Assessing for Suicidal Behavior in Youth Using the Achenbach System of Empirically Based Assessment

Anna R. Van Meter [1], Guillermo Perez Algorta[2], Eric A. Youngstrom[3]

Yana Lechtman[1], Jen K. Youngstrom[3], Norah C. Feeny[4], Robert L. Findling[5]

[1]Ferkauf Graduate School, Yeshiva University, Bronx, NY

[2] Lancaster University, Lancaster, UK

[3]University of North Carolina, Chapel Hill, NC

[4] Case Western Reserve University, Cleveland, OH

[5]Johns Hopkins University/ Kennedy Krieger Institute, Baltimore, MD

Corresponding author:

Anna Van Meter, Ph.D.

1165 Morris Park Avenue, Rousso Building, Bronx, NY 10461

anna.vanmeter@einstein.yu.edu

718.430.3861

Abstract

Objective: This study investigated the clinical utility of the Achenbach System of Empirically Based Assessment (ASEBA) for identifying youth at risk for suicide. Specifically, we investigated how well the Total Problems scores and the sum of two suicide-related items (#18 "Deliberately harms self or attempts suicide" and #91 "Talks about killing self") were able to distinguish youth with a history of suicidal behavior.

Method: Youth (*N*=1117) aged 5 to 18 were recruited for two studies of mental illness. History of suicidal behavior was assessed by semi-structured interviews (KSADS) with youth and caregivers. Youth, caregivers, and a primary teacher each completed the appropriate form (YSR, CBCL, TRF, respectively) of the ASEBA. Areas under the curve (AUCs) from ROC analyses and diagnostic likelihood ratios (DLRs) were used to measure the ability of both Total Problems *T*-scores, as well as the summed score of two suicide-related items, to identify youth with a history of suicidal behavior.

Results: The Suicide Items from the CBCL and YSR performed well (AUCs=.85 and .70, respectively). The TRF Suicide Items did not perform better than chance, AUC=.45. The AUCs for the Total Problems scores were poor-to-fair (.33-.65). The CBCL Suicide Items outperformed all other scores (*p*s=.04 to <.0005). Combining the CBCL and YSR items did not lead to incremental improvement in prediction over the CBCL alone.

Conclusion: The sum of two questions from a commonly used assessment tool can offer important information about a youth's risk for suicidal behavior The low burden of this approach could facilitate wide-spread screening for suicide in an increasingly at-risk population.

Keywords: suicide, diagnostic accuracy, ROC, evidence based assessment

Assessing for Suicidal Behavior in Youth Using the Achenbach System of Empirically Based Assessment

Suicide, especially among young people, occurs much more frequently than we would like to believe [Centers for Disease 1]. The most recent data from the Centers for Disease Control indicate a disquieting increase in the number of suicides, particularly among girls aged 10-14 [2]. A nationwide survey of youth [3] grades 9–12 found that 16% of students reported seriously considering suicide, 13% reported having a suicidal plan, and 8% reported trying to take their own life in the 12 months preceding the survey. For every death by suicide, there may be as many as 25 suicide attempts [4] – a figure that is likely higher among youth, who tend to complete fewer suicides than adults [5,6,3]. This statistic is supported by the fact that each year approximately 157,000 young people (ages 10-24) receive medical care for self-inflicted injuries at emergency departments across the U.S. [1]. Rates appear similar or even higher throughout Europe and Asia [7].

Suicidal thoughts and behaviors affect a concerning number of young people, and we lose nearly 5,000 young people between the ages of 10 and 24 per year to suicide [Centers for Disease 1], but our methods for identifying those at risk are limited. In spite of great efforts to prevent suicide, rates have increased in nearly every demographic group in the past 15 years [2]. A completed suicide has consequences that extend far beyond the family of the person lost; schools must manage fear and grief in the community, other parents must face difficult questions about their own children, and peers are at higher risk due to suicide clustering and contagion [8,9]. Identified risk factors for suicide among youth include psychiatric illness, family history of suicidality, sexual and physical abuse, sexual orientation, bullying, and previous suicidal attempt or behavior [10]. However, a recent meta-analysis raises questions about the predictive validity of suicidal ideation and past suicide behaviors [11]; the presence of a single, specific risk factor is not enough to forecast with certainty who will engage in suicidal behavior. In order to make meaningful predictions about who is at risk, and to quantify that risk, it is important to be able to

combine multiple risk factors, each of which may shift risk by only a small amount, in order to determine how aggressively to intervene.

Although suicide is a leading cause of death among people under 24 [1], it is rare enough that clinicians working with young people may not routinely screen for suicidal thoughts and / or may not have experience working with suicidal youth [12]. Low probability events with large consequences can be particularly dangerous because warning signs that do exist may not be noticed: Because the outcome – suicide – is not on the radar for most pediatricians or mental health providers, they could miss important signs that, if assessed, might have prevented the attempt. Similarly, beyond just considering the potential for suicide it is important that the risk assessment is conducted in a systematic way; when clinical assessments do not take a structured approach, the likelihood that something will be missed is much higher [13].

Unfortunately, because there are no risk factors that provide decisive certainty about suicide risk by themselves; clinicians must weigh protective factors and risk factors to reach a decision about whether a certain individual is in imminent risk to engage in suicidal behavior [12]. Though one might assume a low tolerance for risk when making the decision about hospitalization due to the potential consequences of suicidal thoughts and behaviors, these decisions are not easy – hospitalization means taking a child out of school and away from his/her family, it means risking stigma and ostracism from peers, it entails expense and inconvenience, it may damage the therapeutic relationship, and could increase the risk of suicide, according to some studies [14]. It is a difficult and high stakes decision, so it would be valuable to optimize the process in any ways that increase accuracy. Research has consistently shown that clinicians are poor at weighing and integrating different pieces of information and that actuarial approaches to data integration routinely outperform clinical judgment in diagnostic decision making tasks [15-17]. Part of the reason that clinicians tend to be less accurate in their diagnostic decisions than actuarial methods is that we humans rely on heuristics that can bias our impressions. For example – the availability heuristic [18] could lead a clinician working in an outpatient clinic to

underestimate a client's risk – "I've never seen a suicidal child before, so this one probably isn't either" – or it could lead a hospital-based clinician to over-estimate risk – "Most youth on the unit are suicidal, she must be, too." In both cases, the clinician may arrive at the wrong conclusion, in a high stakes situation, because past experience is influencing the decision. Heuristics can help us to efficiently process large amounts of information, but often lead to errors in clinical situations [19-21]; this may be particularly true when there are multiple pieces of information to consider. In the case of a suicide assessment, the youth is likely to have a number of *both* risk and protective factors – family history of depression, suicidal ideation, female sex, an involved and warm family. How would you weigh each piece of information? Which is most important? Using an actuarial approach could significantly improve the accuracy of predictions made about suicide risk, both by limiting the influence of bias and by objectively integrating the expected change in the probability of future suicide attempt associated with each known risk/protective factor [22,23].

