

STUDY FUNDING: The study was funded by Janssen Scientific Affairs LLC, Titusville, New Jersey.

DISCLOSURES: Dr. Sheehan is the author and copyright holder of the Sheehan-Suicide Tracking Scale (S-STs), the Sheehan-Suicidality Tracking Scale Clinically Meaningful Change Measure Version (S-STs CMCV), the Pediatric versions of the S-STs, the Sheehan Disability Scale (SDS), and the Suicidality Modifiers Scale, is a co-author of the Suicide Plan Tracking Scale (SPTS), and owns stock in Medical Outcomes Systems, which computerized the InterSePT Scale for Suicidal Thinking (ISST-Plus) and the S-STs used in this study; Dr. Alphas is employed by Janssen Scientific Affairs, Titusville, New Jersey, is a stockholder of Johnson & Johnson, and is a co-author and copyright holder of the ISST-Plus; Dr. Mao is an employee of Janssen Research & Development, LLC, Titusville, New Jersey, who provided funding for this study; Mr. Q. Li has no conflicts of interest relevant to the content of this article; Ms. May has no conflicts of interest relevant to the content of this article; Ms. Bruer has no conflicts of interest relevant to the content of this article; Dr. McCullumsmith has no conflicts of interest relevant to the content of this article; Mr. Gray is an employee of Medical Outcome Systems, which computerized the ISST-Plus and the S-STs used in this study; Mr. X. Li has no conflicts of interest relevant to the content of this study; Dr. Williamson is an employee and stockholder in Johnson and Johnson, Janssen Scientific Affairs's parent company; and Ms. Lee is an employee of Janssen Research & Development, LLC, Titusville, New Jersey, who provided funding for this study.

ADDRESS CORRESPONDENCE TO:

David V. Sheehan, MD, MBA; E-mail: dsheehan@health.usf.edu

KEY WORDS: Suicide scale, suicide assessment, suicide attempt, suicide, suicidal ideation, suicidal behavior, suicidality, suicide risk, concurrent validity, validation, C-SSRS, C-CASA, FDA 2012 Draft Guidance Document, ISST-Plus, S-STs

Comparative Validation of the S-STs, the ISST-Plus, and the C-SSRS for Assessing the Suicidal Thinking and Behavior FDA 2012 Suicidality Categories

by **DAVID V. SHEEHAN, MD, MBA; LARRY D. ALPHS, MD, PhD; LIAN MAO, PhD; QIN LI; ROBERTA S. MAY, MA; EMILY H. BRUER, M.Ed; CHERYL B. MCCULLUMSMITH, MD, PhD; CHRISTOPHER R. GRAY, BS; XIAOHUA LI; and DAVID J. WILLIAMSON, PhD**

Dr. Sheehan is Distinguished University Health Professor Emeritus, University of South Florida College of Medicine, Tampa, Florida; Dr. Alphas is with Janssen Medical Affairs, LLC, Titusville, New Jersey; Dr. Mao is with Janssen Research & Development, LLC, Titusville, New Jersey; Mr. Q. Li is Director, Statistical Programming, Regeneron Pharmaceuticals, Inc., Basking Ridge, New Jersey; Ms. May is with the University of Alabama, Birmingham, Alabama; Ms Bruer is with the University of Alabama, Department of Psychiatry and Behavioral Neurobiology, Birmingham, Alabama; Dr. McCullumsmith is with the University of Cincinnati Department of Psychiatry and Behavioral Neuroscience, Cincinnati, Ohio; Mr. Gray is with Medical Outcomes Systems, Jacksonville, Florida; Mr. X. Li is with St. Vincent East Hospital, St. Vincent Health System, Birmingham, Alabama; and Dr. Williamson is with the University of South Alabama College of Medicine, Departments of Psychiatry and Neurology, Birmingham, Alabama, and Janssen Medical Affairs, LLC, Titusville, New Jersey.

Innov Clin Neurosci. 2014;11(9–10):32–46

ABSTRACT

Objective: This exploratory study examines the concurrent validity for mapping symptoms of suicidal ideation, self-harm, and suicidal behavior as recorded on the InterSePT Scale for Suicidal Thinking-Plus, the Sheehan-Suicidality Tracking Scale (clinician- and patient-rated and reconciled patient/clinician

versions), and the Columbia–Suicide Severity Rating Scale to the 11 United States Food and Drug Administration–Classification Algorithm of Suicide Assessment (September 2012) categories.

Method: Forty subjects with varying degrees of suicidal ideation and behavior severity (from not present to extremely severe) were

recruited from inpatient, outpatient, and emergency room settings. Each patient was interviewed using all three scales (InterSePT Scale for Suicidal Thinking-Plus, the Sheehan-Suicidality Tracking Scale, and the Columbia–Suicide Severity Rating Scale) on the same day. The scales were administered in a random sequence by three independent raters who were blind to the ratings on the other scales.

Results: The Sheehan-Suicidality Tracking Scale and the InterSePT Scale for Suicidal Thinking-Plus show acceptable agreement with the Columbia–Suicide Severity Rating Scale in detecting the presence or absence of the 2012 Food and Drug Administration-Classification Algorithm of Suicide Assessment categories 1, 5, 6, 10, and 11 (passive ideation; active ideation with method, intent, and plan; completed suicide; preparatory actions; and self-injurious behavior) but not of categories 2, 3, and 4 (3 other active suicidal ideation combination categories) or to 8 and 9 (aborted and interrupted attempt). Despite the significant disagreement between the Columbia–Suicide Severity Rating Scale on the one side and the InterSePT Scale for Suicidal Thinking-Plus and the Sheehan-Suicidality Tracking Scale on the other in the ability to accurately map to the 2012 Food and Drug Administration-Classification Algorithm of Suicide Assessment categories on some items, there was close agreement between the InterSePT Scale for Suicidal Thinking-Plus and the Sheehan-Suicidality Tracking Scale on these categories.

Conclusion: The results of this exploratory study invite discussion and debate about the validity of the Columbia–Suicide Severity Rating Scale and its ability to accurately assess key active suicidal ideation categories, since it disagrees so much with the other two standardized scales that agree so closely with each other.

INTRODUCTION

What if two test instruments do not line up with a reference or gold

standard? Does this mean that the test instruments are flawed or could it be that the reference itself is limited and the test instruments tap into a more accurate reality? These are quandaries we address in this article describing our efforts to validate two suicidality assessment scales against the gold standard scale. The study was developed in the context of increasing calls for standardized instruments to assess suicidal ideation and behavior (SIB) in central nervous system (CNS) trials.

To understand the purposes of the study, it is helpful to give some background. In 2010, the United States Food and Drug Administration (FDA) in its “Draft Guidance for Industry on the Prospective Assessment of Occurrence of Suicidality in Clinical Trials” (FDA 2010 Draft Guidance) introduced and recommended the Columbia Classification Algorithm of Suicide Assessment (C-CASA), a classification algorithm developed at Columbia University that classifies various types of suicidal and non-suicidal events into nine categories.² In 2012, the FDA introduced a second, modified version of the C-CASA for public comment (referred to here as the FDA-CASA 2012).³ This classification system changed the number of FDA required categories to the 11 used in the Columbia–Suicide Severity Rating Scale (C–SSRS), a rating scale that is now widely used in clinical trials.⁴ In addition, the FDA 2012 Draft Guidance designated the C–SSRS as the standard for the assessment of SIB for United States regulatory data collection, stipulating further that while it would consider other suicidality assessment scales as alternatives to the C–SSRS, any such new scale must map to the new FDA-CASA 2012 categories and should provide validity and reliability study data comparing itself to the C–SSRS. These decisions were not without controversy. Some observers felt that the C–SSRS required more testing before it should be conferred gold

standard status.^{5,6} Concern has also been voiced that C–SSRS categories do not line up with the FDA-CASA 2012 categories when the titles, definitions, and probe questions are carefully scrutinized.^{7,8}

The phenomenology of SIB is complex, and alternative approaches to collecting information that address needs that cannot be fully addressed with the C–SSRS would be valuable.^{9,10}

Two alternative suicidality scales that do map to both the 2010 C-CASA and the 2012 FDA-CASA are the InterSePT Scale for Suicidal Thinking-Plus (ISST-Plus) and the Sheehan-Suicidality Tracking Scale (S-STs).^{11–13} The ISST-Plus is an iterative evolution of the original ISST scale used as the primary outcome measure in the InterSePT study that was the basis for the regulatory approval of clozapine as a treatment for suicidal behaviors in schizophrenia and schizoaffective disorder.¹⁴ The original ISST was developed long before the development of the C-CASA, and it did not map to C-CASA categories. Its authors modified it into the ISST-Plus to permit mapping to the C-CASA. Following publication of the FDA 2012 Draft Guidance, a mapping table was developed to show how the ISST-Plus could be mapped to the FDA-CASA 2012 categories. The current S-STs is an iterative evolution of the original 2009 S-STs scale.¹¹ It has also been adapted to map to the 2010 C-CASA and the FDA-CASA 2012.

