (1999). The feasibility of routine outcome measures in mental health. *Social Psychiatry and Psychiatric Epidemiology, 34, 243-249*.

Sourander, A., Helstela, L. & Helenius, H. (1999). Parent-adolescent agreement on emotional and behavioral problems. *Social Psychiatry Psychiatric Epidemiology, 34, 657-663*.

Sun, J. & Stewart, D. (2007). Development of population-based resilience measures in the primary school setting. *Health Education, 7(6), 575–99.*

UK Government (2004). Children Act, Volume 31. London: The Stationery Office.

Table 1. Completeness and score distribution of individual CORS items and total CORS in community and clinic sample

	Complete responses (%)		Measures of distribution							
			Community sample				Clinic sample			
Item	Community sample	Clinic sample	Mean (SD)	Median	Skewness	Histogram	Mean (SD)	Median	Skewness	Histogram
Item 1: How am I doing?	7783 (99.35)	2621 (100.00%)	78.23 (24.67)	86.00	-1.06		5.59 (2.54)	5.50	-.10	
Item 2: How are things in my family?	7751 (98.94)	2619 (99.92%)	82.01 (24.68)	96.00	-1.35		6.23 (2.58)	6.60	-.39	
Item 3: How am I doing at school?	7719 (98.53)	2606 (99.43%)	75.41	81.00	-.90		5.39 (2.91)	5.20	-.08	

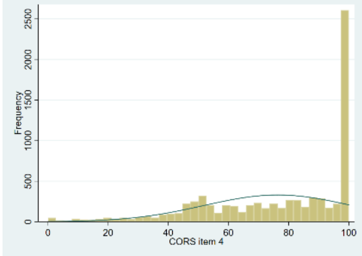
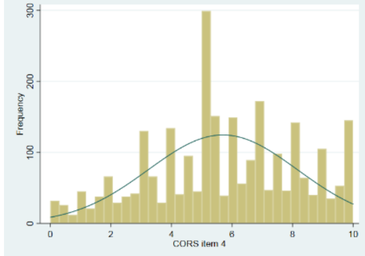
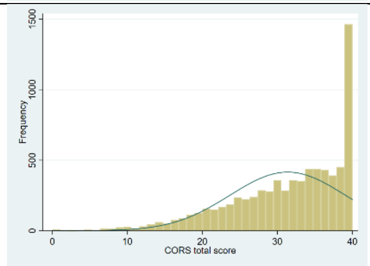
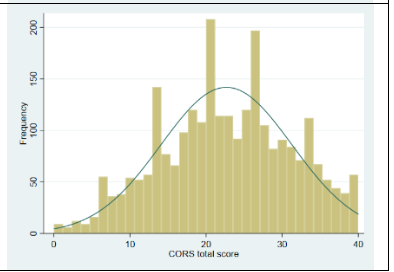
Item 4: How is everything going?	7735 (98.74)	2619 (99.92%)	76.40 (24.59)	83.00	-.87		5.70 (2.47)	5.50	-.15	
Total CORS	7609 (97.13)	2604 (99.35%)	31.37 (7.69)	33.10	-.93		22.67 (8.82)	23.00	-.09	

Table 2. EFA rotated loadings and CFA standardised loadings

Item	EFA factor loadings	CFA factor loadings
Item 1: How am I doing?	.76	.76
Item 2: How are things in my family?	.62	.58
Item 3: How am I doing at school?	.64	.65
Item 4: How is everything going?	.91	.91

Table 3. Correlations between CORS items

Item	Item 1	Item 2	Item 3	Item 4
Item 1: How am I doing?				
Item 2: How are things in my family?	.43 (n= 7716)			
Item 3: How am I doing at school?	.52 (n=7695)	.34 (n= 7663)		
Item 4: How is everything going?	.68 (n= 7709)	.55 (n= 7689)	.57 (n= 7667)	

All correlations are significant at $p < 0.001$ level

Table 4. Pearson correlations between total CORS scores and other scales

	CORS (psychosocial functioning) *(n)*
Emotional problems (MAMF)	-.54* *(n= 7087)*
Behavioural problems (MAMF)	-.38* *(n= 7145)*
Health-related quality of life (EQ 5D-Y)	-.48* *(n= 6901)*
Health today (EQ 5D-Y)	.43* *(n= 7179)*
Family connection (SRS)	.41* *(n= 7160)*
School connection (SRS)	.39* *(n= 7134)*
Community connection (SRS)	.36* *(n= 7082)*
Participation in home and school life (SRS)	.40* *(n= 7089)*
Participation in community life (SRS)	.16* *(n= 7101)*
Self esteem (SRS)	.49* *(n= 7157)*
Empathy (SRS)	.21* *(n= 7189)*
Problem solving (SRS)	.44* *(n= 7112)*
Goals and aspirations (SRS)	.37* *(n= 7122)*
Peer support (SRS)	.35* *(n= 6855)*

