

# School-Based Screening for Suicide Risk: Balancing Costs and Benefits

Michelle Scott, PhD, MSW, Holly Wilcox, PhD, Yanling Huo, MS, J. Blake Turner, PhD, Prudence Fisher, PhD, and David Shaffer, FRCP, FRCPsych

Adolescents commonly keep their thoughts about suicide to themselves and many suicide attempts go unrevealed to parents and other adults<sup>1-4</sup>; furthermore, adolescents rarely seek treatment on their own.<sup>5,6</sup> Proactive screening programs for adolescent suicidality rely on the identification of the principal risk factors for completed suicide (i.e., current suicidal ideation, previous attempt behavior, and the presence of a mood, anxiety, or substance use disorder).<sup>5,6</sup>

One proactive screening program, the Columbia Teen Screen program, has employed a school-based screening approach, the Columbia Suicide Screen (CSS), that involves administering a self-completion form with questions about risk factors for suicide such as suicidal ideation, prior suicide attempts, depression, anxiety, and substance use. Students who screen positive (stage 1) are then seen by a clinician for a secondary confirmatory evaluation (stage 2) and, if indicated, the student is case managed to an appropriate referral. With its original algorithm, the CSS has been shown to identify 75% of students considered to be at high risk for suicide<sup>7</sup> and a third of students who had unspecified mental health problems that were not already known to school professionals.<sup>8</sup> Criticisms that the approach generated many false positives were based on our previous reports<sup>7</sup> that were limited to identifying high-risk cases<sup>9-11</sup> and ignored the fact that screening for suicidal ideation and behaviors will commonly reveal nonsuicidal mental illnesses that have never been disclosed.

Much of the cost of screening comes from providing confirmatory evaluations to students identified during the initial part of a 2-stage procedure. Falsely identifying students who do not have a significant mental health problem adds to the cost of screening. To minimize this problem, these costs need to be weighed against the benefits of identifying students considered to be at high risk for suicide along with those who are not deemed to be at high risk for suicide but who do have an

**Objectives.** We examined the effects of a scoring algorithm change on the burden and sensitivity of a screen for adolescent suicide risk.

**Methods.** The Columbia Suicide Screen was used to screen 641 high school students for high suicide risk (recent ideation or lifetime attempt and depression, or anxiety, or substance use), determined by subsequent blind assessment with the Diagnostic Interview Schedule for Children. We compared the accuracy of different screen algorithms in identifying high-risk cases.

**Results.** A screen algorithm comprising recent ideation or lifetime attempt or depression, anxiety, or substance-use problems set at moderate-severity level classed 35% of students as positive and identified 96% of high-risk students. Increasing the algorithm's threshold reduced the proportion identified to 24% and identified 92% of high-risk cases. Asking only about recent suicidal ideation or lifetime suicide attempt identified 17% of the students and 89% of high-risk cases. The proportion of nonsuicidal diagnosis-bearing students found with the 3 algorithms was 62%, 34%, and 12%, respectively.

**Conclusions.** The Columbia Suicide Screen threshold can be altered to reduce the screen-positive population, saving costs and time while identifying almost all students at high risk for suicide. (*Am J Public Health.* 2010;100:1648-1652. doi:10.2105/AJPH.2009.175224)

undiagnosed but significant, impairing, and treatable mental health condition. Second-stage evaluations that fail to confirm the need for clinical referral are therefore a necessary but at times onerous burden.

We report, for the first time, how varying the items and threshold of the items that determine whether an adolescent screens positive affects the accuracy and the program burden of the CSS. Our research questions were: (1) What effect does altering the scoring algorithm of the CSS have on identifying adolescents at high risk for suicide? and (2) What effect does altering the scoring algorithm of the CSS have on reducing the burden of confirmatory evaluations for a screening setting?