STUDY OBJECTIVES

The primary objectives of the analyses in this exploratory study were to evaluate the concordance between the ISST-Plus and the C–SSRS and the concordances between each version of the S-STs (clinician version, patient version, and reconciled patient/clinician version) and the C–SSRS in mapping symptoms of suicidal ideation, self-harm, and suicidal behavior using FDA-CASA 2012 categories. Secondary objectives included

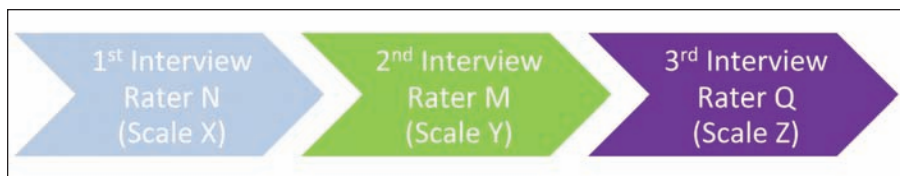


FIGURE 1. Study scale administration rate

1) assessing the comparative administration times of the three instruments; 2) examining the concordance between the ISST-Plus and the S-STS, and 3) describing combinations of suicidal ideation that were captured on the S-STS and the ISST-Plus in this study but not on the C-SSRS.

METHODS

Sample. Forty adult subjects identified as having SIB with varying degrees of severity across the full range of suicidality from “not present” to “extremely severe” were recruited from inpatient, outpatient, and emergency room settings. Since accidents may be suicide attempts in disguise and the FDA required adjudication for suicidality of all accidents that occurred in antidepressant trials, we recruited an additional five subjects who had been involved in recent accidents. None of these subjects were suicidal. Participation was voluntary. The study was approved by the Institutional Review Board of the University of Alabama at Birmingham, and all subjects gave informed consent before the study interviews took place.

Interviewers and training. Five raters, all qualified mental health professionals with experience working with suicidal patients, were used. The authors of each of the three scales trained and certified all of the raters and provided them with training slides and materials developed for each of the suicide assessment instruments (ISST-Plus, S-STS, and C-SSRS).

Study design and procedures. Before administering the three scales, the research team collected demographic information and

assessed each subject using a Clinical Global Impression Scale for Severity of Suicidality (CGI-SS) to ensure that the sample was balanced across the full range of SIB severity.

Scale administration. Each patient was interviewed and rated on each scale (ISST-Plus, S-STS, and C-SSRS) on the same day. The S-STS patient-rated scale was first completed by all patients directly into the laptop computer prior to the administration of any clinician-rated scales. The three clinician-rated scales (ISST-Plus, S-STS, and C-SSRS) were then administered in a predetermined random order sequence by three independent raters who were blind to the ratings on the other scales (Figure 1).

Direct data entry (electronic data capture) at the time of the visit was used to collect the ISST-Plus and S-STS data. This system precluded the possibility of missing values, double entries, legibility problems, and transcription errors. Since the C-SSRS did not have an equivalent direct entry format, all of the data for this scale were collected on the paper version and the categories endorsed were subsequently entered into the database for analysis.

The clinician completed the clinician-rated version of the S-STS blind to the prior patient-rated version. When the clinician saved the clinician ratings, the laptop computer immediately generated a version of S-STS displaying any discrepancies between the clinician and patient versions and asked both patient and clinician to continue the interview to come to agreement in reconciling any differences (S-STS reconciled version). This led to the generation of data on the three variants of the S-STS (clinician-rated, patient-rated,

and reconciled versions). It provided the study team with an opportunity to investigate the relative merits of each of these three possible approaches to assessing suicidality.

The time frame for SIB assessment for all three instruments was the past seven days. Video recordings of subjects were completed with separate consent.

Instrument mapping process. Figure 2 shows the FDA categories used for mapping. For completeness, we included the 11 categories in the FDA 2012 Draft Guidance and four additional categories from the original C-CASA and the FDA 2010 Draft Guidance that are of interest to regulatory agencies and are likely to occur in real-world experience with classification of suicide-like behavior (last 4 categories in Figure 2).

The authors of each of the scales (ISST-Plus, S-STS, and C-SSRS) provided detailed instructions on the algorithms they used to map item responses on their scales to the FDA-CASA 2012 categories. (Table 1). Mapping was performed, using these algorithms, with a computer-coded procedure for the S-STS and ISST-Plus and by a trained individual at the site for the C-SSRS.

Statistical analysis. We assessed baseline demographics and clinical characteristics using descriptive statistics. We used two approaches to evaluate agreement between the test instruments and the C-SSRS: Cohen’s Kappa and the area under the receiver operating characteristic curve (AUC). Cohen’s Kappa is a chance-corrected measure of agreement that ranges from 0 to 1.¹⁵ Shrout et al,¹⁶ after Fleiss,¹⁷ suggest that Kappa values greater than approximately 0.75 indicate excellent agreement beyond chance, values below approximately 0.40 represent poor agreement beyond chance, and values in between indicate fair to good agreement beyond chance. We report the Kappa values with 95-percent confidence intervals (CI).

While Kappa is often used as a measure of agreement, it is

dependent on prevalence and can be low even when there is high concordance on low-prevalence conditions. The AUC, interpreted as the probability that a randomly selected clinical case will score higher on the test than a non-case, has been proposed to correct this problem¹⁸ and can be used in situations in which the predictor is a dichotomy. In this case, the AUC equals $(SN+SP)/2$.¹⁹ Following Agresti,²⁰ we considered the AUC to be excellent evidence of concordance if 0.90 or greater, good evidence of concordance if between 0.80 and 0.90, acceptable although only average if between 0.70 and 0.80, and poor if below 0.70.²⁰

For each of the category comparisons with the C-SSRS, we report the following additional information shown in Table 2:^{21,22} absolute numbers of participants in each cell of the cross-tabulation (those positive for a category on the test instrument and the C-SSRS, those negative on both scales, those positive on the test instrument but negative on the C-SSRS, those negative on the test instrument but positive on the C-SSRS) and McNemar's test of the significance of the comparison. We also report sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). These calculations require a "gold standard" measure of the presence of a construct (e.g., illness, symptom). Given the current putative status of the C-SSRS per FDA publications, it was used as the gold standard criterion measure. Thus, sensitivity is defined here as the probability of the test (i.e., the ISST-Plus or the S-STS, respectively) identifying a patient as meeting criteria for a particular FDA category (using the C-SSRS as the criterion). Specificity is defined as the probability of the test concluding that the subject does *not* meet criteria for the FDA category (using the C-SSRS results as the criterion). The PPV is defined as the probability that a subject who is identified as meeting criteria for a

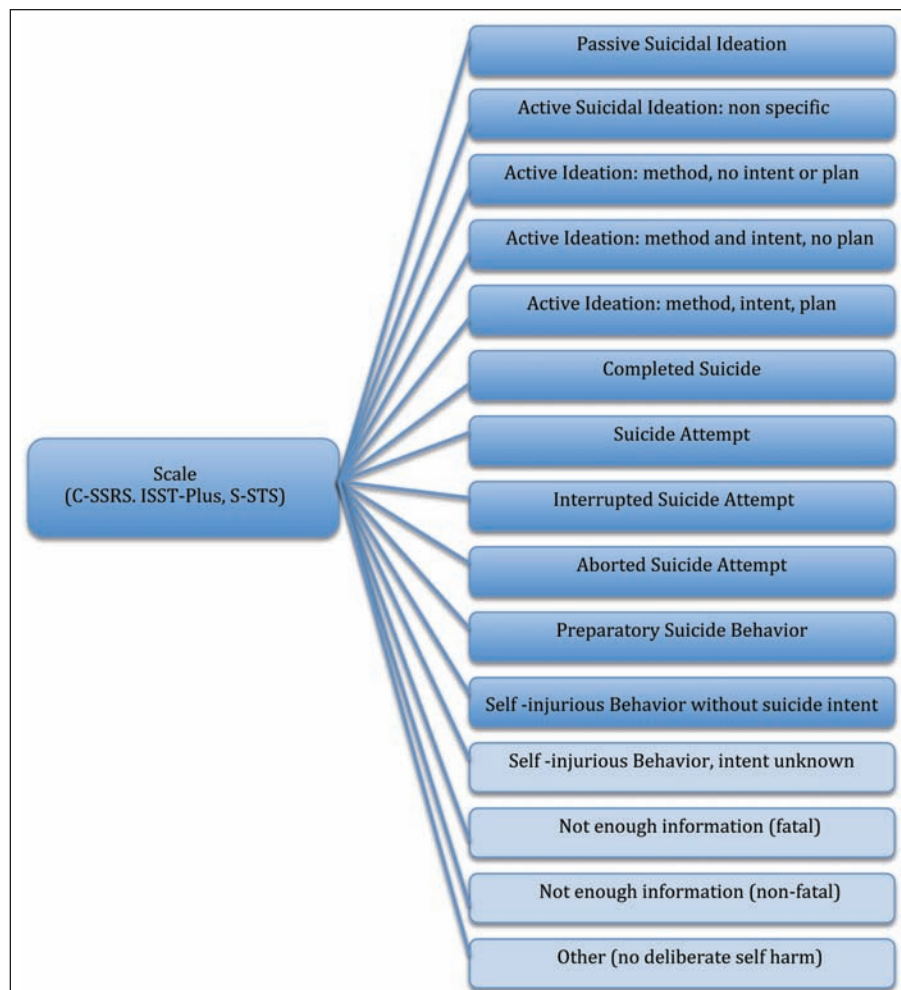


FIGURE 2. FDA categories to which suicide assessment scales should map per FDA Guidance 2012 and additional categories from FDA Guidance 2010

category on the test actually has the condition (using the C-SSRS as the criterion). In contrast, the NPV is the probability that a subject identified as not having the condition does not meet criteria based on the results of the C-SSRS.