* $p < 0.0001$

Table 5. Summary of levels of invariance established between subgroups

	Level of invariance				
	Configural	Metric	Partial scalar	Scalar	Residual
Gender	✓	✓	✓	-	-
SEN	✓	✓	✓	-	-
FSM	✓	✓	✓	✓	-
Age group	✓	✓	✓	✓	✓
EAL	✓	✓	✓	-	-

SEN= Special Educational Needs, FSM= Free School Meals, EAL= English as additional language

Table 6. Model fit indices for final invariance models

	χ^2 (df)	CFI	TLI	RMSEA (90% CI)	SRMR
Gender	86.20 (9)	.99	.98	.05 (.04-.06)	.02
SEN	82.14 (11)	.99	.99	.04 (.03-.05)	.02
FSM	104.61 (12)	.99	.99	.05 (.04-.05)	.05
Age group	62.89 (7)	.99	.98	.05 (.04-.06)	.02
EAL	81.07 (9)	.99	.98	.05 (.04-.06)	.03

SEN= Special Educational Needs, FSM= Free School Meals, EAL= English as Additional Language

Table 7: Comparisons of sequential nested models, for DIF analysis according to respondent gender

Model	χ^2 (df)	CFI	TLI	RMSEA (90% CI)	SRMR	BIC	Model comparison	$^{\Delta}\chi^2$ (df)	$^{\Delta}$CFI	$^{\Delta}$RMSEA	$^{\Delta}$SRMR	$^{\Delta}$BIC
M1: Configural Invariance	64.85 (4)	.99	.97	.06 (.05-.08)	.02	275252	-	-	-	-	-	-
M2: Metric Invariance	74.92 (7)	.99	.98	.05 (.04-.06)	.02	275234.9	M1	10.07 (3)	.00	-.01	.00	-17.13
M3: Scalar Invariance	137.71 (10)	.97	.97	.06 (.05-.07)	.04	275282.8	M2	62.69 (3)***	-.02	.01	.02	47.98
M3a: Partial scalar Invariance (item 3 intercept freely estimated)	86.20 (9)	.99	.98	.05 (.04-.06)	.02	275223.9	M2	11.28 (2)*	.00	.00	.00	-10.93
M: Residual Invariance	-	-	-	-	-	-	-	-	-	-	-	-
M4: Structural invariance (factor variance)	99.27 (10)	.99	.98	.05 (.04-.06)	.05	275232.5	M3a	13.07 (1)***	.00	.00	.03	8.53
M5: Structural invariance (factor variance and means)	153.55 (11)	.98	.97	.06 (.05-.07)	.09	275295.5	M4	54.28 (1)***	-.01	.01	.04	63.03

Note. N = 7822; male n = 3326; female n = 4496.

* p<.05 *** p<.0001

Table 8: Comparisons of sequential nested models for DIF analysis according to presence of SEN

Model	χ^2 (df)	CFI	TLI	RMSEA (90% CI)	SRMR	BIC	Model comparison	$^{\Delta}\chi^2$ (df)	$^{\Delta}$CFI	$^{\Delta}$RMSEA	$^{\Delta}$SRMR	$^{\Delta}$BIC
M1: Configural Invariance	61.78 (4)	.99	.97	.06 (.05-.08)	.02	272659.4	-	-	-	-	-	-
M2: Metric Invariance	71.16 (7)	.99	.98	.05 (.04-.06)	.02	272640.4	M1	9.38 (3)	.00	-.01	.00	-18.959
M3: Scalar Invariance	100.39 (10)	.99	.98	.05 (.04-.06)	.03	272641.9	M2	29.23 (3)***	.00	.00	.01	1.503
M3a: Partial scalar Invariance (item 3 intercept freely estimated)	78.19 (9)	.99	.99	.05 (.04-.05)	.02	272624.4	M2	7.03 (2)	.00	.00	.00	-16.01
M: Residual Invariance	-	-	-	-	-	-	-	-	-	-	-	-
M4: Structural invariance (factor variance)	80.27 (10)	.99	.99	.04 (.03-.05)	.02	272616.5	M3a	2.08 (1)	.00	-.01	-.01	-7.866
M5: Structural invariance (factor variance and means)	82.14 (11)	.99	.99	.04 (.03-.05)	.02	272607.6	M4	1.87 (1)	.00	.00	.00	-8.91