## METHODS

We established a sample frame of 2583 9th-grade through 12th-grade students attending 7 ethnically and ability-diverse high schools in the New York Metropolitan Area from 1991 to 1994. Parents were contacted by letters endorsed by the school principal asking them to

return the letter only if they preferred that their child not participate in the Board of Education- and school-approved screening program; 182 (7%) parents refused consent, 332 students (12.9%) whose parent had passively consented declined to participate, and 340 (13.2%) students were absent during screening, leaving 1729 (67%) who completed the CSS during a regular class period. Consent procedures—active versus passive consent—have been shown to impact participation rates, with active consent having a lower overall participation rate and a decline in subgroups of students shown to be at high risk for depression.<sup>12</sup> Of the participating students, the mean age was 15.4 years (SD=1.4; range=11-19), 57% were female, 56% were White, 18% were African American, 13% were Hispanic, and 13% were another race/ethnicity. Participants were spread evenly across 9th through 12th grades. A full description of the sample and measures are reported elsewhere.<sup>7</sup>

## Measures

Screening was conducted with the CSS,<sup>7</sup> a 43-item self-completion form that includes 32

general health questions and 11 items that address known risk factors for adolescent suicide, including a lifetime history of a suicide attempt and suicidal ideation within the past 3 months (both responded to with yes or no). Five global questions indicated possible but impairing depression, anxiety, and substance abuse within the past 3 months. Questions were prefaced by: “How much of a problem have you had with . . .” with responses on a 1 to 5 (no problem to very bad problem) Likert scale. Specific conditions were: feeling unhappy or sad, feeling nervous or worried, being worried that you are very cut off from other people or are withdrawing more and more into yourself, being worried that your feelings get too easily hurt or that you are losing your temper a lot, that you are often grouchy and that even little things seem to make you mad or upset, and using drugs or alcohol or both. If a question was endorsed as a bad or very bad problem, further questions were asked about whether the student would like help or was already receiving professional help for the problem.

Diagnostic criteria against which the screen was compared were derived from the youth version of the Diagnostic Interview Schedule for Children version 2.3 (DISC).<sup>13</sup> The DISC is a highly structured diagnostic interview assessing *Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition (DSM-III-R)* diagnoses in the past 6 months.<sup>13–15</sup> Eight diagnostic modules were administered (major depression or dysthymia including questions about suicidal ideation and prior attempt, alcohol, marijuana, other substances, agoraphobia, panic disorder, social phobia, and generalized anxiety disorder). A DISC diagnosis required impairment.

### Procedure

Endorsement of the following CSS items were used to determine the need for a confirmatory assessment: (1) suicidal ideation in the past 3 months, or (2) a lifetime history of a suicide attempt, or (3) 3 or more of 5 symptoms (i.e., unhappiness or sadness, anxiety, social withdrawal, irritability, or substance use) rated as “bad” or “very bad,” or (4) indicating a wish to speak to a professional for any problem on the CSS that was rated “bad” or “very bad.” Of the 1729 screened students, 489 (28.3%) met these criteria. The remainder

(n=1240 of 1729) were classified as screen-negative.<sup>6</sup>

*Diagnostic or confirmatory assessment.* All screen-positive students and a subgroup of screen-negative students, frequency-matched by using the grade-by-gender-by-ethnic-group distribution of the screen-positive students, were approached for diagnostic confirmatory assessments. Seventy-three percent (n=356 of 489) of all screen-positive (participation was incomplete as a result of absenteeism and logistical constraints of time available) and 23% (n=285 of 1240) of the screen-negative students were ultimately interviewed on the DISC 2.3. After the student had completed the CSS screen, questionnaires were scored. If the student screened positive, the DISC assessment was carried out—within 1 or 2 days if the student indicated suicidal thoughts or prior attempt, or within 1 to 14 days of the screening for the remainder. With the exception of those who endorsed suicidality, screen-positive students and screen-negative students were spaced throughout the evaluation phase, based largely on logistical arrangements (e.g., class schedules and available time).

Because screen-positive students were oversampled for participation in the DISC interview, screen-negative students were matched to the gender, ethnicity, and grade distribution of the screen-positive students, and the sampling procedures varied by school site. The sample of 641 who completed the DISC interview was weighted with the sampling fractions in each school-by-gender-by-screen-status stratum of the original screened sample.