An evidence based approach to assessment includes using well-validated assessment tools in combination with actuarial methods to inform diagnostic decisions [22,23]. Diagnostic assessments often include screening tools designed to determine what symptoms an individual is experiencing. Similarly, assessment tools are recommended to track client progress in order to determine how treatment is working [24]. However, once the assessment has been made and a score generated, it is not usually clear how that score influences an individual's risk for a certain outcome. Diagnostic likelihood ratios (DLR), which are derived from scores on assessment tools, and other factors influencing risk (family history, sex), are one way of quantifying the risk associated with a certain score/risk factor, but these are rarely published in psychiatry [23]. Making DLRs for commonly-used assessment tools easy for clinicians to access and understand could have a positive impact on the accuracy of the assessment of childhood disorders, including suicide risk [25,23]. With the DLRs corresponding to a client's assessment scores or other risk factors, a clinician can use a nomogram – a simple paper-and-pencil tool – to combine multiple

DLRs and determine the probability, based on all available information, that the individual will attempt suicide [26]. This approach is more consistent and accurate than clinical judgment [16,27,28] and can eliminate the influence of cognitive biases/heuristics by providing probabilities based on data rather than interpretation. Another important benefit of the nomogram is that it can easily include multiple sources of information, which is especially important when considering suicide, because few risk factors carry substantial risk and so the risk factors should be combined, when possible, to best estimate overall risk. For example, history of suicide behavior or ideation is related to future risk, but the effect size is not large [11]. However, by combining this information with other known risk factors, like female sex and older age, a more accurate and meaningful assessment of suicide risk probability could be made.

One of the most commonly used assessment tools for youth – both as an initial screening tool and to measure treatment response – is the Achenbach System of Empirically Based Assessment (ASEBA), which includes youth, parent, and teacher forms [29,30]. Because it is popular, ASEBA is more likely to inform an evidence based assessment (EBA) approach than other questionnaires, which might require a change in the "typical" process [31]. Consequently, we have chosen to investigate the diagnostic utility of the ASEBA system for the identification of youth at risk for suicide. The CBCL has been studied previously in regard to its ability to assess for suicide risk; Jones et al. [32] found that, among youth with epilepsy, the CBCL Total Problems Scale did an adequate job of identifying risk for suicidal ideation (sensitivity = .79, specificity= .64). Related, scales from the YSR have been evaluated as predictors of future suicide behavior; specifically, the social problems scale of the YSR was associated with suicide attempt four years later in a sample of adolescents [33]. However, the ASEBA system has not been evaluated as a tool for identifying youth at risk for suicide in a psychiatric sample. Our goals were to (1) determine whether the ASEBA system has validity for the identification of youth at risk for suicide, based on the presence of suicidal ideation (current or past) or previous suicide attempt, in a mental health treatment seeking population, (2) compare youth self-report (YSR),

parent (CBCL), and teacher (TRF) reports to see whether there are differences in the utility associated with each, and (3) provide DLRs for scores on the ASEBA scales, so that clinicians can easily incorporate scores into an evidence based suicide assessment.

Based on previous findings, we expected that the CBCL Total Problems scale would identify youth at risk for suicide based on suicidal ideation or suicidal behavior/attempt, either now or during the past six months. Additionally, we predicted that looking at specific, suicide-related items on the CBCL, rather than the broad scale, would improve specificity. We anticipated that the YSR Total Problems and YSR suicide-related items would also distinguish youth with suicide risk, and that caregiver report and youth self report would perform better than the teacher report [34]. Finally, we predicted that scale performance would be moderated by the presence of a mood disorder diagnosis, by child age, and sex; previous research has shown that mood disorder diagnoses are associated with suicide in youth [35], and that both male sex and older age are associated with greater risk for suicide [36].

Method

Participants

Youths aged 5 to 18 years and their parent/caregiver were recruited for two studies of childhood mental illness. Exclusion criteria included pervasive developmental disorder or mental retardation, non-English speaking caregiver or youth. Sample A (N=731) was recruited from an urban community mental health center [37]. Sample B (N=386) was recruited from a psychiatric research center with a focus on bipolar disorders, and a high rate of referrals from an affiliated adult mood disorders clinic [34,38].

Measures

Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS). All youth participants and their caregiver were interviewed by a research assistant using

the Schedule for Affective Disorders and Schizophrenia for School-Age Children- Present and Lifetime version (K-SADS-PL [39]), supplemented in one project with the Washington University KSADS mood modules [40]. All interviewers were required to achieve item-level kappa >0.85 on ten interviews with an experienced rater before conducting KSADS interviews on their own. Full details on the diagnostic process have been reported previously [34,38]. The K-SADS includes a section on suicide that assesses for recurrent thoughts of death, suicidal ideation (including frequency), and suicidal acts (including intent and lethality). Summary responses on this section, which took into account youth report, parent report, and known patient history (e.g., previous hospitalization for suicide attempt) formed the diagnostic standard for suicide risk against which the ASEBA measures were assessed. In order to meet criteria for suicidal ideation, the summary response had to be above threshold, indicating recurrent thoughts of death. Similarly, in order to meet criteria for a suicide attempt, the summary score had to reflect self injurious behavior with the intent to die.

Child Behavior Checklist (CBCL). Caregivers completed the 118-item CBCL about their child [41,29]. Items were rated from 0 (*Not True (as far as you know)*) to 2 (*Very True or Often True*). When the study was initiated, the 1991 version was used; later we switched to the 2001 version when it became available [34] – item changes were minimal. All versions of the Achenbach checklists ask whether the item content happened "now or within the past six months." The present study focused on two questions related to suicide: (18) "Deliberately harms self or attempts suicide" and (91) "Talks about killing self." Additionally, we looked at the Total Problems scale to determine whether the Suicide Items offered any improvement for detecting suicide risk. Reliability was good, Cronbach's alpha=.93 for Total Problems in the combined sample.

Youth Self Report (YSR). Participants between the ages of 11 and 18 years completed the YSR [42,29]. The YSR is largely equivalent to the CBCL; items used for the present study

were (18) “I deliberately try to hurt or kill myself” and (91) “I think about killing myself.” Reliability for the Total Problems scale of the YSR was good, Cronbach’s alpha=.96.

Teacher Report Form (TRF). Families selected the teacher who knew their child best, asking her or him to complete the Achenbach TRF [29,43]. The items used for this study were (18) “Deliberately harms self or attempts suicide” and (91) “Talks about killing self.” The Total Problems scale had good reliability, Cronbach’s alpha=.97.