Table 9: Comparisons of sequential nested models for DIF analysis according to eligibility for FSM

Model	χ^2 (df)	CFI	TLI	RMSEA (90% CI)	SRMR	BIC	Model comparison	$\Delta\chi^2$ (df)	ΔCFI	ΔRMSEA	ΔSRMR	ΔBIC
M1: Configural Invariance	62.74 (4)	.99	.97	.06 (.05-.08)	.02	272824.4	-	-	-	-	-	-
M2: Metric Invariance	73.55 (7)	.99	.98	.05 (.04-.06)	.02	272807.2	M1	10.81 (3)	.00	-.01	.00	-17.26
M3: Scalar Invariance	88.18 (10)	.99	.99	.05 (.04-.05)	.03	272788.3	M2	14.63 (3)*	.00	.00	.01	-18.86
M4: Residual Invariance	174.76 (14)	.97	.98	.05 (.05-.06)	.05	272880.7	M3	86.58 (4) ***	-.02	.00	.02	92.33
M5: Structural invariance (factor variance)	94.39 (11)	.99	.99	.04 (.04-.05)	.04	272785.2	M4	6.21 (1)*	.02	-.01	-.01	-3.11
M6: Structural invariance (factor variance and means)	104.61 (12)	.99	.99	.05 (.04-.05)	.05	272786.9	M5	10.22 (1)**	.00	.01	.01	1.72

Table 10: Comparisons of sequential nested models for DIF analysis according to age group

Model	χ^2 (df)	CFI	TLI	RMSEA (90% CI)	SRMR	BIC	Model comparison	$\Delta\chi^2$ (df)	ΔCFI	ΔRMSEA	ΔSRMR	ΔBIC
M1: Configural Invariance	7.35 (4)	.99	.97	.02 (.00-.03)	.02	376107.83						
M2: Metric Invariance	11.41 (7)	.99	.98	.01 (.00-.03)	.04	376160.77	M1	-4.06 (3)	.00	.01	-.02	-52.94
M3: Scalar Invariance	16.37 (10)	.98	.98	.01 (.00-.02)	.04	376139.77	M2	-4.96 (3)	.01	.00	.00	21
M4: Residual Invariance	19.07 (14)	.98	.99	.01 (.00-.02)	.06	376280.69	M3	-2.7 (4)	.00	.00	-.02	-140.92
M5: Structural invariance (factor variance)	21.48 (15)	.98	.98	.01 (.00-.02)	.09	376349.50	M4	-2.41 (1)	.00	.00	-.03	-68.81
M6: Structural invariance (factor variance and means)	23.85 (16)	.98	.98	.01 (.00-.02)	.09	376360.20	M5	-2.37 (1)	.00	.00	.00	-10.7

Table 11: Comparisons of sequential nested models for DIF analysis according to whether English is a first language or not

Model	χ^2 (df)	CFI	TLI	RMSEA (90% CI)	SRMR	BIC	Model comparison	$^{\Delta}\chi^2$ (df)	$^{\Delta}$CFI	$^{\Delta}$RMSEA	$^{\Delta}$SRMR	$^{\Delta}$BIC
M1: Configural Invariance	59.85 (4)	.99	.97	.06 (.05-.07)	.02	271608	-	-	-	-	-	-
M2: Metric Invariance	69.22 (7)	.99	.98	.05 (.04-.06)	.03	271591.4	M1	9.37	.00	-.01	.01	-16.619
M3: Scalar Invariance	96.63 (10)	.99	.98	.05 (.04-.06)	.04	271592.8	M2	27.41***	.00	.00	.01	1.43
M3a: Partial scalar Invariance (item 2 intercept freely estimated)	81.07 (9)	.99	.98	.05 (.04-.06)	.03	271581.9	M2	11.85*	.00	.00	.00	-9.497
M: Residual Invariance	-	-	-	-	-	-	-	-	-	-	-	-
M4: Structural invariance (factor variance)	92.57 (10)	.99	.98	.05 (.04-.06)	.06	271591.5	M3a	11.50**	.00	.00	.03	9.578
M5: Structural invariance (factor variance and means)	138.05 (11)	.98	.98	.06 (.05-.06)	.09	271645.5	M4	45.48***	-.01	.01	.03	54.029