*Reference standard criteria.* Two reference standards were defined with the DISC 2.3. High-risk for suicide was defined as ideation (past 6 months) or lifetime attempt in the presence of a current (past 6 months) *DSM-III-R* mood, anxiety, or substance-use disorder with impairment. Clinically significant mental health problem was defined as any ideation (past 6 months), lifetime attempt, or current (past 6 months) *DSM-III-R* mood, anxiety, or substance use diagnosis with impairment.

*Construction of screening algorithms.* Responses to CSS questions about suicidal ideation and attempt were merged into a single scale with higher values given to attempts over ideation and to recent ideation and attempts within the past 3 months over more distal

occurrences. An emotional symptom scale was constructed by summing the number of symptom items (i.e., unhappiness or sadness, irritability, social withdrawal, anxiety, and substance use) endorsed as greater than or equal to 3 (i.e., a medium, bad, or very bad problem). These 2 components were combined with Boolean operators (and/or) to create the screening algorithms.

Two considerations informed our selection of the algorithms’ thresholds. One priority was to minimize program burden—defined as the resources dedicated to the confirmatory assessment for screen-positive students. For this reason, we decided the ratio of screen-positive to total number screened should not exceed 35% in order to not overtax the resources of screening sites. We deemed this proportion appropriate on the basis of our experiences with screening. The second priority was to miss as few students with mental health problems as possible. Thus, we decided that the sensitivity of the screen for capturing students at high risk for suicide should be no less than 75% and the sensitivity for those with clinically significant psychological conditions should be at least 65%.

Three algorithms met these standards. One of these required recent (past 3 months) suicidal ideation and a lifetime prior attempt; however, students who had made a prior attempt could be screened negative if they did not report recent suicidal ideation. We decided that ignoring a student’s endorsement of a prior attempt would be unacceptable for ethical reasons. The 2 other CSS algorithms (Table 1) that met the criteria were felt to be clinically realistic, and practical for school settings:

- High-threshold CSS algorithm (algorithm A) requires any recent suicidal ideation, or prior lifetime suicide attempt, or 4 or more emotional items rated as a medium, bad, or very bad problem.
- Low-threshold CSS algorithm (algorithm B) requires any recent suicidal ideation, or prior lifetime suicide attempt, or 3 or more emotional items rated as a medium, bad, or very bad problem.

Although it did not meet our 2 considerations for algorithm selection, for contrast, we also examined a third algorithm (algorithm C)

**TABLE 1—Comparison of Columbia Suicide Screen (CSS) Algorithms and Diagnostic Outcome Criteria for Screening High School Students for High Suicide Risk and Any Significant Clinical Criteria DISC 2.3, New York, NY, 1991–1994**

CSS Algorithm	No. of Students	High Risk for Suicide Criterion <sup>a</sup>			Clinical Criterion <sup>b,c</sup>			
		Sensitivity, Weighted Proportion (95% CI)	Specificity, Weighted Proportion (95% CI)	PPV, Weighted Proportion (95% CI)	Sensitivity, Weighted Proportion (95% CI)	Specificity, Weighted Proportion (95% CI)	PPV, Weighted Proportion (95% CI)	
High threshold (algorithm A), n = 280 <sup>d,e</sup>	78	92.0 (79.1, 97.5)	78.2 (74.5, 81.4)	25.0 (18.8, 32.4)	197	68.8 (61.4, 75.8)	88.4 (85.0, 91.1)	68.6 (60.9, 75.4)
Low threshold (algorithm B), n = 355 <sup>e,f</sup>	81	96.1 (84.5, 99.4)	61.0 (56.9, 65.0)	16.3 (12.2, 21.4)	216	81.5 (74.7, 86.9)	70.9 (66.5, 75.0)	50.7 (44.6, 56.8)
Any lifetime attempt or recent ideation (algorithm C), n = 219 <sup>e</sup>	75	89.1 (75.6, 95.9)	87.1 (84.0, 89.7)	35.7 (27.1, 45.2)	178	58.1 (50.0, 65.5)	95.6 (93.2, 97.2)	82.6 (74.1, 88.8)
High threshold or ask for help (algorithm D), n = 344 <sup>e</sup>	83	99.2 (88.9, 100.0)	72.1 (68.2, 75.6)	22.4 (16.9, 28.9)	212	76.5 (69.0, 82.7)	82.5 (78.5, 85.8)	61.2 (54.1, 68.0)

