



The Columbia–Suicide Severity Rating Scale (C–SSRS): Has the “Gold Standard” Become a Liability?

FUNDING: There was no funding for the development and writing of this article.

FINANCIAL DISCLOSURES:

J. Giddens is the author and copyright holder of the Suicide Plan Tracking Scale (SPTS) and is a named consultant on the Sheehan-Suicidality Tracking Scale (S-STSS), the Sheehan-Suicidality Tracking Scale Clinically Meaningful Change Measure Version (S-STSS CMCM), the Pediatric versions of the S-STSS, and the Suicidality Modifiers Scale; Dr. K. Sheehan is the spouse of Dr. D. Sheehan, who is the author and copyright holder of the S-STSS, the S-STSS CMCM, the Pediatric versions of the S-STSS, the Sheehan Disability Scale (SDS), and the Suicidality Modifiers Scale, is a co-author of the SPTS, the Mini International Neuropsychiatric Interview (MINI), and owns stock in Medical Outcomes Systems, which has computerized the MINI and the S-STSS. She has no other conflicts to report; and Dr. D. Sheehan is the author and copyright holder of the S-STSS, the S-STSS CMCM, the Pediatric versions of the S-STSS, the SDS, and the Suicidality Modifiers Scale, is a co-author of the SPTS, the Mini International Neuropsychiatric Interview (MINI), and owns stock in Medical Outcomes Systems, which has computerized the MINI and the S-STSS.

ADDRESS CORRESPONDENCE TO:

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

KEY WORDS: Suicide scale, suicide assessment, suicide risk, suicide attempt, suicide, suicidal ideation, suicidal behavior, suicidality, C-SSRS, FDA 2012 Draft Guidance Document

by **JENNIFER M. GIDDENS; KATHY HARNETT SHEEHAN, PhD; and DAVID V. SHEEHAN, MD, MBA**

J. Giddens is the Co-founder of the Tampa Center for Research on Suicidality, Tampa, Florida; Dr. K. Sheehan is Associate Professor Emeritus at the University of South Florida College of Medicine, Tampa, Florida; and Dr. D. Sheehan is Distinguished University Health Professor Emeritus at the University of South Florida College of Medicine, Tampa, Florida.

Innov Clin Neurosci. 2014;11(9–10):66–80

ABSTRACT

Objective: The Columbia–Suicide Severity Rating Scale has become the gold standard for the assessment of suicidal ideation and behavior in clinical trials. Criticism of the instrument has been mounting. We examine whether the instrument meets widely accepted psychometric standards and maps to the United States Food and Drug Administration’s most recent 2012 algorithm for assessment of suicidal phenomena. Our goal is to determine if the Columbia–Suicide Severity Rating Scale should be retained as the preferred instrument for assessment of suicidal ideation and behavior.

Method: Standard psychometric criteria dictate that categorizations to avoid type I and type II errors should be comprehensive and address the full spectrum (i.e., all dimensions) of a phenomenon. The

criteria should also be well defined and consistent, and the wording throughout should be unambiguous. We examine the Columbia–Suicide Severity Rating Scale in terms of these criteria.

Results: The Columbia–Suicide Severity Rating Scale does not address the full spectrum of suicidal ideation or behavior. As a result, it has the potential to miss many combinations of suicidal ideation and behavior that present to clinicians in practice (type II error). Potential misclassifications (type I and II errors) are compounded by flawed navigation instructions; mismatches in category titles, definitions, and probes; and wording that is susceptible to multiple interpretations. Further, the Columbia–Suicide Severity Rating Scale in its current form does not map to the 2012 Food and Drug Administration’s draft classification

algorithm for suicidal ideation and behavior.

Conclusion: The evidence suggests that the Columbia–Suicide Severity Rating Scale is conceptually and psychometrically flawed and does not map to the Food and Drug Administration’s new standards. A new gold standard for assessment of suicidality may be warranted.

INTRODUCTION

In 2012, the United States Food and Drug Administration (FDA) made the Columbia–Suicide Severity Rating Scale (C–SSRS)¹ the preferred instrument—the “gold standard”—for measuring suicidal ideation and behavior in clinical trials going forward. But has the gold standard become a liability? In this paper, we trace the making of the C–SSRS into a gold standard. We then evaluate the instrument in the context of widely accepted psychometric criteria and the FDA’s own most recent algorithm for classifying suicidal ideation and behavior.