Figure 1. Path diagram for final measurement invariance model according to gender (unstandardized)

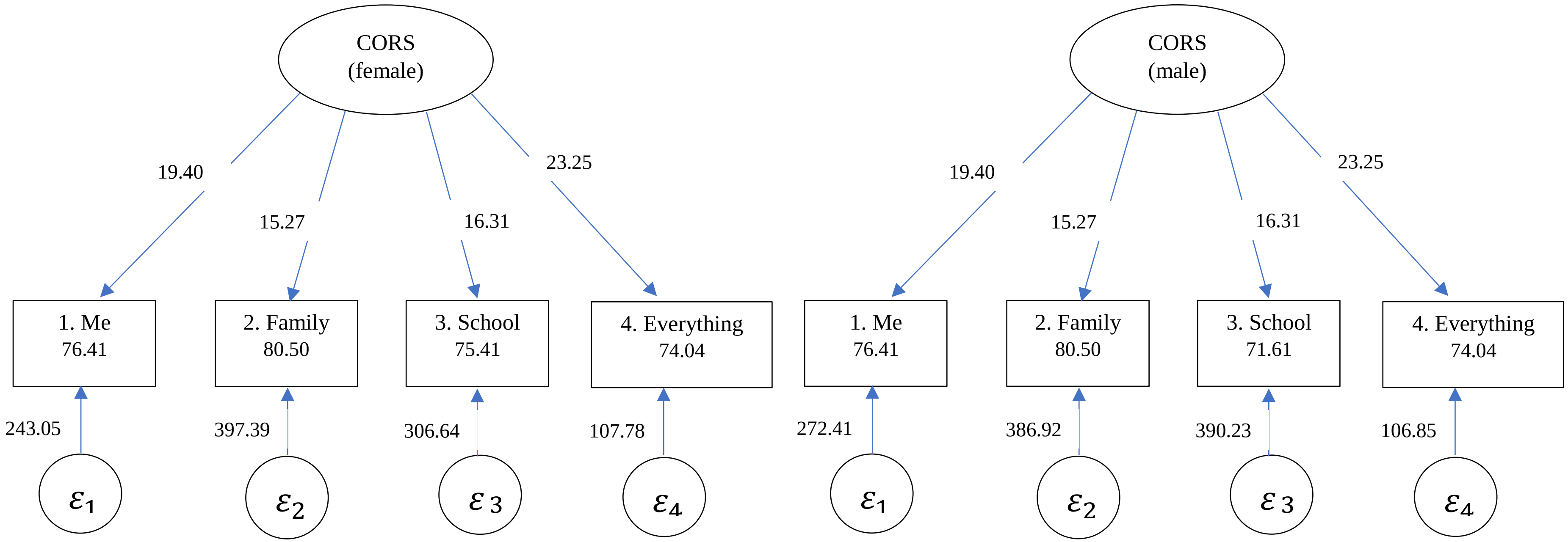


Figure 2. Path diagram for final measurement invariance model according to presence/absence of SEN (unstandardized)