Note. DISC = Diagnostic Interview Schedule for Children; PPV = positive predictive value. The sample size of students who completed the DISC 2.3 Evaluation Outcome Criteria was n = 641. The sample was weighted to represent the original 1729 screened students.  
<sup>a</sup>The sample size was n = 85. There were 20 male students and 65 female students.  
<sup>b</sup>The sample size was n = 254. There were 91 male students and 163 female students.  
<sup>c</sup>Any suicidal ideation (past 6 months), or lifetime suicide attempt, or current (past 6 months) *Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition*, mood, anxiety, or substance use disorder.  
<sup>d</sup>Any recent ideation or lifetime attempt, or 4 or more emotional items rated as a "medium," "bad" or "very bad" problem.  
<sup>e</sup>This number represents the unweighted frequency of students who met criteria for this algorithm and who completed the DISC interview. Unweighted frequencies and weighted proportions are presented.  
<sup>f</sup>Any recent ideation or lifetime attempt, or 3 or more emotional items rated as a "medium," "bad" or "very bad" problem.

that required only recent suicidal ideation or a lifetime prior attempt. Finally, we examined the impact of including “asking for help” as an indicator of positive status in the high-threshold algorithm (algorithm D).

**Statistical Analyses**

For each algorithm, we used the 2 DISC-based reference criteria to calculate sensitivity (i.e., the proportion of students with the outcome criterion who were screened positive), specificity (i.e., the proportion of students without the outcome criterion who were screened negative), and positive predictive value (i.e., the proportion of screen-positive students who met the outcome criterion). SPSS version 16 (SPSS, Chicago, IL) was used in all analyses.

**RESULTS**

Table 1 shows the sensitivity, specificity, and positive predictive value for the high-threshold (algorithm A) and low-threshold (algorithm B) CSS algorithms, as well as suicidal ideation or prior attempt algorithm (algorithm C) and an

algorithm combining the high-threshold criteria or endorsement of asking for help (algorithm D), against the 2 reference standard criteria from the DISC: (1) high risk for suicide (n=85; male=20; female=65) and (2) any clinically significant mental health conditions (n=254; male=91; female=163).

**Comparison of High- and Low-Threshold Algorithms**

As shown in Table 1, when high suicide risk was the criterion, both high- and low-threshold algorithms were highly sensitive. The additional benefit of the low-threshold algorithm was modest (96.1% vs 92.0%). Specificity was reduced from 78.2% to 61.0% and positive predictive value was reduced from 25.0% to only 16.3%. However, the high-threshold algorithm identified only 24% of the original 1729 screened students (not shown) as positive compared with the low-threshold algorithm, which identified 35% of the original 1729 screened students (not shown). Thus, when the goal was to identify students at high risk for suicide, the high-threshold algorithm seemed to be indicated,

particularly when there was a strong motivation to avoid false-positives.

The pattern was somewhat different with respect to identifying clinical conditions. Although even the least-specific algorithm identified few more than half of students with ideation, or attempt, or mood disorders, or anxiety disorders, or substance use disorders, the proportions were greatly increased when the low threshold was used. Furthermore, Table 2 shows the sensitivities for the 4 algorithms when one is identifying specific mental health conditions such as suicidal ideation, suicide attempts, and those who reported a diagnosis of mood, anxiety, or substance-use disorder with and without any suicidal ideation or attempts. The 4 algorithms performed similarly when one was identifying students who reported suicidal ideation, students who reported prior attempts, and students who reported suicidal ideation or attempt as well as met criteria for a *DSM-III-R* mood, anxiety, or substance use disorder. However, the low-threshold algorithm had the highest sensitivity (62.1%), identifying the greatest proportion of nonsuicidal