We also report overall agreement between the clinician-rated versions of ISST-Plus and S-STS. In addition, we tabulate the combinations of active suicidal ideation that are captured on the ISST-Plus and S-STS but not on the C-SSRS.

RESULTS

Study subjects. Forty-five subjects were interviewed. The mean (standard deviation [SD]) age of the sample was 39.9 (15.0) years, with a range of 19 to 73 years. For gender, 44.4 percent of the subjects self-

identified as male. For race, 73.3 percent self-identified as white, 24 percent as black, and 2.2 percent as mixed race. The mean age of first psychiatric symptom was 16.6 (10.0) years with a range of 5 to 42 years. The mean age of first psychiatric treatment was 24.5 (13.3) years, with a range of 5 to 58 years. The mean age of first psychiatric hospitalization was 30.8 (14.7) years, with a range of 10 to 68 years. The mean number of past suicide attempts was 2.1 (2.8) with a range of 0 to 13. Suicide severity distribution is shown in Table 3. Table 3 illustrates that the sample studied had suicide severities well distributed across the full spectrum of severity.

Agreement between the (ISST-Plus and C-SSRS) and (S-STS and C-SSRS). Table 4, together

TABLE 1. Mapping algorithms (How the ISST-Plus, S-STs, and C-SSRS map to FDA-CASA 2012)

FDA-CASA 2012 CODE NUMBER	FDA-CASA 2012 CLASSIFICATION	ISST-PLUS MAP TO FDA-CASA 2012*	S-STs MAP TO FDA-CASA 2012	C-SSRS MAP TO FDA-CASA 2012
1	"Passive suicidal ideation"	Yes if: Part I, Item 1≥1	A positive response to 2	A "yes" response to item 1: "Wish to be Dead"
2	"Active suicidal ideation: nonspecific (no method, intent, or plan)"	Yes if: (Part I, Item 4≥1 OR Part I, Item 5≥1) AND Part I Items 10, 11, 12=0	A positive response to 3 AND a negative response to 4, 5, 6, and 7	A "yes" response to item 2: "Non-specific Active Suicidal Thoughts"
3	"Active suicidal ideation: method, but no intent or plan"	Yes if: (Part I, Item 4≥1 OR Part I, Item 5≥1) AND Part I, Item 10≥1; AND Part I, Items 11, 12=0	A positive response to 3 and 4 AND a negative response to 5, 6, and 7	A "yes" response to item 3: "Active Suicidal Ideation With Any Methods (no plan) Without Intent to Act"
4	"Active suicidal ideation: method and intent, but no plan"	Yes if: (Part I, Item 4≥1 OR Part I, Item 5≥1) AND Items 10 and 12≥1; AND Item 11=0	A positive response to 3 and 4 and (6 or 7) and a negative response to 5	A "yes" response to Item 4: "Active Suicidal Ideation With Some Intent to Act, Without Specific Plan"
5	"Active suicidal ideation: method, intent, and plan"	Yes if: (Part I, Item 4≥1 OR Part I, Item 5≥1) AND Part I, Items 10, 11, and 12≥1	A positive response to 3, 4, and 5 and (6 or 7)	A "yes" response to item 5: "Active Suicidal Ideation With Specific Plan and Intent"
6	"Completed suicide"	Yes if: Part III coded as 6	A "yes" response to 13	NA
7	"Suicide attempt"	Yes if: "Yes" on Part II, Item 1.2 OR "yes" on Part II, Item 1.3	A positive response to 10 or 16 OR a "yes" response to 1b	A "yes" response to "Actual Attempt"
8	"Interrupted suicide attempt"	Yes if: Part II, Item 2.3≥1	A positive response to 8 with at least one Level 3 on 12	A "yes" response to "Interrupted Attempt"
9	"Aborted suicide attempt"	Yes if: Part II, Item 2.2≥1	A positive response to 8 with at least one Level 2 on 12	A "yes" response to "Aborted Attempt"
10	"Preparatory acts toward imminent suicidal behavior"—"not counting aborted or interrupted attempts"	Yes if: Part II, Item 1.4=Yes	A positive response to 8 with at least one Level 1 on 12	A "yes" response to "Preparatory Acts or Behavior"
11	"Self-Injurious Behavior Without Suicidal Intent"	Yes if: Part II, Item 1.1=Yes; AND Part II, Item 1.2=No	A positive response to 9 OR a positive response to 1a and a negative response to 1b	A "yes" response to "Has subject engaged in Non-Suicidal Self-Injurious Behavior?"
12	"Self-injurious behavior, intent unknown"	Yes if: Part II, Item 1.1=Yes; AND Part II, Item 1.2=unknown/NA	A positive response to 1a with 1b and 6 and 7 and 10 unanswered	NA
13	"Not enough information (fatal)"	Yes if: Part III scored 8	A "yes" response to 14	NA
14	"Not enough information (nonfatal)"	Yes if: Part II, Item 1.0=Unknown/NA OR Part II, Item 1.1=Unknown/NA	Yes if question 17 is "yes" and FDA-CASA 2012 code numbers 1–13 are ALL "no"	NA
15	"Other (accidental, psychiatric medical), no deliberate self-harm"	Yes if: Part II, Item 1.0=Yes AND Part II, Item 1.1=No	(A positive or blank response to 1 and a negative response to questions 2, 3, 4, 5, 6, 7, 8, 9, 10, 13, and 14) OR a positive response to 15	NA

*For each separate event, only the most severe categorization was counted. All separate events were assessed.

ISST-Plus: InterSePT Scale for Suicidal Thinking-Plus; S-STs: Sheehan-Suicidality Tracking Scale; C-SSRS: Columbia-Suicide Severity Rating Scale; FDA-CASA 2012: United States Food and Drug Administration-Classification Algorithm of Suicide Assessment; NA: not applicable

with Figures 3 to 7, show the results of the tests of agreement between the ISST-Plus and the C-SSRS and the three versions of the S-STs for each FDA category with the exception of completed suicide (#6) for which no subjects met criteria. The five subjects who had been involved in a recent accident are not included in these results for two reasons. First, their injuries were such that they could not complete the patient-rated version of the S-STs. Second, none had any

suicidality and the C-SSRS does not capture this category. The C-SSRS was not designed to categorize such subjects, and thus cannot be used as a standard reference for them. On the ISST-Plus and S-STs, these five subjects mapped to category 15 (other injury or overdose, no suicide intent).

Agreement between the ISST-Plus and C-SSRS. AUC values were good to excellent (0.80–1.00) for "Passive ideation" (category #1), "Suicidal attempt" (#7), and

"Preparatory acts" (#10). The AUC was acceptable (0.70–0.80) for "Active suicidal ideation: method, intent, and plan" (#5) and "Self-Injurious Behavior Without Suicidal Intent" (#11). AUC values were poor (<0.70) for categories "Active suicidal ideation: nonspecific" (#2), "Active suicidal ideation: method, but no intent or plan" (#3), "Active suicidal ideation: method and intent, but no plan" (#4), "Interrupted suicide attempt" (#8), and "Aborted suicide attempt" (#9).

Kappa values were acceptable to high for categories 1, 5, 7, 10, and 11 but low for the other categories. Sensitivity was high (≥ 0.80) for “Passive ideation” (#1), “Suicide attempt” (#7), and “Preparatory acts” (#10). Sensitivity was acceptable (0.40–0.79) for “Active suicidal ideation: method, intent, and plan” (#5) and “Self-Injurious Behavior Without Suicidal Intent” (#11). Sensitivity was low for the remaining categories of active suicidal ideation (#2, #3, and #4) as well as for interrupted and aborted attempt categories (#8 and #9). In contrast, specificity was high (> 0.80) for every category on the ISST-Plus. NPV was high for all of the categories with the exception of the active ideation categories 2, 3, and 4. PPV values were high for categories 1, 3, 5, and 7 and acceptable for categories 10 and 11, but low for categories 4, 8, and 9.

Agreement between the S-STS and C-SSRS. AUC values were good to excellent (0.80–1.00) for all three versions of the S-STS for “Passive ideation” (#1), “Active suicidal ideation: method, intent, and plan” (#5), “Suicidal attempt” (#7), and “Self-Injurious Behavior Without Suicidal Intent” (#11). The AUC was acceptable (0.70–0.80) for “Preparatory acts” (#10) on the clinician version of the S-STS but not on the patient version or reconciled version. AUC values were poor (< 0.70) for categories “Active suicidal ideation: nonspecific” (#2), “Active suicidal ideation: method, but no intent or plan” (#3), “Active suicidal ideation: method and intent, but no plan” (#4), “Interrupted suicide attempt” (#8) and “Aborted suicide attempt” (#9).