Procedure

All procedures were approved by the Institutional Review Board. In both samples, the caregivers and youth completed an informed consent process and were compensated for their time. All measures included in the present study were collected at baseline. Procedures, as they relate to the present study, were largely the same for both samples; youths and their caregiver completed the K-SADS interview. The Longitudinal Evaluation of All Available Data (LEAD) standard of diagnosis, which takes into account the K-SADS interview, family history, prior treatment history, and clinical judgment, but not self- or parent-report questionnaires, was used to make all diagnoses [44]. Kappa was 0.91 comparing the LEAD diagnosis with the K-SADS diagnosis [34].

Analytic Plan

Chi-squared and *t*-tests compared the demographic and clinical characteristics of youth with a history of suicide behavior to those without. We estimated the area under the curve (AUC) from receiver operating characteristic (ROC) analyses for each of the Suicide Items (CBCL, YSR, and TRF) and each of the Total Problems scales to quantify the diagnostic efficiency of each for identifying youth with a history of suicide behavior and for identifying youth with current or past suicidal ideation. We then compared the AUCs using Venkatraman’s permutation test, which compares two ROC curves at every operating point, rather than just comparing the

overall AUCs [45]; consequently, this test is more sensitive to differences – even when the overall AUC estimates are the same – than other tests by which two ROC curves may be compared. This allowed us to determine whether one or more scales outperformed the others.

We tested potential moderators of scale performance, including sample (A or B), sex, age (<11, 11+), presence of a depression diagnosis (Y or N), and presence of a bipolar disorder diagnosis (Y or N). For those moderator variables that were significant predictors in logistic regression predicting history of suicide behavior, AUCs were calculated separately, in order to determine whether scale performance was significantly affected.

After determining which moderator variables affected scale performance, we calculated diagnostic likelihood ratios (DLRs) for the scale scores that were best at identifying risk for suicide. The DLRs were based on scale scores that yielded the best balance of sensitivity and specificity from ROC curves [46]. DLRs are a way by which to understand how an individual's score on a measure changes the probability the s/he is at risk for suicide; DLRs less than 1.0 are associated with lower than average odds, whereas between 2 and 5 represent a small increase of the odds and DLRs between 5.0 and 10.0 represent a moderate increase in risk, and DLRs greater than 10 can be clinically decisive [26].

We used SPSS-Version 22.0 or the pROC package[46] in R for all analyses. The missing at random assumption for missing data was tenable, so analyses used listwise deletion.

Results

Demographics

The demographic characteristics of Sample A and Sample B were slightly different, Sample A was slightly younger (p=.027) and had more minority youth (p<.001). Sample B had more youth with a history of suicide behavior (p<.001) and more youth with a bipolar or depression diagnosis (both ps<.001). Table S1 reports demographic information for each sample separately.

The average age of the participants in the combined sample was 11.34 (3.33), 40% (*n*=451) were female, and 51% (*n*=566) were Black. Three percent of youth (*n*=33) in the combined sample had a history of suicide behavior and 7% (*n*=78) had current or past suicidal ideation per the KSADS interview. Table 1 presents demographic characteristics of youth with a history of suicidal behavior and youth with no history of suicidal behavior. Youth with a history of suicide attempt were, on average, older (p=.025), more likely to be female (p<.0005), more likely to be black (p=.007), and more likely to have a mood disorder diagnosis (ps<.01).

Diagnostic Efficiency Statistics

We first investigated whether the Suicide Items and Total Problems *T*-Scores performed differently in samples A and B. Venkatraman's test compared the AUCs for each scale identifying youth with history of suicide behavior and for each scale identifying youth with suicidal ideation. There were no differences in the performance of any scale between the two samples (all ps>.08). Therefore, subsequent analyses were done with the two samples combined in order to improve the generalizability of our results.

The Suicide Items from each questionnaire were related, but each assesses a distinct risk factor for suicide; the CBCL items were moderately correlated (r=.45, p<.001), as were the YSR (r=.63, p<.001) and TRF (r=.40, p<.001) items. AUCs for the Suicide Items identifying youth with a history of suicide behavior from the CBCL (.85) and YSR (.70) were good (see Table 2). The AUC for the TRF Suicide Items identifying youth with a history of suicide behavior was poor (.45). The AUCs for the Total Problems scores identifying youth with a history of suicide behavior were poor-to-fair (.33-.65). The CBCL Suicide Items score had the largest AUC; Venkatraman's test compared AUCs for all the other scale scores against the CBCL Suicide Items score. The CBCL Suicide Items outperformed all the other scale scores for identifying youth with a history of suicide behavior (ps=.04 to <.0005).

We also calculated the AUCs for the Suicide Items scores for identifying youth with suicidal ideation; these ranged from good (.79 CBCL; .72 YSR) to fair (.61 TRF). The AUCs

from the Total Problems *T*-scores were not as good; ranging from poor (.53 CBCL; .53 TRF) to fair (.69 YSR). The AUC for the CBCL Suicide Items was the highest, consequently all the other scales were compared to it using Venkatraman's test. The CBCL Suicide Items outperformed all the Total Problems *T*-Scores (*p*s =.013 to <.001), but it was not better than the YSR Suicide Items (*p*=.159) or the TRF Suicide Items (*p*=.072) at identifying youth with suicidal ideation.

Moderator analyses indicated that the Suicide Items might perform differently depending on whether or not the youth had a depressive disorder diagnosis. Consequently, we calculated the AUCs separately for each of the Suicide Items and for the Total Problems *T*-Scores for youth with a depressive disorder and for those with no depressive disorder (including bipolar disorder) diagnosis. Venkatraman's test compared AUCs for each scale in youth with a depressive disorder diagnosis against those who did not have a depressive disorder diagnosis (see Table S2). With the exception of the CBCL Total Problems *T*-Score for identifying suicidal ideation, there were no significant differences in the AUCs for any scale. The CBCL Total Problems *T*-Score performed better in youth with a depressive disorder diagnosis than in youth without a depressive disorder diagnosis (*p*=.029) for identifying youth with suicidal ideation. Scale performance was not significantly moderated by youth bipolar diagnoses, sex, or age. The results comparing AUCs for youth with a bipolar diagnosis to those without, for males and females, and for youth under 11 to those 11 and older are available upon request.