**TABLE 2—Identification Sensitivity of Specific Mental Health Conditions Among High School Students, by CSS Algorithm Used: New York, NY, 1991–1994**

Mental Health Conditions	High Threshold (Algorithm A), %	Low Threshold (Algorithm B), %	Suicidal (Algorithm C), %	High Threshold or Ask for Help (Algorithm D), %
Any recent ideation (n = 145)	88.8	89.8	86.4	91.8
Prior attempt (n = 79)	81.5	87.5	81.5	84.6
Ideation or attempt and anxiety, substance use, or mood disorder (n = 85)	92.0	96.1	89.2	76.5
Anxiety, substance use, or mood disorder; no ideation or attempt (n = 79)	33.6	62.1	11.7	47.4

Note. CSS=Columbia Suicide Screen. Unweighted frequencies and weighted proportions are presented. Disorder determined by Diagnostic Interview Schedule for Children version 2.3, *Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition* with impairment.<sup>13-15</sup>

students who were experiencing an impairing mood, anxiety, or substance use disorder, compared with the other algorithms.

**Effect of Screening Only for Recent Suicidal Ideation or a Prior Attempt**

The bottom row of Table 1 displays the performance of the simplest of screens: endorsement of any lifetime suicide attempt or recent suicide ideation. When high suicide risk was the criterion, this screen performed nearly as well as the high-threshold algorithm in terms of sensitivity (89.1% vs 92.0%), and its specificity and positive predictive value were substantially better. The program burden imposed by this screen—i.e., the rate of screen positives—was only 17.2% (n=298; unweighted frequency of students identified out of 1729 students originally screened). When any significant mental health condition was the criterion, however, the sensitivity of this simple screen was very poor (58.1%; Table 1).

**An “Asking for Help” Algorithm**

As originally designed, the CSS included an additional rating option for each question if it was rated as a bad or very bad problem: “Are you so concerned about this that you would like to speak to a professional?” All students who indicated that they wanted to speak to a professional were approached for a second-stage evaluation. Including the “asking for help” item in the definition of screen-positive (high-threshold or ask for help) improved sensitivity for students at high risk for suicide from 92% to 99.2% (5 additional cases identified) and identified an additional 15 clinically significant cases,

improving sensitivity for clinical conditions from 69% to 77% (Table 1). Nonetheless, the improved sensitivity came at the cost of an increase in program burden from 24% to 28% requiring evaluation of an additional 87 students, with an associated increase in unnecessary confirmatory evaluations to 59 and 49 for high-risk and clinical conditions, respectively.

**DISCUSSION**

Varying the scoring algorithms for the CSS during school-based screening for suicide risk had a major effect on the proportion of students rated positively (i.e., who would require further evaluation to determine clinical need, if any). However, even algorithms that reduced the proportion of screen-positive students by half had little effect on the identification of students at high risk for suicide, defined as students currently experiencing suicidal ideation or those who had made a previous suicide attempt and who currently met criteria for a mood, anxiety, or substance-abuse diagnosis with impairment (all of which are risk factors for suicide).<sup>5,6</sup> The main cost of up-rating the algorithm’s threshold was in the smaller proportion of students who had a diagnosis without suicidality that would be identified; 66% of these students would be missed with the high-threshold algorithm.

If program goals were to maximize identification of anxiety, mood, and substance-use disorders, disorders that might cause considerable social and academic impairment and that frequently go unidentified and untreated in adolescents, then the low-threshold algorithm is indicated.

The high-threshold algorithm was in some respects the least satisfactory of the algorithms. It was only marginally less sensitive to high suicide risk than asking 2 questions about past suicide attempts and current suicidal ideation. Although it identified more of the clinically significant conditions, its sensitivity remained lower than that of the low-threshold algorithm.