Kappa values were acceptable to excellent for all of the categories on all three S-STS versions with the following exceptions: kappa values were low for the active suicidal ideation categories 2, 3, and 4 and for “Aborted attempt” (#9) on all three S-STS versions. Kappa was also low for “Interrupted attempt” (#8) and “Preparatory behavior” (#10) on

TABLE 2. Definitions of tests of agreement		Standard Reference (C-SSRS)	
		Present	Absent
Test Instrument	Present	A	B
	Absent	C	D

Sensitivity: When present according to the C-SSRS (standard reference scale), how often does the test instrument detect the condition? $A/(A + C)$

Specificity: When absent according to the C-SSRS (standard reference scale), how often does the test instrument say the condition is absent? $D/(D + B)$

Positive Predictive Value (PPV): When present according to the test instrument, how often does the C-SSRS (standard reference scale) agree? $A/(A + B)$

Negative Predictive Value (NPV): When absent according to the test instrument, how often does the C-SSRS (standard reference scale) agree? $D/(C + D)$

Area Under the Receiver Operating Characteristic Curve (AUC): What is the probability that a randomly selected patient with the trait is more likely to be categorized as having the trait relative to a randomly selected subject without the trait? When both the predictor and the criterion variables are binary, this is calculated by $(\text{Sensitivity} + \text{Specificity})/2$. ROC curve analysis is used to evaluate the precision of the best-fitting prediction equations of each scale relative to the C-SSRS and may be the best overview assessment of the data (see references 21 and 22).

C-SSRS: Columbia–Suicide Severity Rating Scale

the clinician version. Sensitivity was high (≥ 0.80) for “Passive ideation” (#1), “Active suicidal ideation: method, intent, and plan” (#5), and “Self-Injurious Behavior Without Suicidal Intent” (#11) on all three versions. It was also high for “Suicide attempt” (#7) on the patient and reconciled versions. Sensitivity was acceptable (0.60 on the patient version and 0.40 on the clinician version) for “Preparatory acts” (#10). Sensitivity was low for the remaining categories of active suicidal ideation (#2, #3, and #4) as well as for the interrupted and aborted attempt categories (#8 and #9) on all three versions.

Specificity, on the other hand, was high (> 0.80) for the active ideation categories (#2, #3, and #4), for the clinician version of “Active ideation” (#5), as well as for the interrupted and aborted attempt categories (#8 and #9) and for “Suicide attempt” (#7). It was acceptable but somewhat lower (0.70–0.80) for “Passive ideation” (#1) on all three

versions. This result is undoubtedly a function of the S-STS detecting more passive ideation than the C-SSRS, since it asks about different subtypes of passive suicidal ideation than does the C-SSRS.²³ Taken together, these results indicate that the S-STS and the C-SSRS were in general agreement as far as ruling out most of the categories of active ideation, but not on ruling in categories 2, 3, and 4. NPVs were acceptable to high for all of the categories with the exception of the two active ideation categories: “Active suicidal ideation: nonspecific” (#2) and “Active suicidal ideation: method, but no intent or plan” (#3). PPVs were high on patient-rated version for categories 1, 4, 7, 8, and 9, on the clinician and reconciled versions for categories 1, 3, 7, and 11, but low for categories 2 and 3.

The patient-rated S-STS was quite discrepant from the C-SSRS in mapping categories 2, 3, and 4 and to a lesser extent in mapping categories 8 and 9. The S-STS reconciled scale

TABLE 3. Suicide severity in subjects (N=45)

SEVERITY OF SUICIDAL IDEATION AND BEHAVIORS (CGI-SS)	SUBJECTS N (%)
Not suicidal, recent accident	5 (11.1)
Not suicidal, but self-harm (on the border of being suicidal)	7 (15.6)
Questionably suicidal	5 (11.1)
Mildly suicidal	6 (13.3)
Moderately suicidal	6 (13.3)
Severely suicidal	5 (11.1)
Extremely suicidal	6 (13.3)
Recently attempted suicide	5 (11.1)

CGI-SS: Clinical Global Impression Scale for Severity of Suicidality

tended to be closer to the original S-STS clinician-rated scale, but this was not consistently so across all categories.

Agreement between S-STS and ISST-Plus. Figure 8 shows the agreement for the comparison between the clinician-rated versions of the ISST-Plus and the S-STS on the FDA-CASA 2012 categories (blue columns) for which data could be collected. Raw concordance was 80 percent or higher for all of the categories. This concordance was surprising since the two scales were developed independently, have very different source origins, and at face inspection appear different from each other in approach, lines of

questioning, and format used to elicit information.

Agreement between the S-STS patient-rated, clinician-rated, and the reconciled scale versions. The results shown in Table 4 and Figures 3 to 7 suggest that the patient-rated version of the S-STS performed similarly to the clinician-rated S-STS with a few exceptions. The frequencies of “Active suicidal ideation: method, intent and plan” (#5) and of “Preparatory acts” (#10) were also somewhat higher on the patient-rated version (24 subjects vs. 18 and 10 vs. 6, respectively).

Almost twice as many subjects (17 vs. 8) endorsed “Self-Injurious Behavior Without Suicidal Intent” (#11) on the patient-rated version compared to the clinician-rated version. Since some patients who had suicidal ideation in the past week may have engaged in non-suicidal self injury (e.g., to masochistically relieve tension) and this injury may have led to their emergency room visit, it cannot be concluded that the clinician made the correct assessment and that the patient was hiding the truth. In other words, the discrepancy between ratings of the patient and clinician on this item can be interpreted in opposite ways. Both may have been correct some of the time.

This finding lends support to the idea that the patient-rated S-STS may not be any less valid an approach than the clinician-rated S-STS (or even the “reconciled”) version of the S-STS. This finding needs to be replicated in much larger samples and in different clinical and cultural settings.

Suicidality combinations captured by ISST-Plus and S-STS but not by the C-SSRS. Sixty-seven percent of the patients had combinations of suicidal ideation phenomena (ideation, method, intent, plan) that did not fit any of the existing FDA-CASA 2012 or C-SSRS categories. As described in greater detail in two other articles,^{7,8} the FDA-CASA 2012 and the C-SSRS only cover five of the 16 possible

combinations of active ideation, method, intent, and plan and only six of 32 possible combinations of passive suicidal ideation, active suicidal ideation, method, intent, and plan. In contrast, the S-STS and the ISST-Plus both capture all these possible combinations.

Tables 5 and 6 display the combinations captured by the S-STS, but not captured by the C-SSRS, and the combinations captured by both. Table 5 displays the data using the clinician-rated S-STS for all 45 suicidal and non-suicidal patients while Table 6 shows the data using the patient-rated S-STS for only those 40 patients who had SIB. Among the 45 suicidal and non-suicidal patients, 67 percent had combinations of suicidal ideation, method, intent, or plan detected by the clinician-rated S-STS, and 76 percent had combinations detected by the clinician-rated ISST-Plus, for which no unique combination category exists on the C-SSRS. Among the 40 suicidal subjects, 80 percent had combinations of suicidal ideation, method, intent, or plan detected by the patient-rated S-STS, for which no unique combination category exists on the C-SSRS. Examples include the combination of active ideation, with intent, but with no method or plan.

Duration of interviews. The computer clock tracked the duration of the ISST-Plus, S-STS patient-rated, S-STS clinician-rated, and the S-STS reconciled version. The taped videos were used to calculate the duration of the C-SSRS (since the C-SSRS was only done on paper). Table 7 shows the duration of the 34 interviews (out of a total of 45) for which there was comparable data across all three scales and their variants. Videos were not done for five study participants, an additional two were unable to complete the patient-rated version because of injuries, one was unable to complete the reconciliation because of the subject’s emotional state, and three of the remaining 37 participants (8.1%) did not need a reconciliation