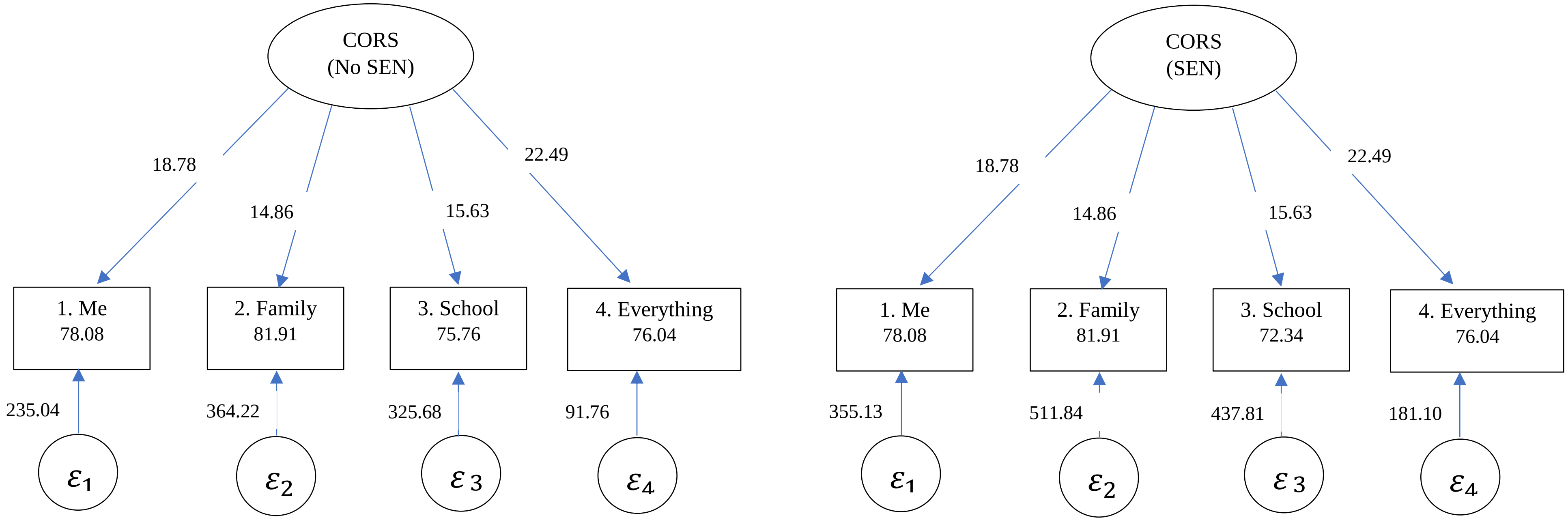


Figure 3. Path diagram for final measurement invariance model according to eligibility for FSM (unstandardized)

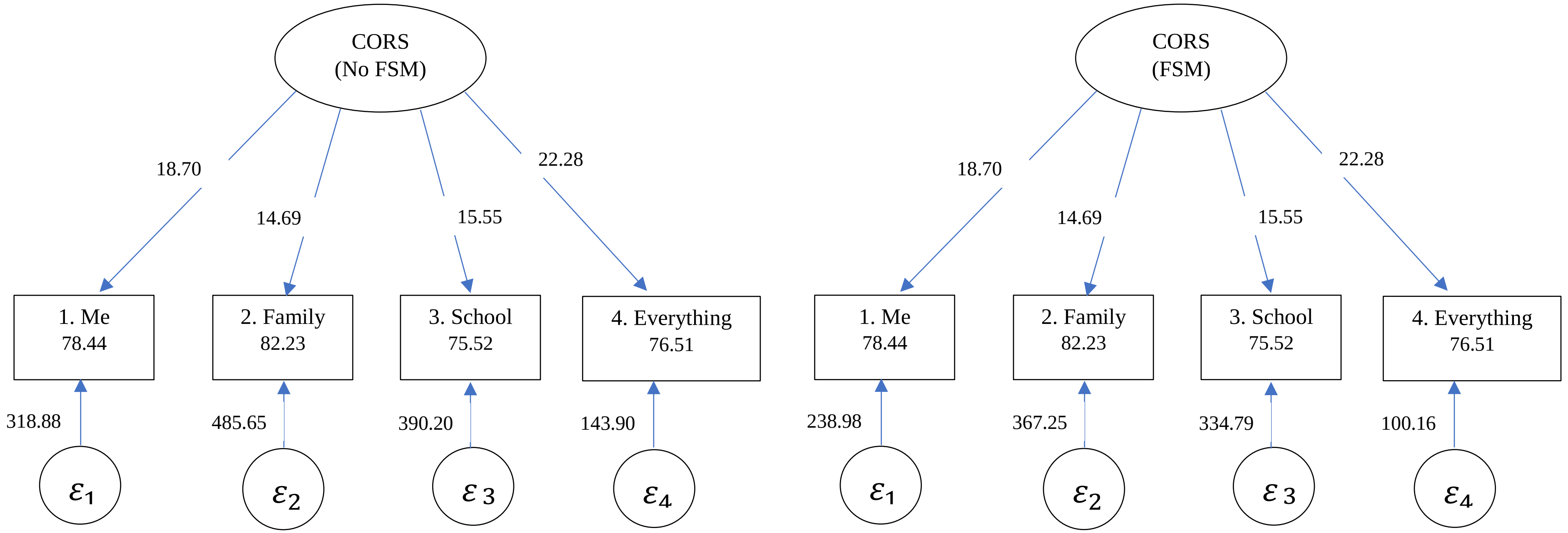


Figure 4. Path diagram for final measurement invariance model according to Age Group (unstandardized)