There are several acceptable approaches to identify students at risk for emotional and behavioral problems. Multiple-gate screening—that is, having multiple ways, such as combining teacher nomination, screening, and record review to identify students with emotional and behavioral problems—is a commonly accepted approach in preschool<sup>16-18</sup> and, more recently, in secondary-school<sup>19</sup> students. Two-stage screening procedures (i.e., universal screening with a follow-up rule-out procedure) are an important tool for identifying individuals in the community who may be in need of treatment. The first-stage screening instrument, however, must be designed with 2 conflicting goals in mind; because only positive screens will be followed with a diagnostic interview, it is important to miss as few individuals with the target condition as possible while not overburdening the screening site with secondary interviews that will not capture true cases. It appears that this can be achieved by setting the threshold of the algorithm to the appropriate level. By evaluating the cost and benefits of different CSS algorithm thresholds we hope to have provided institutions with some guidance regarding the point at which, for them, the best balance of these conflicting goals lies.



The generalizability of our findings may be limited. The CSS algorithms were created posthoc of data collection; however, criteria for their inclusion was set a priori (i.e., the maximum number of secondary evaluations allowed was 35%). This limit in allowable secondary evaluations was a practical consideration made in an effort to reduce the burden on screening sites; however, it would have also limited the algorithm's sensitivity.

Screening for suicide risk is an important public health intervention to identify both students at risk for suicide and students with clinical mental health disorders, such as depression, anxiety, and substance use, that would otherwise go undetected. We have demonstrated that altering the thresholds of the CSS does not dramatically affect the sensitivity of the screen on the one hand, but can greatly reduce its cost and burden on the other hand. Furthermore, most students identified during screening for suicide risk are experiencing psychiatric conditions that, by themselves, can cause significant impairment in school and social functioning and therefore merit recognition and treatment. A growing body of research<sup>20–22</sup> has demonstrated that these conditions can be successfully treated, resulting in an improved quality of life. Furthermore, because these psychiatric conditions are potential precursors of suicide, early identification of these problems is also advisable from a suicide prevention perspective. ■

About the Authors

At the time of the study, Michelle Scott, Yanling Huo, Blake Turner, Prudence Fisher, and David Shaffer were with the Department of Child Psychiatry, College of Physicians and Surgeons, Columbia University, New York, NY. Holly Wilcox was with the Johns Hopkins School of Medicine, Baltimore, MD.

Correspondence should be sent to Michelle Scott, School of Social Work, Monmouth University, 400 Cedar Ave, West Long Branch, NJ 07764-1898 (e-mail: mscott@monmouth.edu). Reprints can be ordered at <http://ajph.org> by clicking the "Reprints/Eprints" link.

This article was accepted November 25, 2009.

Contributors

M. Scott co-conceptualized the article topic, directed analyses, and led the writing. H. Wilcox assisted with data collection and edits of article. Y. Huo conducted data analysis. J.B. Turner supervised data analysis, assisted with interpretation of data, and participated in writing and editing the article. P. Fisher assisted with data collection, interpretation of data, and editing the

article. D. Shaffer originated the study and its design, co-conceptualized the article topic, and participated in writing and editing the article.

Acknowledgments

This study was funded by the Centers for Disease Control and Prevention (grant R49/CCR 202598), the National Institute of Mental Health (grants P30MH-43878 and T32MH-16434), and the Carmel Hill Fund at Columbia University.

Human Participant Protection

The screening materials and consent procedure were approved by the New York State Psychiatric Institute's institutional review board, the Review Board of the New York City Board of Education, and the Review Board of the Archdiocese of New York.