TABLE 4. Percent agreement, n (%), kappa, and AUC for ISST-Plus vs. C-SSRS and S-STSS vs. C-SSRS		ISST-PLUS (N = 40)		S-STSS (PATIENT) (N = 40)		S-STSS (CLINICIAN) (N = 40)		S-STSS (RECONCILED) (N = 39)	
		No	Yes	No	Yes	No	Yes	No	Yes
FDA-CASA 2012 Category 1: Passive suicidal ideation									
C-SSRS	No	4 (10.0)	0	3 (7.5)	1 (2.5)	3 (7.5)	1 (2.5)	3 (7.7)	1 (2.6)
	Yes	1 (2.5)	35 (87.5)	1 (2.5)	35 (87.5)	2 (5.0)	34 (85.0)	2 (5.1)	33 (84.6)
% agreement*		39/40 (97.5)		38/40 (95.0)		37/40 (92.5)		36/39 (92.3)	
P value**		0.317		1.000		0.564		0.564	
Kappa (95% CL)		0.875 (0.635, 1.000)		0.722 (0.357, 1.000)		0.625 (0.237, 1.000)		0.624 (0.235, 1.000)	
AUC		0.986		0.861		0.847		0.846	
FDA-CASA 2012 Category 2: Active suicidal ideation: nonspecific									
C-SSRS	No	10 (25.0)	2 (5.0)	12 (30.0)	0	10 (25.0)	2 (5.0)	11 (28.2)	1 (2.6)
	Yes	26 (65.0)	2 (5.0)	28 (70.0)	0	28 (70.0)	0	27 (69.2)	0
% Agreement*		12/40 (30.0)		12/40 (30.0)		10/40 (25.0)		11/39 (28.2)	
P value**		<0.001		-		<0.001		<0.001	
Kappa (95% CL)		-0.061 (-0.212, 0.090)		-		-0.103 (-0.246, 0.040)		-0.052 (-0.154, 0.050)	
AUC		0.452		0.500		0.417		0.458	
FDA-CASA 2012 Category 3: Active suicidal ideation: method, but no intent or plan									
C-SSRS	No	12 (30.0)	1 (2.5)	13 (32.5)	0	13 (32.5)	0	13 (33.3)	0
	Yes	23 (57.5)	4 (10.0)	27 (67.5)	0	25 (62.5)	2 (5.0)	24 (61.5)	2 (5.1)
% Agreement*		16/40 (40.0)		13/40 (32.5)		15/40 (37.5)		15/39 (38.5)	
P value**		<0.001		-		<0.001		<0.001	
Kappa (95% CL)		0.071 (-0.090, 0.189)		-		0.049 (-0.021, 0.120)		0.053 (-0.023, 0.128)	
AUC		0.536		0.500		0.537		0.538	
FDA-CASA 2012 Category 4: Active suicidal ideation: method and intent, but no plan									
C-SSRS	No	17 (42.5)	1 (2.5)	18 (45.0)	0	17 (42.5)	1 (2.5)	17 (43.6)	1 (2.6)
	Yes	22 (55.0)	0	20 (50.0)	2 (5.0)	20 (50.0)	2 (5.0)	20 (51.3)	1 (2.6)
% Agreement*		17/40 (42.5)		20/40 (50.0)		19/40 (47.5)		18/39 (46.2)	
P value**		<0.001		<0.001		<0.001		<0.001	
Kappa (95% CL)		-0.050 (-0.148, 0.047)		0.083 (-0.003, 0.196)		0.032 (-0.114, 0.179)		-0.007 (-0.137, 0.122)	
AUC		0.472		0.545		0.518		0.496	
FDA-CASA 2012 Category 5: Active suicidal ideation: method, intent, and plan									
C-SSRS	No	20 (50.0)	3 (7.5)	16 (40.0)	7 (17.5)	19 (47.5)	4 (10.0)	18 (46.2)	5 (12.8)
	Yes	5 (12.5)	12 (30.0)	0	17 (42.5)	3 (7.5)	14 (35.0)	1 (2.6)	15 (38.5)
% Agreement*		32/40 (80.0)		33/40 (82.5)		33/40 (82.5)		33/39 (84.6)	
P value**		0.480		0.008		0.705		0.102	
Kappa (95% CL)		0.584 (0.329, 0.840)		0.660 (0.445, 0.875)		0.645 (0.406, 0.883)		0.694 (0.473, 0.914)	
AUC		0.788		0.848		0.825		0.860	
FDA-CASA 2012 Category 7: Suicidal attempt									
C-SSRS	No	30 (75.0)	0	30 (75.0)	0	30 (75.0)	0	29 (74.4)	0
	Yes	2 (5.0)	8 (20.0)	2 (5.0)	8 (20.0)	3 (7.5)	7 (17.5)	2 (5.1)	8 (20.5)
% Agreement*		38/40 (95.0)		38/40 (95.0)		37/40 (92.5)		37/39 (94.9)	
P value**		0.157		0.157		0.083		0.157	
Kappa (95% CL)		0.857 (0.666, 1.000)		0.857 (0.666, 1.000)		0.778 (0.542, 1.000)		0.856 (0.664, 1.000)	
AUC		0.900		0.900		0.850		0.900	
FDA-CASA 2012 Category 8: Interrupted suicide attempt									
C-SSRS	No	32 (80.0)	0	32 (80.0)	0	31 (77.5)	1 (2.5)	31 (79.5)	0
	Yes	8 (20.0)	0	5 (12.5)	3 (7.5)	6 (15.0)	2 (5.0)	8 (20.5)	0
% Agreement*		32/40 (80.0)		35/40 (87.5)		33/40 (82.5)		31/39 (79.5)	
P value**		-		0.025		0.059		-	
Kappa (95% CL)		-		0.490 (0.130, 0.850)		0.286 (-0.082, 0.653)		-	
AUC		0.500		0.688		0.609		0.500	
FDA-CASA 2012 Category 9: Aborted suicide attempt									
C-SSRS	No	36 (90.0)	0	36 (90.0)	0	35 (87.5)	1 (2.5)	35 (89.7)	0
	Yes	4 (10.0)	0	3 (7.5)	1 (2.5)	3 (7.5)	1 (2.5)	4 (10.3)	0
% Agreement*		36/40 (90.0)		37/40 (92.5)		36/40 (90.0)		35/39 (89.7)	
P value**		-		0.083		0.317		-	
Kappa (95% CL)		-		0.375 (-0.156, 0.906)		0.286 (-0.214, 0.786)		-	
AUC		0.500		0.625		0.611		0.500	

TABLE 4, continued. Percent agreement, n (%), kappa, and AUC for ISST-Plus vs. C-SSRS and S-STS vs. C-SSRS

	ISST-PLUS (N = 40)		S-STS (PATIENT) (N = 40)		S-STS (CLINICIAN) (N = 40)		S-STS (RECONCILED) (N = 39)		
	No	Yes	No	Yes	No	Yes	No	Yes	
FDA-CASA 2012 Category 10: Preparatory acts toward imminent suicidal behavior									
C-SSRS	No	25 (62.5)	5 (12.5)	26 (65.0)	4 (10.0)	28 (70.0)	2 (5.0)	29 (74.4)	0
	Yes	1 (2.5)	9 (22.5)	4 (10.0)	6 (15.0)	6 (15.0)	4 (10.0)	10 (25.6)	0
% Agreement*	34/40 (85.0)		32/40 (80.0)		32/40 (80.0)		29/39 (74.4)		
P value**	0.102		1.000		0.157		-		
Kappa (95% CL)	0.647 (0.396, 0.898)		0.467 (0.151, 0.782)		0.385 (0.048, 0.722)		-		
AUC	0.867		0.733		0.667		0.500		
FDA-CASA 2012 Category 11: Self-injurious behavior without suicidal intent									
C-SSRS	No	27 (67.5)	2 (5.0)	22 (55.0)	7 (17.5)	29 (72.5)	0	29 (74.4)	0
	Yes	5 (12.5)	6 (15.0)	1 (2.5)	10 (25.0)	3 (7.5)	8 (20.0)	1 (2.6)	9 (23.1)
% Agreement*	33/40 (82.5)		32/40 (80.0)		37/40 (92.5)		38/39 (97.4)		
P value**	0.257		0.034		0.083		0.317		
Kappa (95% CL)	0.521 (0.215, 0.826)		0.571 (0.320, 0.822)		0.795 (0.576, 1.000)		0.930 (0.796, 1.000)		
AUC	0.738		0.834		0.864		0.950		

* % Agreement: raw concordance
 ** P value based on McNemar's test of significance

ISST-Plus: InterSePT Scale for Suicidal Thinking-Plus; S-STS (Patient): Sheehan-Suicidality Tracking Scale patient-rated version; S-STS (Clinician): Sheehan-Suicidality Tracking Scale clinician-rated version; C-SSRS: Columbia-Suicide Severity Rating Scale; AUC: Area under Receiver Operating Curve

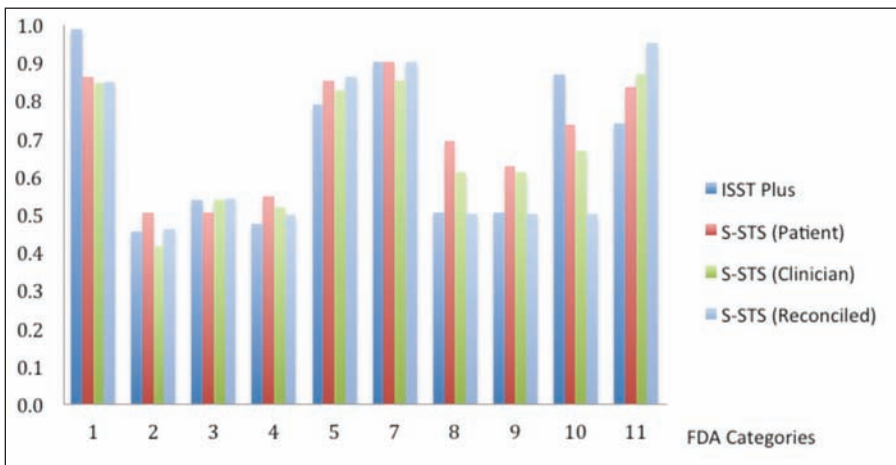


FIGURE 3. Area under curve (AUC) for the ISST-Plus and S-STS by FDA-CASA 2012 category relative to the C-SSRS

ISST-Plus: InterSePT Scale for Suicidal Thinking-Plus; S-STS (Patient): Sheehan-Suicidality Tracking Scale patient-rated version; S-STS (Clinician): Sheehan-Suicidality Tracking Scale clinician-rated version; S-STS (Reconciled): Sheehan-Suicidality Tracking Scale reconciled version; C-SSRS: Columbia-Suicide Severity Rating Scale

version done because their scores on the patient- and clinician-rated versions of the S-STS were identical. For 34 of the 37 participants (91.9%), there was at least one score difference between the patient- and clinician-rated versions of the S-STS. Consequently, the reconciliation version of S-STS was needed for these 34 subjects. The reconciliation between patient and clinician versions required less than three

minutes, even in suicidal subjects. Overall, the results show that assessments of suicidality can be done in a short time frame using any of these scales and that the patient-rated version of the S-STS takes approximately the same length of time as the clinician versions.

