

Shaffer, D., Scott, M., Wilcox, H., Maslow, C., Hicks, R., Lucas, C. P., 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 and Adolescent Psychiatry*, 43(1), 71-79.

keywords: suicide, adolescents, Columbia SuicideScreen, Beck Depression Inventory

According to the National Institute of Mental Health, suicide by youth ages 15-24 was the third leading cause of death among adolescents in 2002. It is also estimated that there may be 8-25 attempts for every adolescent suicide death. These statistics reveal the importance of identifying youth who are at risk of suicide and to provide them with treatment. This article reports on the validity and test-retest reliability of the Columbia SuicideScreen (CSS). The CSS is a brief, 11-item self-report instrument designed to assess known risk factors for suicide among adolescents. The instrument was found to have "good sensitivity and reasonable specificity" (p. 71). However, youth who screen positive on the CSS for being at risk of suicide should be evaluated with a second instrument in order to reduce the number of false-positives (i.e. students who are not at risk of suicide) identified by the CSS.

Participants were drawn from a convenience sample of 9th-12th grade students attending seven New York high schools ($N = 1,729$). Each grade represented about 25% of the sample. The mean age of the sample was 15.4 years ($SD = 1.4$), and over half of the students were female (57%). Caucasians comprised 56% of the sample, followed by African-Americans (18%), Hispanics (13%), and other ethnicities (13%).

All students completed the CSS and Beck Depression Inventory (BDI) during school hours. To avoid an emphasis on suicidality, items from the CSS and BDI were embedded within a larger health survey. With regard to the CSS, students were deemed at risk of suicide if they endorsed any of the following key risk items: (a) any suicidal ideation within the last three months; (b) any suicide attempt at any time; (c) three out of five emotional items (i.e., unhappy, withdrawal, irritability, anxiety, substance abuse); or (d) if they asked to speak to a professional about their emotional problems. Of the total sample, 28% ($n = 489$) of the students endorsed one of the items and therefore met criteria for being at risk of suicide.

In order to assess the validity of the CSS, 73% ($n = 356$) of students who were identified as being at risk of suicide were group matched (by grade, gender, and ethnicity) with 285 students who were not at risk. The Diagnostic Interview for Children (DISC-2.3) was then administered to each group, and students needing further evaluation or treatment were identified at this time. The DISC-2.3 for risk of suicide includes: suicidal ideation or prior attempt, and a DSM-III-R diagnosis for major depression, dysthymia, or substance abuse. The level of risk of suicide detected by the DISC-2.3 was used to establish the criterion validity of the CSS.

Results indicate that the CSS algorithm for *suicidal ideation* or *previous attempt*, along with intermediate or higher scores for *unhappy*, *withdrawal*, *irritability* or *anxiety*, provided the best concordance with the DISC-2.3 criterion. The most balanced algorithm had a sensitivity of 0.75, a specificity of 0.83, and a positive predictive value ranging from 16-33% (p. 75). Thus, in comparison to the BDI, the CSS was more sensitive to detecting adolescents who are at risk of suicide, and was more specific in its ability to reduce the number of false-positives. The CSS was also more sensitive than the BDI in predicting suicide as measured by the DISC-2.3 validity criterion.

The power of the CSS to detect adolescents who are at risk of suicide, along with their specific risk factors, may be due to the longer time frame (three months) covered by the CSS in comparison to the BDI (one week). Further, the CSS asks about both suicidal ideation and suicidal behaviors, and the BDI does not. However, like other suicide risk instruments, the CSS "has the potential for having high sensitivity at the

Continued...

expense of specificity.” (p. 75). Thus, even though the specificity of the CSS was “generally better than for other instruments,” (p. 75), the authors recommend a second-stage evaluation to increase specificity.

Test-retest reliability was determined with 85 students who were re-administered the CSS and BDI eight days after their first administration. Test-retest reliabilities were 0.48(k) for suicidal ideation, and 0.58(k) for suicide attempt. With the exception of ethnicity, this sample of students did not differ significantly from the overall sample. There were more Hispanic and African-American students in the test-re-test sample than in the original sample. Compared to the BDI, the CSS had slightly lower test-retest reliability, but had greater sensitivity and specificity against the DISC-2.3 validity criterion.

In conclusion, the CSS shows promise as an instrument that can detect suicidal ideation and other at-risk factors for suicide among teenagers as a first step toward identifying students who need evaluation or treatment. A second-stage evaluation should increase the ability of clinicians and researchers to accurately detect these teens. Further, the CSS can be administered in a school setting, which should increase the odds of identification and treatment of high school youth.