References

1. Velez CN, Cohen P. Suicidal behavior and ideation in a community sample of children: maternal and youth reports. *J Am Acad Child Adolesc Psychiatry.* 1988;27(3):349–356.
2. Young TL, Zimmerman R. Clueless: parental knowledge of risk behaviors of middle school students. *Arch Pediatr Adolesc Med.* 1998;152(11):1137–1139.
3. Loeber R, Green SM, Lahey BB. Mental health professionals' perceptions of the utility of children, mothers, and teachers as informants of child psychopathology. *J Clin Child Psychol.* 1990;19:136–143.
4. Mojtabai R, Olfson M. Parental detection of youth self-harm behavior. *Suicide Life Threat Behav.* 2008;38(1):60–73.
5. Shaffer D, Gould MS, Fisher P, et al. Psychiatric diagnosis in child and adolescent suicide. *Arch Gen Psychiatry.* 1996;53(4):339–348.
6. Brent DA, Baugher M, Bridge J, Chen T, Chiapetta L. Age- and sex-related risk factors for adolescent suicide. *J Am Acad Child Adolesc Psychiatry.* 1999;38(12):1497–1505.
7. Shaffer D, Scott M, Wilcox H, et al. The Columbia Suicide Screen: validity and reliability of a screen for youth suicide and depression. *J Am Acad Child Adolesc Psychiatry.* 2004;43(1):71–79.
8. Scott MA, Wilcox HC, Schonfeld IS, et al. School-based screening to identify at-risk students not already known to school professionals: the Columbia Suicide Screen. *Am J Public Health.* 2009;99(2):334–339.
9. Hallfors D, Brodish PH, Khatapoush S, Sanchez V, Cho H, Steckler A. Feasibility of screening adolescents for suicide risk in "real world" high school settings. *Am J Public Health.* 2006;96(2):282–287.
10. Gilbody S, Sheldon T, Wessely S. Should we screen for depression? *BMJ.* 2006;332(7548):1027–1030.
11. Lenzer J. Bush's plan to screen for mental health meets opposition in Illinois. *BMJ.* 2004;329(7474):1065
12. Chartier M, Vander Stoep A, McCauley E, Hertine JR, Tracy M, Lymp J. Passive versus active parental permission: implications for the ability of school-based depression screening to reach youth at risk. *J Sch Health.* 2008;78(3):157–164.
13. Shaffer D, Fisher P, Dulcan MK, et al. The NIMH Diagnostic Interview Schedule for Children, version 2.3

(DISC-2.3): description, acceptability, prevalence rates, and performance in the MECA Study (Methods for the Epidemiology of Child and Adolescent Mental Disorders Study). *J Am Acad Child Adolesc Psychiatry.* 1996;35(7):865–877.

14. Schwab-Stone ME, Shaffer D, Dulcan MK, Jensen PS, Fisher P, Bird HR. Criterion validity of the NIMH Diagnostic Interview Schedule for Children, version 2.3 (DISC-2.3). *J Am Acad Child Adolesc Psychiatry.* 1996;35(7):878–888.

15. *Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition.* Washington, DC: American Psychiatric Association; 1987.

16. Feil EG, Severson HH, Walker HM. Screening for emotional and behavioral delays: the Early Screening Project. *J Early Interv.* 1998;21(3):252–266.

17. Feil EG, Walker H, Severson H, Ball A. Proactive screenings for emotional and behavioral concerns in Head Start preschools: promising practices and challenges with applied research. *Behav Disord.* 2000;26(1):13–25.

18. Walker HM, Severson HH, Nicholson F, Kehle T, Jensen WR, Clark E. Replication of the Systematic Screening of Behavioral Disorders (SSBD) for the identification of at-risk children. *J Emot Behav Disord.* 1994;2:66–77.

19. Caldarella P, Young E, Richardson M, Young B, Young R. Validation of the Systematic Screening for Behavior Disorders in middle and junior high school. *J Emot Behav Disord.* 2008;16(2):105–117.

20. Kennard B, Silva S, Vitiello B, et al. Remission and residual symptoms after short-term treatment in the Treatment of Adolescents with Depression Study (TADS). *J Am Acad Child Adolesc Psychiatry.* 2006;45(12):1404–1411.

21. The TADS Team. The Treatment for Adolescents with Depression Study (TADS): long-term effectiveness and safety outcomes. *Arch Gen Psychiatry.* 2007;64(10):1132–1144.

22. Mufson L, Dorta KP, Wickramaratne P, Nomura Y, Olfson M, Weissman MM. A randomized effectiveness trial of interpersonal psychotherapy for depressed adolescents. *Arch Gen Psychiatry.* 2004;61(6):577–584.