DISCUSSION

This study is the first examination of the concurrent validity of the

ISST-Plus and S-STS in relation to the C-SSRS. The study has several important strengths: 1) assessments were all made by different raters who were blind to the results of the other interviews (avoiding potential rater bias); 2) clinician interviews were randomly sequenced (mitigating against possible order effects); and 3) direct data capture and computer-coded mapping precluded missing values, double entries, legibility problems, and transcription errors for ISST-Plus and S-STS data.

Overall, there was good agreement on all three scales for some categories and poor agreement for others. The greatest disagreement between the test instruments and the C-SSRS was seen in patients with “intermediate levels” of active suicidal ideation (categories 2, 3, and 4) (i.e., categories short of including all 3 of method, intent, and plan). Specifically the C-SSRS endorsed category 2 (“Non-specific active suicidal thoughts”) 13 or more times as often as the ISST-Plus and S-STS with almost two-thirds (65%) of the 40 patients mapping to this category on the C-SSRS compared to five percent or less on the ISST-Plus and the three versions of the S-STS. Similar patterns were observed for the active ideation categories 3 and 4

with the C-SSRS endorsing these categories 20 or more times as often with 55 percent or more of subjects mapping to these categories compared to 2.6 percent or less on the other two scales.

These discrepancies could be interpreted to mean that the ISST-Plus and S-STS both under-endorse certain categories of active ideation or they could be interpreted to mean, conversely, that the C-SSRS over-endorse these phenomena. We suspect that the latter interpretation is more plausible for the following reason: The C-SSRS has a logical flaw in that it requires that a subject answer “Yes” to question 2 (“Non-specific Active Suicidal Thoughts”) to proceed to subsequent active ideation questions. However, if the subject answers “Yes” to question 2, questions on active ideation (3, 4, and 5) should theoretically be answered “No.” This is because the “Yes” to question 2 is predicated on not having “thoughts of ways to kill oneself / associated methods, intent, or plan during the assessment period.”⁴ In addition, the probe question for “Non-Specific Active Suicidal Thoughts” is, “Have you actually had any thoughts of killing yourself?” A “Yes” response to this question could, we think, map to very specific rather than *non-specific* active suicidal thoughts in clinical practice. That it does not in this case is likely to lead to substantial inflation of false positives on this category in the C-SSRS.⁸ Neither the S-STS nor the ISST-Plus have such navigation flaws or mismatches between probe questions and FDA-CASA 2012 categories.

There is an additional problem. Not all combinations of active suicidal ideation are captured in the C-SSRS and by extension the FDA-CASA 2012. As pointed out here and in two companion articles, as many as 26 combinations of suicidal ideation out of 32 possible combinations are excluded in the C-SSRS and FDA-CASA 2012.^{7,8} These combinations are captured by the other two scales. In different words, it is not that the ISST-

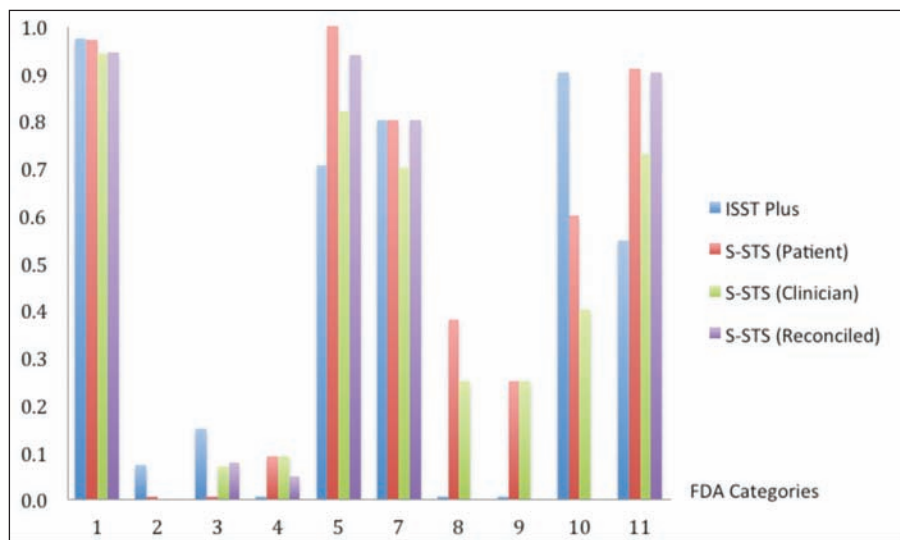


FIGURE 4. Sensitivity of the ISST-Plus and the S-STS by FDA-CASA 2012 category using the C-SSRS

ISST-Plus: InterSePT Scale for Suicidal Thinking-Plus; S-STS (Patient): Sheehan-Suicidality Tracking Scale patient-rated version; S-STS (Clinician): Sheehan-Suicidality Tracking Scale clinician-rated version; S-STS (Reconciled): Sheehan-Suicidality Tracking Scale reconciled version; C-SSRS: Columbia-Suicide Severity Rating Scale

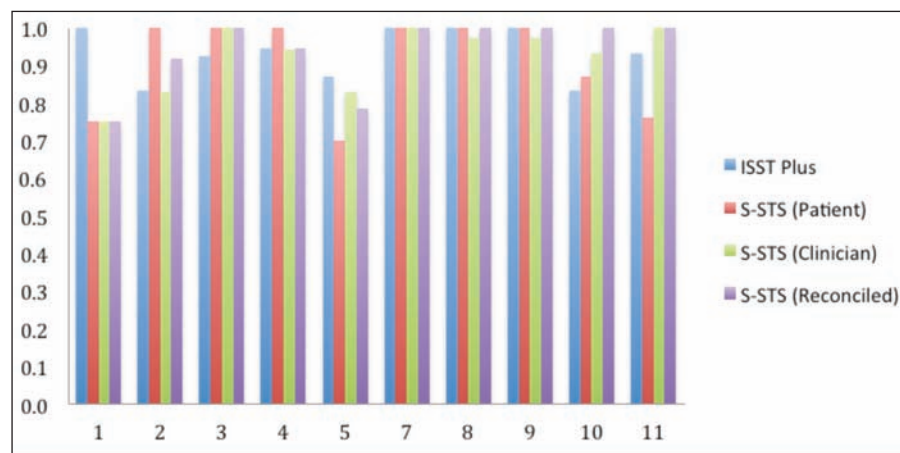


FIGURE 5. Specificity of ISST-Plus and S-STS by FDA-CASA category using the C-SSRS

ISST-Plus: InterSePT Scale for Suicidal Thinking-Plus; S-STS (Patient): Sheehan-Suicidality Tracking Scale patient-rated version; S-STS (Clinician): Sheehan-Suicidality Tracking Scale clinician-rated version; S-STS (Reconciled): Sheehan-Suicidality Tracking Scale reconciled version; C-SSRS: Columbia-Suicide Severity Rating Scale

Plus and S-STS under-endorse active ideation; rather, they are more specific. On the S-STS and the ISST-Plus, the components that make up these combinations are disaggregated at the interview and data acquisition level. They are later aggregated into all 32 possible combinations by the computer after entry. This is not the case on the C-SSRS, which does not

capture all the components making up the combinations separately. It only captures six of the possible 32 combinations directly on the scoring form, rather than disaggregating all the elements and later recombining them. Raters have told the authors of this article that they and the patients encountered difficulties in rating the complex combinations of suicidal

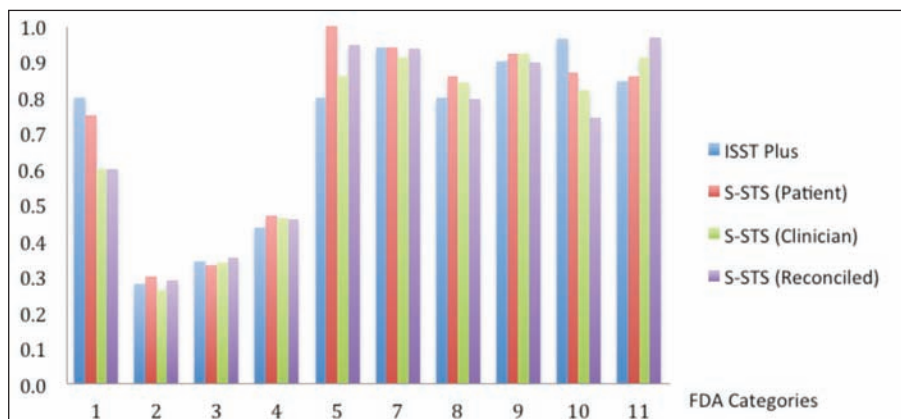


FIGURE 6. Negative Predictive Value (NPV) for each scale by FDA-CASA 2012 classification relative to the C-SSRS

ISST-Plus: InterSePT Scale for Suicidal Thinking-Plus; S-STs (Patient): Sheehan-Suicidality Tracking Scale patient-rated version; S-STs (Clinician): Sheehan-Suicidality Tracking Scale clinician-rated version; S-STs (Reconciled): Sheehan-Suicidality Tracking Scale reconciled version; C-SSRS: Columbia-Suicide Severity Rating Scale