The CBCL Suicide Items performed as well or better than the other scales we tested for identifying youth with a history of suicide behavior or suicidal ideation. Therefore, in the interest of parsimony, we chose to calculate and report the diagnostic likelihood ratios for the CBCL Suicide Items. In addition, because there may be situations in which parent report is not available, we also determined the DLRs for the YSR Suicide Items. The DLRs for both scales were calculated in order to find the best balance between sensitivity and specificity [47] for identifying history of suicide behavior and for identifying suicidal ideation (see Table 3). We determined that two DLRs –low risk and elevated risk – provided the best way by which to identify youth at risk

for suicide behavior, whether due to a history of suicide behavior or to suicidal ideation. A total score of one (out of four) on the CBCL Suicide Items was associated with a substantial increase in the probability of either history of suicide behavior (DLR=2.84) or suicidal ideation (DLR=2.67). Scores of zero could be helpful for ruling out either suicide behavior (DLR=0.13) or suicidal ideation (DLR=0.29). The DLRs associated with the YSR Suicide Items were not quite as influential as those associated with the CBCL Suicide Items, but did moderately increase the likelihood of suicide behavior (DLR=2.43) or ideation (DLR=2.72). We also calculated the DLRs for the sum of YSR and CBCL Suicide Item scores. Again, a score of one (out of a possible eight) offered the best balance between sensitivity and specificity. The DLRs for the sum score were similar to those for the CBCL alone for both attempt (DLR=2.62) and ideation (DLR=2.58). We used logistic regression to test whether there was benefit to combining scores from the CBCL and YSR using the nomogram; there was not (p=.43).

Discussion

Results suggest that the CBCL Suicide Items can be helpful at identifying youth who have engaged in suicidal behaviors and are at risk for future suicide attempt. Their performance also remained consistent across sex, age, and bipolar disorder status. The CBCL and YSR are commonly administered in mental health settings; checking the two suicide-related items would not add any burden to the family or clinician during an intake appointment, but could yield very important information. Importantly, the total problems scores for the CBCL and YSR did not perform as well as the suicide-specific items – this means that it is incumbent upon the clinician to check these items, rather than simply relying on the score profile to indicate risk. This is a process many clinicians will be familiar with, having learned to check pathognomonic items on other questionnaires. The difference might be remembering to think about suicide assessment when working with young people; an important advantage of using the suicide items from the CBCL and YSR is that they are already being administered. Other brief suicide assessments, like the Ask Suicide Questions (ASQ [48]) are also effective for identifying suicide risk, but require

an extra assessment. Additionally, because the Suicide Items are brief and easy to understand, they can be administered regularly to check for the presence of suicidal thoughts or behaviors. It is also important to recognize that for some sensitive topics, like suicidal thoughts and behaviors, adolescents might be more likely to disclose their experiences on a self-report than in a face-to-face interview [49-51]; an initial assessment with a broad self-report, like the YSR, followed by a more detailed evaluation may be best for elucidating risk.

The Suicide Items do not constitute a thorough suicide assessment, but as a screening tool, they provide helpful information for identifying those at risk. The CBCL suicide scale outperformed the YSR suicide scale for identifying suicide attempt in older youth (11-18). This may be related to hesitancy, on the part of the youth, to disclose suicide behaviors and suggests the importance of interviewing caregivers when evaluating youth suicide risk. Interestingly, the YSR suicide scale was equivalent to the CBCL suicide scale for identifying suicidal ideation; it may be that disclosing suicidal thoughts/behaviors becomes less likely as they become more severe. When an individual and/or her parents endorse one or both items, the odds of suicidal behavior more than double; and the clinician may want to administer a more comprehensive assessment, such as the Columbia SuicideScreen [52]. Screening for suicidality, like other psychiatric problems, can make an evaluation more efficient and effective; evidence based assessment recommends a twelve-step process to inform all aspects of psychological services, including diagnosis, treatment, process evaluation, outcome measurement, and maintenance monitoring [53]. To illustrate how one might use the CBCL and YSR Suicide Items to screen for suicide risk, we offer the following vignette:

Case example

Shelly is a 14-year-old, Caucasian girl who was brought to your clinic by her mother, Sharon, due to concerns about her irritable mood and morbid ideation, including saying things like, “Wouldn’t it be better if I had never been born.” When Shelly and Sharon come to your office, Sharon reports that Shelly has become “difficult to live with” and says that she and her

partner “walk on egg shells” around Shelly. Shelly says very little in response to her mother’s report, only that “Mom and Morgan get on my nerves.” You ask Sharon to complete the CBCL while you meet with Shelly privately. One-on-one, Shelly is a little more talkative and acknowledges that she does not get along well with her mother and her mother’s partner, but says that she is doing well in school and does not have “any problems except that I have to live with those two.” When you ask about her mother’s report that she has made comments about wishing she had never been born, Shelly says she doesn’t remember saying that and denies any past or current suicidal thoughts. She reports that she has not engaged in self harm and would “never do something like that.” After meeting with Shelly, you ask her to complete the YSR while you talk with Sharon. Sharon reports that Shelly is “minimizing” the extent of the problems and says that she is “really worried.” She confirms that Shelly’s school performance is fine and that she is still active socially, but says that she is a “completely different person” at home. Sharon also reports that her father died by suicide, and that she is concerned that Shelly might do something to hurt herself.

Shelly’s score on the CBCL total problems scale *T*-score was a 66, in the borderline clinical range. Her mother endorsed the item “Talks about killing self,” as “Very often true,” but did not endorse “Deliberately harms self or attempts suicide,” resulting in a score of 2 on the CBCL Suicide Items. On the YSR, Shelly’s total problems *T*-score was a 53, in the normal range. She did not endorse either suicide item, giving her a score of 0 on the YSR Suicide Items.

We can combine the information we have about Shelly’s risk for suicide to estimate the probability that she will engage in suicidal behavior in the proximal few months. First, we know that the prevalence of suicidal behavior in youth Shelly’s age is about 10% in our clinic, higher than the 3% reported for youth under 14 in the general population (versus around 8% among high school aged youth in the general population). Second, Shelly is a girl, putting her at higher risk for suicide attempt, about 3 times higher than a boy of the same age. Third, Shelly has a family history of suicide, which increases her risk by about 2.5 times. Finally, we have her score on the

CBCL Suicide Items, which is associated with a DLR of 2.84. Because there is no incremental benefit to including *both* the CBCL and YSR DLRs, we will focus on the CBCL. If you were meeting with a youth in a setting where parent report was not available, you could use the YSR DLR instead.

Rather than trying to weigh this information in our head to determine the risk, we can use a probability nomogram (a paper and pencil tool; see Van Meter et al., 2014 [54] for a complete description) to combine these pieces of information. Starting with a blank nomogram (Figure 1), we put the prevalence of suicide behavior on the left side – this is our best estimate of the probability of Shelly engaging in suicidal behavior, without any other information. Next, we add the likelihood ratio associated with female sex to the middle line. Drawing a line from .10 through 3, we arrive at a revised probability of 25% on the right side. We then move 25% to the left side as our new probability. We can now add the family history risk factor, by drawing a line from .25 through 2.5, resulting in a new probability of about 45%. We repeat the process to incorporate the DLRs for the CBCL Suicide Items, drawing from 45% through 2.84 (CBCL DLR+) to arrive at a posterior probability of 70%. We can add the risk factors in any order, the result will be the same.