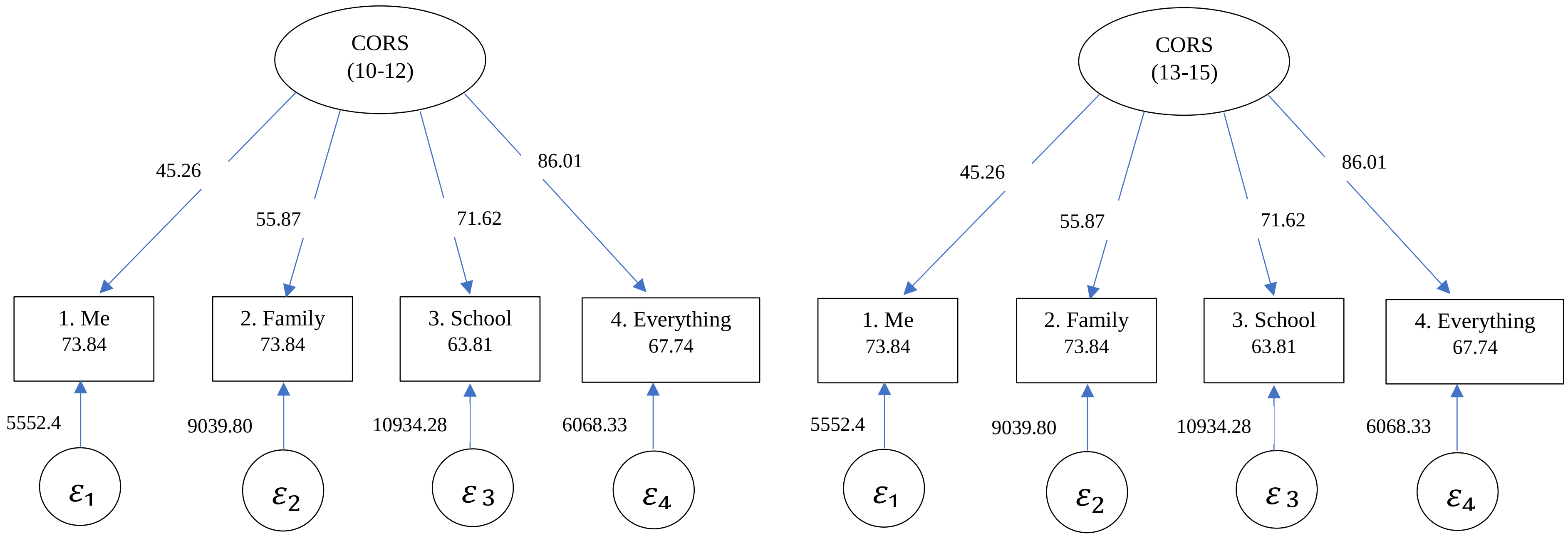


Figure 5. Path diagram for final measurement invariance model according to whether English is a first or additional language (unstandardized)