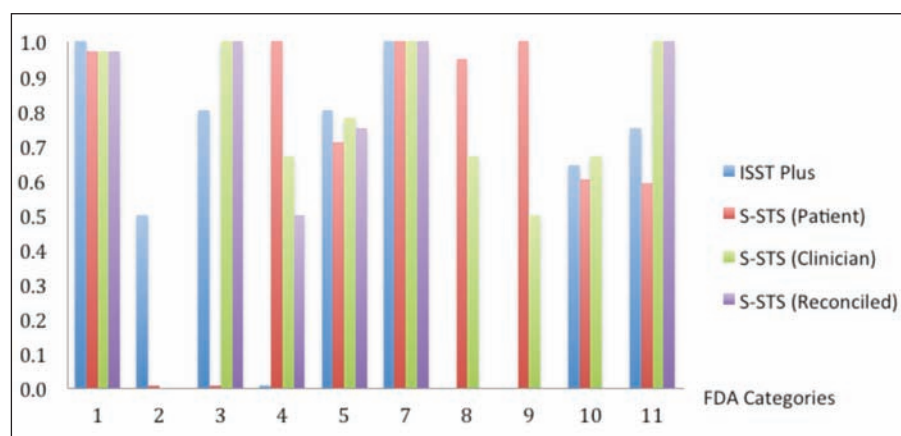


FIGURE 7. Positive Predictive Value (PPV) for each scale by FDA-CASA 2012 classification relative to the C-SSRS

ISST-Plus: InterSePT Scale for Suicidal Thinking-Plus; S-STs (Patient): Sheehan-Suicidality Tracking Scale patient-rated version; S-STs (Clinician): Sheehan-Suicidality Tracking Scale clinician-rated version; S-STs (Reconciled): Sheehan-Suicidality Tracking Scale reconciled version; C-SSRS: Columbia-Suicide Severity Rating Scale

phenomena to categories 2, 3, and 4 in the C-SSRS.

Less severe but still significant disagreement is seen on “intermediate levels” of suicide attempts (i.e., between preparatory acts and actual attempt). As with the ideation categories, there is general agreement among all scales when the C-SSRS categorizes patients as not falling into categories 8 and 9 (interrupted and aborted suicide attempt, respectively). However, when the

C-SSRS rates patients as meeting criteria for these categories, the ISST-Plus and the S-STs are more likely to agree with each other but to disagree with the C-SSRS. It was our impression that delineation of each of these three categories from each other is not precise enough in the C-SSRS and the FDA-CASA 2012 and leaves too much latitude for raters and patients to vary in their interpretations and abilities to respond to questions about these phenomena.

The differences may be a function of ambiguity in the C-SSRS navigation instructions and mismatches between probe questions, the title names, and the category definitions for the C-SSRS suicidal ideation categories. They may also be a function of rater and patient difficulty in reliably rating the complex combinations on these four active suicidal ideation categories as they appear on the C-SSRS. Differences in categorization observed for the three scales may be related to differences in the instruments’ approaches to collecting SIB data, problems in FDA-CASA 2012 classification, variability in patient reports, and rater reliability. The findings suggest a need to conduct larger-scale psychometric studies, including inter-rater and intra-rater reliability testing at additional sites. These studies should include a diverse sample of psychiatric and medical co-morbidities and subjects with diverse ethnic and religious backgrounds.

Limitations. The study had several limitations: 1) the sample size of the study was modest; 2) the Kappa values might have been affected by low base rates; 3) the results for FDA-CASA 2012 categories 6, 12, 13, and 14 could not be analyzed since these items did not occur in the subjects surveyed; and 4) emphasis in the C-SSRS on suicidal patients meant that it could not be used as a standard reference for category 15 (other, [no deliberate self-harm]).

CONCLUSION

Concordance of the ISST-Plus and S-STs to the C-SSRS was acceptable for categories 1, 5, 6, 10, and 11, but not for the active suicidal ideation categories 2, 3, 4, or for the aborted and interrupted suicide attempt categories 8 and 9.

Based on the results of this exploratory study, we cannot recommend merging data collected using these scales into a common database for meta-analyses. Indeed

the magnitude of the discrepancies calls into question the validity of the gold standard if it disagrees so much with the two other standardized instruments that agree so closely with each other. Modification of the C-SSRS to address the problems raised by Giddens et al⁸ and to capture the additional missing combinations might significantly narrow these discrepancies. However such modifications would be considerable and consequently would require that the C-SSRS so modified be revalidated against some other standard.

The lack of agreement on these categories should not be taken to mean that the ISST-Plus and S-STS are weak and the C-SSRS is strong. Rather it invites debate and discussion over whether these alternative instruments better tap into the full spectrum of suicidal phenomena that exist and whether the C-SSRS in its existing form should continue to be the reference standard.

ACKNOWLEDGMENTS

The authors acknowledge Samantha White, BS; Courtney Blair MA; and Jaymee Nelson, MD for their work as raters in this study and Kathy Harnett Sheehan, PhD; Jennifer M. Giddens, Sheena Hunt, PhD; and Matthew Grzywacz, PhD for providing editorial assistance.

REFERENCES

1. United States Food and Drug Administration, United States Department of Health and Human Services. *Guidance for Industry: Suicidality: Prospective Assessment of Occurrence in Clinical Trials, Draft Guidance*. September 2010. <https://www.federalregister.gov/articles/2010/09/09/2010-22404/draft-guidance-for-industry-on-suicidality-prospective-assessment-of-occurrence-in-clinical-trials>. Accessed October 1, 2014.
2. Posner K, Oquendo, MA, Gould, et al. Columbia Classification

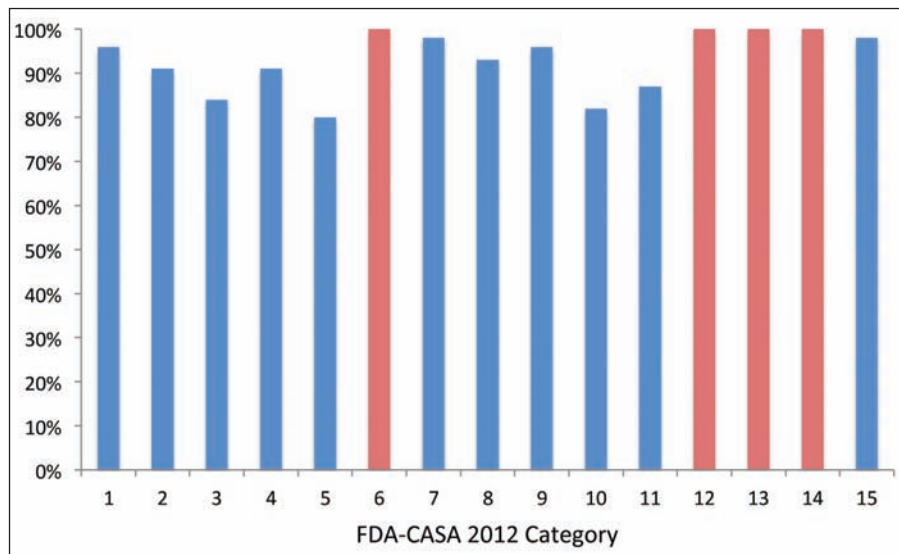


FIGURE 8. Percentage agreement between the clinician-rated version of the S-STS and the ISST-Plus based on FDA-CASA 2012 categories.

Notes: Since categories #6, #12, #13 and #14 on the FDA-CASA 2012 (shown as red bars) (completed suicide; self-injurious behavior intent unknown; fatal, not enough information; non-fatal, not enough information) did not occur in the subjects surveyed, there is 100% agreement.

ISST-Plus: InterSePT Scale for Suicidal Thinking-Plus; S-STS (Clinician): Sheehan-Suicidality Tracking Scale clinician-rated version; C-SSRS: Columbia-Suicide Severity Rating Scale

- Algorithm of Suicide Assessment (C-CASA): classification of suicidal events in the FDA's pediatric suicidal risk analysis of antidepressants, *Am J Psychiatry*. 2007;164:1035-1043.
3. United States Food and Drug Administration, United States Department of Health and Human Services. *Guidance for Industry: Suicidality: Prospective Assessment of Occurrence in Clinical Trials, Draft Guidance*. August 2012. Revision 1. <http://www.fda.gov/downloads/Drugs/Guidances/UCM225130.pdf>. Accessed October 1, 2014.
4. Posner K, Brown GK, Stanley B, et al. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *Am J Psychiatry*. 2011;168:1266-1277.
5. Gutierrez PM. Evaluation of existing psychometric data on the Columbia-Suicide Severity Rating Scale (C-SSRS). Working paper for the Military Suicide Research Consortium. Florida State University. December 6, 2011. https://msrc.fsu.edu/sites/default/files/MSRC_C-SSRS_evaluation.pdf. Accessed October 1, 2014.
6. Vandeppeer M. Health Policy Advisory Committee on Technology, Technology Brief, Columbia-Suicide Severity Rating Scale, HealthPact Emerging Health Technology. State of Queensland, Australia. August 2012. <http://www.health.qld.gov.au/healthpact/docs/briefs/WP114.pdf>. Accessed October 1, 2014.
7. Sheehan DV, Giddens JM, Sheehan KH. Current assessment and classification of suicidal phenomena using the FDA 2012 Draft Guidance document on suicide assessment: a critical review. *Innov Clin Neurosci*. 2014;11(9-10):54-65.
8. Giddens JM, Sheehan KH, Sheehan DV. The Columbia-Suicide Severity

TABLE 5. Combinations of passive and active suicidal ideation, method, intent, and plan detected by the clinician-rated S-STS and by the C-SSRS for all 45 suicidal and non-suicidal subjects