The posterior probability makes it clear that Shelly is at risk for suicidal behavior. Her clinician might next opt to administer a more comprehensive suicide assessment to get a better sense for the severity and content of her suicidal thoughts and to more closely evaluate whether she has engaged in or considered any harmful behaviors. Although Shelly denied suicidal thoughts and the severity of her mood symptoms was not severe, the *combination* of risk factors led to a concerning level of risk. Taking a more typical approach to risk assessment, a clinician might feel confused about how to weight different pieces of information – is Sharon or Shelly giving a more accurate report of Shelly's ideation? By how much does sex or a family history of suicide change the equation? The nomogram, or some other method of combining likelihood ratios, takes the onus of weighting risk off the clinician, and results in a more accurate estimation

[28,55]. In addition, knowing ahead of time what the DLRs are for the CBCL and YSR Suicide Items, and how risk factors like family history and sex affect the probability of suicide, makes it easy to very quickly arrive at a posterior probability, enabling the clinician to conduct further assessment immediately (as necessary) and take the appropriate next steps. This is particularly important for suicide risk; waiting a week or more to score assessments and meet with the family again could result in a critical gap in care.

Limitations.

Using the CBCL or YSR to screen for suicide makes sense because, in many clinics, the Achenbach system is already being used, meaning that no extra time or expense is added. Additionally, scoring two two-point items is easy and quick. However, in spite of these strengths, there are some limitations to this approach. First, the data on which we calculated the DLRs were not longitudinal. Consequently, we were only able to look at *past* suicidal behavior or *past/current* suicidal ideation as our outcomes. The Achenbach items ask about "now or within the past six months." Ideally, we would want to know how well the Suicide Items perform at identifying people at risk for *future* suicide risk. However, past suicide behavior is probably the best established predictor of future behavior [56] and we maximized sensitivity with the DLRs, to favor false positives over false negatives. By definition, we had no completed suicides represented in our sample, so we cannot say if/how the Suicide Items relate to severity of attempt. However, we approach this with the goal of preventing *any* attempt. The sample also covered a large age range, and the rates of suicidal ideation and behavior would be higher if working only with an adolescent sample; however, there was a history of attempts across the age range, and it is helpful to know that the Achenbach items were equally informative across the broad age range. Finally, this sample was treatment seeking, and as such, parents may have been more aware of suicidal thoughts/behaviors than parents in the general population would be.

Conclusion

Our results suggest that the two CBCL Suicide Items are a fast and easy way to assess for suicide risk, associated with an almost *three times higher* risk when either question is endorsed, even as "somewhat or sometimes true." Related, when youth endorse any suicidal thoughts or risky behaviors, resulting in a score of one or more on the two items combined, their risk is more than doubled. Many clinicians do not routinely screen for suicide in their young patients and may not feel comfortable doing so [12]; using the Achenbach measures enables the integration of a suicide screen without any additional time or expense. When any suicidal thoughts or behaviors are endorsed, the likelihood ratios from the CBCL and/or YSR can be combined with other known risk factors to give an accurate estimate of risk and inform next steps. Importantly, this is true independent of diagnosis, indicating that these two questions, can offer a helpful index of risk above-and-beyond diagnostic profile.

Previous reports have not directly addressed the importance of parent report when screening adolescents for suicide [57], but our results suggest that parent report is informative, and when possible, should be collected. In contrast, teacher report was not associated with suicide risk; this indicates that, not only is the extra effort of collecting teacher report unnecessary in a suicide assessment, but also that it may actually impede diagnostic efficiency. As we see suicide rates increasing among some youth, with few theories about why or how to stop it, broad screening for risk may become increasingly important [57] – particularly in school or pediatrician offices to better reach all youth. Counting up these items – or simply inquiring about the presence (current or past) of suicidal ideation or suicidal behavior – requires no special training, scoring software, or psychological insight, but doing so could make a difference in our ability to identify and help those at risk.

Author Note:

Dr. Feeny has no conflict of interest to report.

Dr. Findling receives or has received research support, acted as a consultant and/or served on a speaker's bureau for Actavis, Akili, Alcobra, American Academy of Child & Adolescent Psychiatry, American Psychiatric Press, Bracket, CogCubed, Cognition Group, Coronado Biosciences, Elsevier, Epharma Solutions, Forest, Genentech, GlaxoSmithKline, Guilford Press, Ironshore, Johns Hopkins University Press, KemPharm, Lundbeck, Medgenics, Merck, NIH, Neurim, Novartis, Otsuka, PCORI, Pfizer, Physicians Postgraduate Press, Purdue, Rhodes Pharmaceuticals, Roche, Sage, Shire, Sunovion, Supernus Pharmaceuticals, Syneurx, Takeda, Teva, TouchPoint, Tris, Validus, and WebMD.

Ms. Lechtman has no conflict of interest to report.

Dr. Perez Algorta has no conflict of interest to report.

Dr. Van Meter has no conflict of interest to report.

Dr. J. Youngstrom has no conflict of interest to report.

Dr. E. Youngstrom has consulted with Pearson, Janssen, Otsuka, Lundbeck, and Western Psychological Services about psychological assessment.

References

1. Centers for Disease Control (2015) Youth Suicide. Accessed January 8, 2016

2. Curtin S, Warner M, Hedegaard H (2016) Increase in suicide in the United States, 1999–2014. National Center for Health Statistics. Accessed May 11 2016

3. Nock M, Green J, Hwang I, McLaughlin I, Sampson N, Zaslavsky A, Kessler R (2013) Prevalence, correlates, and treatment of lifetime suicidal behavior among adolescents: Results from the national comorbidity survey replication adolescent supplement. JAMA Psychiatry 70 (3):300-310. doi:10.1001/2013.jamapsychiatry.55

4. Nock M, Borges G, Bromet E, Cha C, Kessler R, Lee S (2008) Suicide and Suicidal Behavior. Epidemiologic Reviews 30:133-154. doi:10.1093/epirev/mxn002

5. Bridge JA, Goldstein TR, Brent DA (2006) Adolescent suicide and suicidal behavior. Journal of Child Psychology and Psychiatry 47 (3-4):372-394. doi:10.1111/j.1469-7610.2006.01615.x

6. Cash SJ, Bridge JA (2009) Epidemiology of youth suicide and suicidal behavior. Current Opinion in Pediatrics 21 (5):613

7. World Health Organization (2009) Global health risks: mortality and burden of disease attributable to selected major risks. WHO Press, Geneva, Switzerland