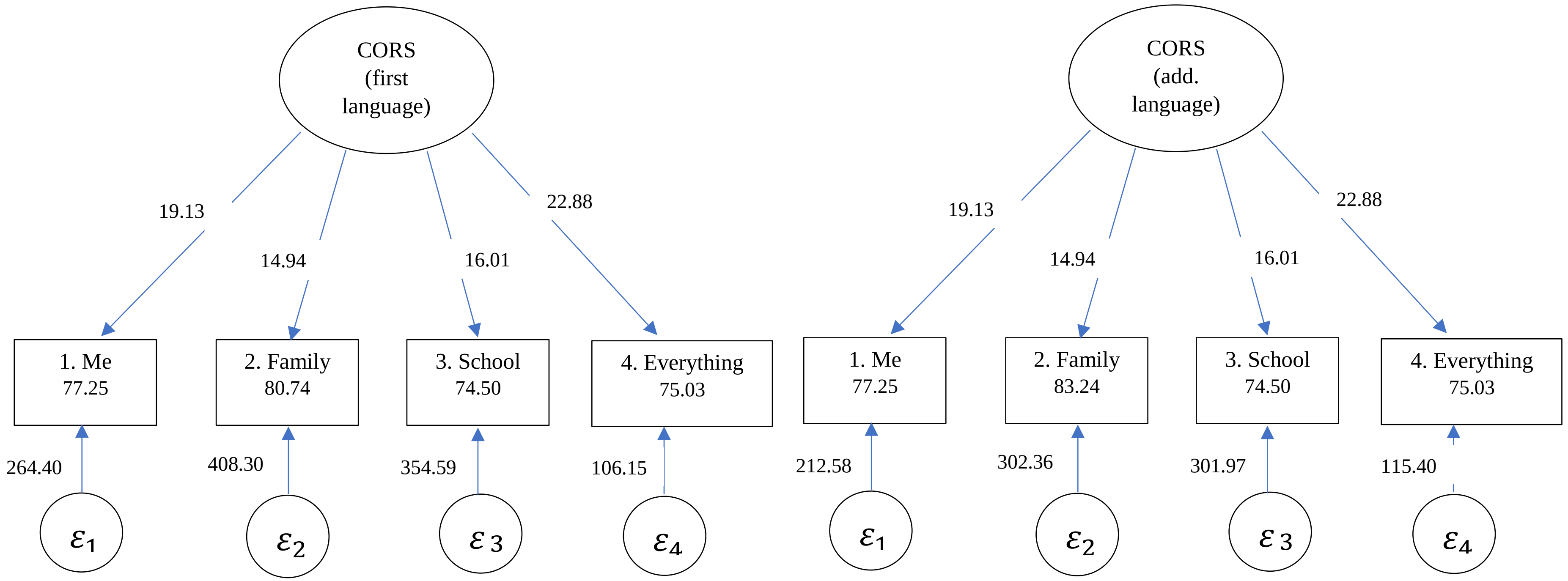


Figure 6. ROC curve for CORS

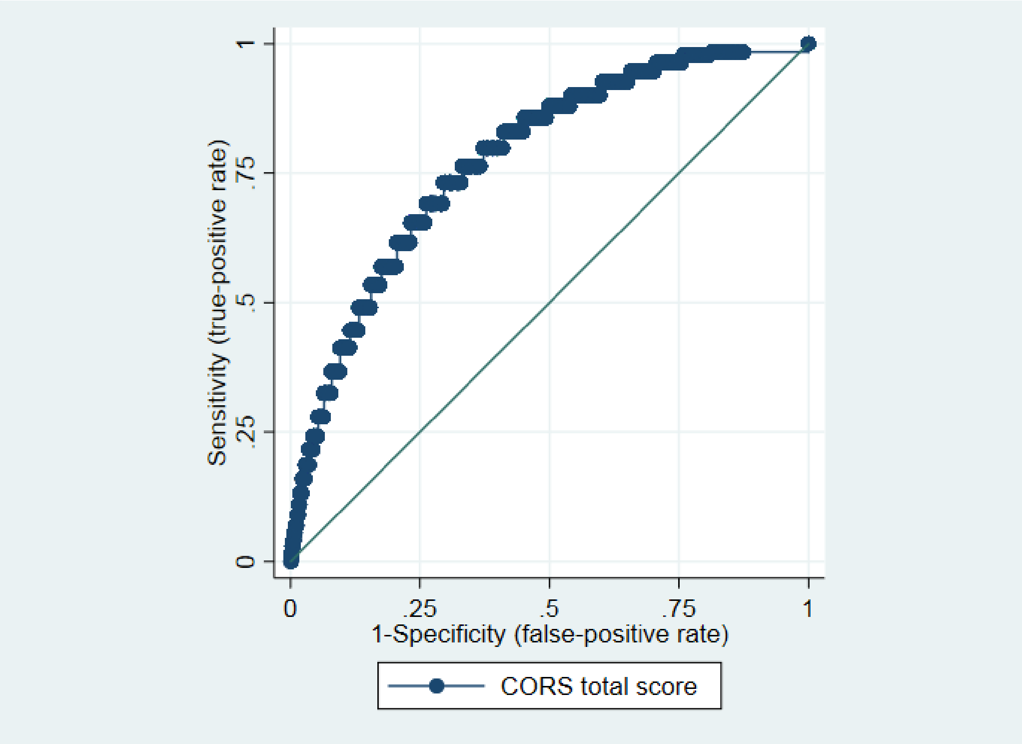


Figure 7. Distribution of CORS scores for community and clinic sample, with vertical line representing the estimated cut-off score of 28.