POSSIBLE COMBINATION NUMBER	PASSIVE IDEATION	ACTIVE IDEATION	METHOD	INTENT	PLAN	COMBINATIONS NOT CAPTURED BY C-SSRS/FDA-CASA 2012 DEFINITIONS	# OF PATIENTS WHO MAP TO THE S-STS COMBINATIONS AND DO NOT MAP TO THE C-SSRS COMBINATIONS	% OF PATIENTS WHO MAP TO THE S-STS COMBINATIONS AND DO NOT MAP TO THE C-SSRS COMBINATIONS	# OF PATIENTS WHO MAP TO THE S-STS COMBINATIONS AND ALSO MAP TO THE C-SSRS COMBINATIONS	% OF PATIENTS WHO MAP TO THE S-STS COMBINATIONS AND ALSO MAP TO THE C-SSRS COMBINATIONS
1	0	0	0	0	0		•	•	8	17.8%
2	0	1	0	0	0		•	•	0	0
3	0	0	0	1	0	✓	0	0	•	•
4	0	1	0	1	0		0	0	•	•
5	0	0	1	0	0	✓	0	0	•	•
6	0	1	1	0	0		•	•	1	2.2%
7	0	0	1	1	0	✓	0	0	•	•
8	0	1	1	1	0	✓	0	0	•	•
9	0	0	0	0	1	✓	0	0	•	•
10	0	1	0	0	1	✓	0	0	•	•
11	0	0	0	1	1	✓	0	0	•	•
12	0	1	0	1	1		•	•	0	0
13	0	0	1	0	1	✓	0	0	•	•
14	0	1	1	0	1	✓	0	0	•	•
15	0	0	1	1	1	✓	1	2.2%	•	•
16	0	1	1	1	1	✓	0	0	•	•
17	1	0	0	0	0		•	•	6	13.3%
18	1	1	0	0	0	✓	2	4.4%	•	•
19	1	0	0	1	0	✓	2	4.4%	•	•
20	1	1	0	1	0	✓	2	4.4%	•	•
21	1	0	1	0	0	✓	1	2.2%	•	•
22	1	1	1	0	0	✓	1	2.2%	•	•
23	1	0	1	1	0	✓	0	0	•	•
24	1	1	1	1	0	✓	3	6.7%	•	•
25	1	0	0	0	1	✓	0	0	•	•
26	1	1	0	0	1	✓	0	0	•	•
27	1	0	0	1	1	✓	0	0	•	•
28	1	1	0	1	1	✓	0	0	•	•
29	1	0	1	0	1	✓	0	0	•	•
30	1	1	1	0	1	✓	0	0	•	•
31	1	0	1	1	1	✓	0	0	•	•
32	1	1	1	1	1	✓	18	40%	•	•
Total	32					26	30	67%	15	33%

S-STS: Sheehan-Suicidality Tracking Scale; C-SSRS: Columbia-Suicide Severity Rating Scale; FDA-CASA 2012: United States Food and Drug Administration-Classification Algorithm for Suicide Assessment

TABLE 6. Combinations of passive and active suicidal ideation, method, intent, and plan detected by the patient-rated S-STs and by the C-SSRS for the 40 suicidal subjects only

POSSIBLE COMBINATION NUMBER	PASSIVE IDEATION	ACTIVE IDEATION	METHOD	INTENT	PLAN	COMBINATIONS NOT CAPTURED BY C-SSRS/FDA-CASA 2012 DEFINITIONS	# OF PATIENTS WHO MAP TO THE S-STs COMBINATIONS AND DO NOT MAP TO THE C-SSRS COMBINATIONS	% OF PATIENTS WHO MAP TO THE S-STs COMBINATIONS AND DO NOT MAP TO THE C-SSRS COMBINATIONS	# OF PATIENTS WHO MAP TO THE S-STs COMBINATIONS AND ALSO MAP TO THE C-SSRS COMBINATIONS	% OF PATIENTS WHO MAP TO THE S-STs COMBINATIONS AND ALSO MAP TO THE C-SSRS COMBINATIONS
1	0	0	0	0	0		•	•	2	5%
2	0	1	0	0	0		•	•	0	0
3	0	0	0	1	0	✓	0	0	•	•
4	0	1	0	1	0		0	0	•	•
5	0	0	1	0	0	✓	0	0	•	•
6	0	1	1	0	0		•	•	0	0
7	0	0	1	1	0	✓	0	0	•	•
8	0	1	1	1	0	✓	0	0	•	•
9	0	0	0	0	1	✓	0	0	•	•
10	0	1	0	0	1	✓	0	0	•	•
11	0	0	0	1	1	✓	0	0	•	•
12	0	1	0	1	1		•	•	0	0
13	0	0	1	0	1	✓	1	2.5%	•	•
14	0	1	1	0	1	✓	0	0	•	•
15	0	0	1	1	1	✓	0	0	•	•
16	0	1	1	1	1	✓	1	2.5%	•	•
17	1	0	0	0	0		•	•	6	15%
18	1	1	0	0	0	✓	0	0	•	•
19	1	0	0	1	0	✓	1	2.5%	•	•
20	1	1	0	1	0	✓	0	0	•	•
21	1	0	1	0	0	✓	2	5%	•	•
22	1	1	1	0	0	✓	0	0	•	•
23	1	0	1	1	0	✓	0	0	•	•
24	1	1	1	1	0	✓	2	5%	•	•
25	1	0	0	0	1	✓	0	0	•	•
26	1	1	0	0	1	✓	0	0	•	•
27	1	0	0	1	1	✓	0	0	•	•
28	1	1	0	1	1	✓	0	0	•	•
29	1	0	1	0	1	✓	0	0	•	•
30	1	1	1	0	1	✓	1	2.5%	•	•
31	1	0	1	1	1	✓	1	2.5%	•	•
32	1	1	1	1	1	✓	23	57.5%	•	•
Total	32					26	32	80%	8	20%

S-STs: Sheehan-Suicidality Tracking Scale; C-SSRS: Columbia-Suicide Severity Rating Scale; FDA-CASA 2012: United States Food and Drug Administration-Classification Algorithm for Suicide Assessment

TABLE 7. Time taken to administer each scale (N=34)

SCALE	DURATION IN MINUTES AND SECONDS (MEAN±SD)
C-SSRS	8:13±3:08
ISST-Plus	14:54±10:23
S-STs	
S-STs patient-rated	8:35±4:47
S-STs clinician-rated	9:13±6:14
S-STs reconciliation (34 needing reconciliation)	2:51±1:28

SD: standard deviation; C-SSRS: Columbia-Suicide Severity Rating Scale; ISST-Plus: InterSePT Scale for Suicidal Thinking-Plus; S-STs: Sheehan-Suicidality Tracking Scale

- Rating Scale (C-SSRS): Has the “Gold Standard” become a liability? *Innov Clin Neurosci.* 2014;11(9-10):66-80.
9. Meyer RE, Salzman C, Youngstrom EA, et al. Suicidality and risk of suicide: definition, drug safety concerns, and a necessary target for drug development: a consensus statement. *J Clin Psychiatry.* 2010;71(8):e1-e21.
 10. Gassmann-Mayer C, Jiang K, McSorley P, et al. Clinical and statistical assessment of suicidal ideation and behavior in pharmaceutical trials. *Clin Pharmacol Ther.* 2011;90(4):554-560.
 11. Coric V, Stock EG, Pultz J, et al. Sheehan-Suicidality Tracking Scale (S-STs): preliminary results from a multicenter clinical trial in generalized anxiety disorder. *Psychiatry (Edgmtont).* 2009;1(6):26-31.
 12. Preti A, Sheehan DV, Coric V, et al. Sheehan Suicidality Tracking Scale (S-STs): reliability, convergent, and discriminative validity in young Italian adults. *Compr Psychiatry.* 2013;54(7):842-849.
 13. Sheehan DV, Giddens JM, Sheehan IS. Status Update on the Sheehan-Suicidality Tracking Scale (S-STs) 2014. *Innov Clin Neurosci.* 2014;11(9-10):93-140.
 14. Lindenmayer JP, InterSePT Study Group, Czobor P, et al. The InterSePT scale for suicidal thinking reliability and validity. *Schizophr Res.* 2003;63(1-2):161-170.
 15. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas.* 1960;20(1):37-46.
 16. Shrout PE, Spitzer R, Fleiss JL. Quantification of agreement in psychiatric diagnosis revisited. *Arch Gen Psychiatry.* 1987;44(2):172-177.
 17. Fleiss HJL, *Statistical Methods for Rates and Proportions, Second Edition.* New York, NY: John Wiley & Sons, 1981.
 18. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol.* 1993;46(5):423-429.
 19. Kessler RC, Akiska HS, Angst J, et al. Validity of the assessment of bipolar spectrum disorders in the WHO CIDI 3.0. *J Affect Disord.* 2006;96(3):259-269.
 20. Agresti A. *Categorical Data Analysis, Second Edition.* Hoboken, NJ: John Wiley & Sons, Inc; 2002.
 21. Kessler RC, Barker PR, Colpe LJ, et al. Screening for serious mental illness in the general population. *Arch Gen Psychiatry.* 2003;60(2):184-189.
 22. Hanley J, McNeil B. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982;143(1):29-36.
 23. Giddens JM, Sheehan DV. Is there any value in asking the question “Do you think you would be better off dead?” in assessing suicide? *Innov Clin Neurosci.* 2014;11(9-10):182-190. ■