8. Joiner TE (1999) The Clustering and Contagion of Suicide. Current Directions in Psychological Science 8 (3):89-92. doi:10.1111/1467-8721.00021

9. Randall JR, Nickel NC, Colman I (2015) Contagion from peer suicidal behavior in a representative sample of American adolescents. Journal of Affective Disorders 186:219-225. doi:http://dx.doi.org/10.1016/j.jad.2015.07.001

10. Dilillo D, Mauri S, Mantegazza C, Fabiano V, Mameli C, Zuccotti GV (2015) Suicide in pediatrics: epidemiology, risk factors, warning signs and the role of the pediatrician in detecting them. Italian journal of pediatrics 41:49. doi:10.1186/s13052-015-0153-3

11. Ribeiro JD, Franklin JC, Fox KR, Bentley KH, Kleiman EM, Chang BP, Nock MK (2016) Self-injurious thoughts and behaviors as risk factors for future suicide ideation, attempts, and death: a meta-analysis of longitudinal studies. Psychological Medicine 46 (02):225-236. doi:doi:10.1017/S0033291715001804

12. Helms SW, Prinstein MJ (2014) Risk assessment and decision making regarding imminent suicidality in pediatric settings. Clinical Practice in Pediatric Psychology 2 (2):176

13. Youngstrom EA (2015) Using Standardized Methods to Assess Suicidal Behavior: The Need is Even Greater Than it Looks (Commentary). The Journal of Clinical Psychiatry 76 (10):e1331-e1332. doi:10.4088/JCP.14com09573

14. Qin P, Nordentoft M (2005) Suicide risk in relation to psychiatric hospitalization: Evidence based on longitudinal registers. Archives of General Psychiatry 62 (4):427-432. doi:10.1001/archpsyc.62.4.427
15. Dawes R, Faust D, Meehl P (1989) Clinical versus actuarial judgment. Science 243 (4899):1668-1674. doi:10.1126/science.2648573
16. Grove WM, Zald DH, Lebow BS, Snitz BE, Nelson C (2000) Clinical versus mechanical prediction: A meta-analysis. Psychological Assessment 12 (1):19-30. doi:10.1037/1040-3590.12.1.19
17. Aegisdottir S, White MJ, Spengler PM, Maugherman AS, Anderson LA, Cook RS, Nichols CN, Lampropoulos GK, Walker BS, Cohen G, Rush JD (2006) The Meta-Analysis of Clinical Judgment Project: Fifty-Six Years of Accumulated Research on Clinical Versus Statistical Prediction. The Counseling Psychologist 34 (3):341-382. doi:10.1177/0011000005285875
18. Croskerry P (2009) Clinical cognition and diagnostic error: applications of a dual process model of reasoning. Advances in Health Sciences Education 14 (0):27-35. doi:10.1007/s10459-009-9182-2
19. Lilienfeld SO, Lynn SJ (2014) Errors/Biases in Clinical Decision Making. In: The Encyclopedia of Clinical Psychology. John Wiley & Sons, Inc. doi:10.1002/9781118625392.wbecp567
20. Jenkins MM, Youngstrom EA (2016) A Randomized Controlled Trial of Cognitive Debiasing Improves Assessment and Treatment Selection for Pediatric Bipolar Disorder. Journal of Consulting and Clinical Psychology. doi:10.1037/ccp0000070
21. Van Meter A, Youngstrom E, Youngstrom J, Feeny N, Findling R (2011) Examining the validity of cyclothymic disorder in a youth sample Journal of Affective Disorders 132 (1-2):55-63. doi:10.1016/j.jad.2011.02.004
22. Youngstrom EA (2012) Future Directions in Psychological Assessment: Combining Evidence-Based Medicine Innovations with Psychology's Historical Strengths to Enhance Utility. Journal of Clinical Child & Adolescent Psychology 42 (1):139-159. doi:10.1080/15374416.2012.736358
23. Youngstrom EA, Choukas-Bradley S, Calhoun CD, Jensen-Doss A (2014) Clinical Guide to the Evidence-Based Assessment Approach to Diagnosis and Treatment. Cognitive and Behavioral Practice (0). doi:http://dx.doi.org/10.1016/j.cbpra.2013.12.005
24. Lambert MJ, Hansen NB, Finch AE (2001) Patient-focused research: Using patient outcome data to enhance treatment effects. Journal of Consulting and Clinical Psychology 69 (2):159-172. doi:10.1037/0022-006x.69.2.159

25. Christon LM, McLeod BD, Jensen-Doss A (2014) Evidence-Based Assessment Meets Evidence-Based Treatment: An Approach to Science-Informed Case Conceptualization. Cognitive and Behavioral Practice (0). doi:http://dx.doi.org/10.1016/j.cbpra.2013.12.004
26. Straus SE, Glasziou P, Richardson WS, Haynes RB (2011) Evidence-based medicine: How to practice and teach EBM. 4th edn. Churchill Livingstone, New York, NY
27. Hamilton JD (2001) Do we under utilise actuarial judgement and decision analysis? Evidence Based Mental Health 4 (4):102-103. doi:10.1136/ebmh.4.4.102
28. Jenkins M, Youngstrom E, Washburn J, Youngstrom J (2011) Evidence-based strategies improve assessment of pediatric bipolar disorder by community practitioners. Professional Psychology: Research and Practice 42 (2):121
29. Achenbach T, Rescorla L (2001) Manual for the ASEBA school-age forms & profiles. Research Center for Children, Youth, and Families, Burlington
30. Pauschardt J, Remschmidt H, Mattejat F (2010) Assessing child and adolescent anxiety in psychiatric samples with the Child Behavior Checklist. J Anxiety Disord 24 (5):461-467. doi:10.1016/j.janxdis.2010.03.002
31. Achenbach TM (2005) Advancing Assessment of Children and Adolescents: Commentary on Evidence-Based Assessment of Child and Adolescent Disorders. Journal of Clinical Child & Adolescent Psychology 34 (3):541-547. doi:10.1207/s15374424jccp3403_9
32. Jones JE, Siddarth P, Gurbani S, Shields WD, Caplan R (2013) Screening for suicidal ideation in children with epilepsy. Epilepsy & Behavior 29 (3):521-526. doi:http://dx.doi.org/10.1016/j.yebeh.2013.09.020
33. Ferdinand RF, van der Ende J, Verhulst FC (2004) Parent-Adolescent Disagreement Regarding Psychopathology in Adolescents From the General Population as a Risk Factor for Adverse Outcome. Journal of Abnormal Psychology 113 (2):198-206. doi:10.1037/0021-843X.113.2.198
34. Youngstrom E, Meyers O, Demeter C, Youngstrom J, Morello L, Piiparinen R, Feeny N, Calabrese J, Findling R (2005) Comparing diagnostic checklists for pediatric bipolar disorder in academic and community mental health settings. Bipolar Disorders 7:507-517. doi:10.1111/j.1399-5618.2005.00269.x
35. GrØHolt B, Ekeberg Ø, WichstrØM L, Haldorsen TOR (1998) Suicide Among Children and Younger and Older Adolescents in Norway: A Comparative Study. Journal of the American Academy of Child & Adolescent Psychiatry 37 (5):473-481. doi:http://dx.doi.org/10.1097/00004583-199805000-00008

36. Brent DA, Baugher M, Bridge J, Chen T, Chiappetta L (1999) Age- and Sex-Related Risk Factors for Adolescent Suicide. Journal of the American Academy of Child & Adolescent Psychiatry 38 (12):1497-1505. doi:http://dx.doi.org/10.1097/00004583-199912000-00010
37. Youngstrom E, Findling R, Calabrese J, Gracious B, Demeter C, DelPorto Bedoya D, Price M (2004) Comparing the diagnostic accuracy of six potential screening instruments for bipolar disorder in youths aged 5 to 17 years. Journal of the American Academy of Child and Adolescent Psychiatry 43 (7):847-858. doi:10.1097/01.chi.0000125091.35109.1e
38. Findling R, Youngstrom E, McNamara N, Stansbrey R, Demeter C, Bedoya D, Kahana S, Calabrese J (2005) Early symptoms of mania and the role of parental risk. Bipolar Disorders 7 (6):623-634. doi:10.1111/j.1399-5618.2005.00260.x
39. Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, Williamson D, Ryan N (1997) Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): Initial reliability and validity data. Journal of the American Academy of Child and Adolescent Psychiatry 36 (7):980-988. doi:10.1097/00004583-199707000-00021
40. Geller B, Zimerman B, Williams M, Bolhofner K, Craney JL, DelBello MP, Soutullo C (2001) Reliability of the Washington University in St. Louis Kiddie Schedule for Affective Disorders and Schizophrenia (WASH-U-KSADS) mania and rapid cycling sections. Journal of the American Academy of Child & Adolescent Psychiatry 40 (4):450-455
41. Achenbach T (1991) Manual for the child behavior checklist. Burlington, VT: University of Vermont, Department of Psychiatry.
42. Achenbach T (1991) Manual for the Youth Self-Report and 1991 profile. Dept. of Psychiatry, University of Vermont,
43. Achenbach T (1991) Teacher Report Form. Department of Psychiatry, University of Vermont, Burlington, VT
44. Spitzer R (1983) Psychiatric diagnosis: Are clinicians still necessary? Comprehensive Psychiatry 24 (5):399-411. doi:10.1016/0010-440X(83)90032-9
45. Venkatraman ES (2000) A Permutation Test to Compare Receiver Operating Characteristic Curves. Biometrics 56 (4):1134-1138. doi:10.2307/2677047
46. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Muller M (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC bioinformatics 12:77. doi:10.1186/1471-2105-12-77

47. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, Müller M (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 12:77. doi:10.1186/1471-2105-12-77

48. Horowitz LM, Bridge JA, Teach SJ, et al. (2012) Ask suicide-screening questions (asq): A brief instrument for the pediatric emergency department. Archives of Pediatrics & Adolescent Medicine 166 (12):1170-1176. doi:10.1001/archpediatrics.2012.1276

49. Joiner TE, Rudd MD, Rajab MH (1999) Agreement between self- and clinician-rated suicidal symptoms in a clinical sample of young adults: Explaining discrepancies. Journal of Consulting and Clinical Psychology 67 (2):171-176. doi:10.1037/0022-006X.67.2.171

50. Kaplan ML, Asnis GM, Sanderson WC, Keswani L, de Lecuona JM, Joseph S (1994) Suicide assessment: Clinical interview vs. self-report. Journal of Clinical Psychology 50 (2):294-298. doi:10.1002/1097-4679(199403)50:2<294::AID-JCLP2270500224>3.0.CO;2-R

51. Bridge JA, Barbe RP, Birmaher B, Kolko DJ, Brent DA (2005) Emergent Suicidality in a Clinical Psychotherapy Trial for Adolescent Depression. American Journal of Psychiatry 162 (11):2173-2175. doi:10.1176/appi.ajp.162.11.2173

52. Shaffer D, Scott M, Wilcox H, Maslow C, Hicks R, Lucas CP, Garfinkel R, Greenwald S (2004) The Columbia SuicideScreen: Validity and Reliability of a Screen for Youth Suicide and Depression. Journal of the American Academy of Child & Adolescent Psychiatry 43 (1):71-79. doi:http://dx.doi.org/10.1097/00004583-200401000-00016

53. Youngstrom E, Van Meter A (in press) Empirically Supported Assessment of Children and Adolescents. Clinical Psychology: Science & Practice

54. Van Meter A, Youngstrom E, Youngstrom JK, Ollendick T, Demeter C, Findling RL (2014) Clinical Decision Making About Child and Adolescent Anxiety Disorders Using the Achenbach System of Empirically Based Assessment. J Clin Child Adolesc Psychol. doi:10.1080/15374416.2014.883930

55. Straus SE, Glasziou P, Richardson WS, Haynes RB (2011) Evidence-based medicine: How to practice and teach it. 4 edn. Churchill Livingstone Elsevier,

56. Meyer RE, Salzman C, Youngstrom EA, Clayton PJ, Goodwin FK, Mann JJ, Alphs LD, Broich K, Goodman WK, Greden JF, Meltzer HY, Normand S-LT, Posner K, Shaffer D, Oquendo MA, Stanley B, Trivedi MH, Turecki G, Beasley CM, Beautrais AL, Bridge JA, Brown GK, Revicki DA, Ryan ND, Sheehan DV (2010) Suicidality and risk of suicide—definition, drug safety concerns, and a necessary target for drug development. J Clin Psychiatry 71 (08):e1-e21. doi:10.4088/JCP.10cs06070blu

57. Horowitz LM, Bridge JA, Pao M, Boudreaux ED (2014) Screening Youth for Suicide Risk in Medical Settings: Time to Ask Questions. American Journal of Preventive Medicine 47 (3, Supplement 2):S170-S175. doi:http://dx.doi.org/10.1016/j.amepre.2014.06.002

Table 1

Demographic and Clinical Information

Variable	No history of suicide behavior *n*=1085	History of suicide behavior *n*=33	Test Statistic	*p*
Age in years (*Mean, SD*)	11.31 (3.31)	12.63 (3.56)	$t(1116)=-2.25$	.025
Female	*n*=428	*n*=23	$X^2(1)=12.18$	<.0005
Race (Black)	*n*=556[a]	*n*=10[b]	$X^2(4)=14.12$	.007
Depression	*n*=159	*n*=11	$X^2(1)=8.67$	.003
Bipolar spectrum disorder	*n*=297	*n*=16	$X^2(1)=7.06$	.008
	Mean (*SD*)			
Number of DSM-IV-TR Axis I diagnoses	3.13 (1.75)	2.58 (1.62)	$X^2(10)=13.53$	.195
CBCL Suicide Items	.51 (0.87)	2.12 (1.27)	$t(32.9)=-7.25$	<.0005
CBCL total problems	68.57 (9.51)	71.70 (9.46)	$t(1113)=-1.86$	.063
YSR Suicide Items	.53 (1.00)	1.21 (1.03)	$t(564)=-2.91$	.004
YSR total problems	58.63 (12.15)	64.52 (9.71)	$t(603)=-2.20$	.028
TRF Suicide Items	.17 (0.55)	0 (0)	$t(368)=0.81$	.418
TRF total problems	62.15 (10.81)	56.00 (9.71)	$t(409)=1.95$	.052

[a]Complete racial breakdown: 556 Black youth, 452 White youth, 50 multiracial youth, 24 Hispanic youth, and 3 Asian youth. [b]Complete racial breakdown: 21 White youth, 10 Black youth, 1 multiracial youth, and 1 Asian youth.

Table 2.

ROC[1] *Analyses Comparing CBCL Suicide Score with Other Achenbach Subscales*

Target Outcome	Achenbach Score	Area Under Curve	95% Confidence Interval		*p* for Difference between AUCs[2]
			Lower	Upper	
Suicide Attempt	CBCL Suicide Items	.85***	.79	.92	
	CBCL total problems	.60*	.49	.71	<.0005
	YSR Suicide Items	.70**	.58	.82	.043
	YSR total problems	.65*	.54	.77	<.0005
	TRF Suicide Items	.45	.25	.64	.0095
	TRF total problems	.33*	.19	.47	.004
Suicide Ideation	CBCL Suicide Items	.79***	.74	.84	
	CBCL total problems	.53	.46	.60	<.0005
	YSR Suicide Items	.72***	.65	.79	.159
	YSR total problems	.69***	.63	.75	.013
	TRF Suicide Items	.61	.49	.74	.072
	TRF total problems	.53	.42	.65	<.0005

* $p<.05$, ** $p<.005$, *** $p<.0005$

[1]Receiver Operating Characteristic; [2]Area Under the Curve – compared to CBCL Suicide Items

Table 3.

Diagnostic Likelihood Ratios Based on Optimal Cut-off Scores for Identification of Suicide Behavior or Ideation

		Score Range					
Suicide Behavior (Base rate = 3%)		Low Risk	Elevated Risk	Sensitivity	Specificity	Positive predictive value	Negative predictive value
CBCL Suicide Items	Score	<1	1+				
	DLR	0.13	2.84	.91	.69	.07	.99
YSR Suicide Items	Score	<1	1+				
	DLR	.44	2.43	.68	.72	.08	.98
Sum of CBCL and YSR Suicide Items	Score	<1	1+				
	DLR	.05	2.62	.97	.63	.07	~1.00
		Score Range					
Suicide Ideation (Base rate = 7%)		Low Risk	Elevated Risk	Sensitivity	Specificity	Positive predictive value	Negative predictive value
CBCL Suicide Items	Score	<1	1+				
	DLR	.29	2.67	.78	.71	.17	.99
YSR Suicide Items	Score	<1	1+				
	DLR	.43	2.72	.68	.75	.24	.95
Sum of CBCL and YSR Suicide Items	Score	<1	1+				
	DLR	.22	2.58	.85	.66	.17	.98

Table S1.

Demographic and clinical data for each sample

Variable	Sample A *n*=731	Sample B *n*=386	Test Statistic	*p*
Age in years (*Mean, SD*)	11.18 (3.26)	11.65 (3.43)	$t(751)=2.21$	.027
Female	*n*=296	*n*=155	$X^2(1)=0.02$	.886
Race (Black)	*n*=512[b]	*n*=54[a]	$X^2(4)=343.01$	<.0005
Depression	*n*=74	*n*=96	$X^2(1)=42.31$	<.0005
Bipolar spectrum disorder	*n*=131	*n*=182	$X^2(1)=107.01$	<.0005
History of suicide	*n*=12	*n*=21	$X^2(1)=12.65$	<.0005
History of suicidal ideation	*n*=57	*n*=21	$X^2(1)=2.15$	.143
	Mean (*SD*)			
Number of DSM-IV-TR Axis I diagnoses	3.72 (1.72)	1.98 (1.14)	$X^2(10)=285.22$	<.0005
CBCL Suicide Items	.51 (.91)	.63 (.96)	$t(1095)=2.31$	.021
CBCL total problems	68.84 (8.66)	68.34 (10.98)	$t(1113)=-.84$	.404
YSR Suicide Items	.55 (.99)	.59 (1.08)	$t(559)=.40$	.688
YSR total problems	58.62 (11.96)	59.33 (12.48)	$t(603)=.67$	.503
TRF Suicide Items	.17 (.55)	.15 (.51)	$t(361)=-.30$	.764
TRF total problems	64.23 (9.50)	58.27 (11.82)	$t(409)=-5.62$	<.0005

[a]Complete racial breakdown: 304 White youth, 54 Black youth, 16 multiracial youth, 11 Hispanic youth, and 2 Asian youth. [b]Complete racial breakdown: 512 Black youth, 169 White youth, 35 multiracial youth, 13 Hispanic youth, and 2 Asian youth.

Table S2.

ROC[1] Analyses Comparing CBCL Suicide Score with Other Achenbach Subscales comparing youth with major depressive disorder to youth without major depressive disorder (minus youth with bipolar disorder)

Target Outcome	Achenbach Score	MDD Area Under Curve	No MDD Area Under Curve	*p* for Difference between AUCs[2]
Suicide Attempt	CBCL total problems	.52	.68	.356
	CBCL Suicide Items	.78	.96	.298
	YSR total problems	.55	.69	.303
	YSR Suicide Items	.57	.88	.224
	TRF total problems	.41	.58	.922
	TRF Suicide Items	.42	.47	.999
Suicide Ideation	CBCL total problems	.64	.57	.029
	CBCL Suicide Items	.73	.79	.483
	YSR total problems	.57	.69	.317
	YSR Suicide Items	.61	.73	.419
	TRF total problems	.58	.67	.686
	TRF Suicide Items	.61	.54	.978

[1]Receiver Operating Characteristic; [2]Area Under the Curve

Figure 1.

Example of how DLRs can be combined using a nomogram to arrive at an estimated probability of suicide behavior