BACKGROUND: THE MAKING OF A GOLD STANDARD

The C–SSRS was developed for a National Institute of Mental Health study of adolescent suicide attempters reported in 2007.² From the outset, the scale adopted definitions from the Columbia Suicide History Form (CSHF),³ an instrument originally designed as a chart extraction tool. The scale was also tied to the Columbia Classification Algorithm of Suicide Assessment (C–CASA), an algorithm commissioned by the FDA in 2007 for retrospective study of adverse events related to use of antidepressants.⁴ In 2010, the FDA sanctioned the use of the nine-category C–CASA and, by extension, approved the C–SSRS as the basis for mapping suicidal ideation and behavior in trials falling under its authority.⁵ More recently, in 2012, the FDA conferred gold standard status on the C–SSRS by labeling it the preferred instrument for

measuring suicidal ideation and behavior in clinical trials going forward.⁶

But was this gold standard status justified? And to what extent does the C–SSRS map to the more recent 2012 classification system the FDA now endorses for monitoring suicidal ideation and behavior?⁶ These are important questions since how suicidal ideation and behavior are measured in clinical trials has serious ramifications. On one level, whatever instrument is chosen as the preferred instrument will determine who is included or excluded in a particular trial. On another level, instrument choice influences whether suicidal ideation and/or behavior are connected to a particular drug (e.g., as an adverse effect). In other words, the choice has serious safety implications.

In the normal course of events, gold standard status is achieved slowly over time. According to Martin Roth, “it took more than a decade” before the Hamilton Rating Scale for Depression (HRSD) was even recognized as a worthy contribution to practice.⁷ It took more than two decades before the HRSD was endorsed by the World Health Organization (WHO) and the FDA, and this only occurred after it had been subjected to enormous scrutiny that included hundreds of studies by clinicians and researchers in the field, as opposed to “top down” endorsement by regulatory agencies and government authorities.⁸

Why was the C–SSRS so quickly endorsed by the FDA? After all, the C–SSRS was not the only suicide rating scale in use at the time. In fact, there were and are numerous other scales in use, including Harkavy-Asnis Suicide Survey,⁹ the InterSePT Scale for Suicidal Thinking (ISST-Plus),¹⁰ the Suicidal Behaviors Questionnaire–Revised,¹¹ the Beck Scale for Suicidal Ideation (BSI or BSS),¹² and the Sheehan-Suicidality Tracking Scale (S-STs)¹³ to name a few. One explanation is that suicide was much in the news, and the FDA

needed to identify a scale as rapidly as possible to be able to show the public that it was on top of this lethal problem. The C–SSRS would have been convenient since the FDA had already adopted the C–CASA for classificatory purposes. In addition, as the FDA 2012 Draft Guidance specifically notes, the C–SSRS definitions had by then been adopted by the United States Center for Disease Control and Prevention (CDC),¹⁴ and there was evidence from the Columbia University website that the C–SSRS was frequently requested by national and international agencies, including various branches of the United States Military, the Israeli Defense Force, Health Canada, Japanese National Institute of Mental Health, Korean Association for Suicide Prevention, the United States Department of Education, the CDC, the FDA, and the WHO.¹⁵ But is “top-down” recognition by government authorities equal to rigorous scientific scrutiny?

MINIMAL SCRUTINY

Despite its gold standard conferral, the C–SSRS has been relatively under scrutinized. As Chappell et al¹⁶ observed, published papers on the reliability and validity of the scale have “lagged behind” its widespread acceptance. Moreover, what published data do exist have been inconsistent for inter-rater reliability^{17,18} and limited for validity.¹⁹ Although two validation reports have been published,^{1,18} one is a retrospective analysis of data from three clinical trials that were not specifically designed as tests of the psychometric properties of the instrument and neither was prospectively designed to compare C–SSRS with alternative suicide assessment tools. Additionally, both studies are limited by small sample sizes and both have been deemed by at least one government expert as “low quality studies (level III-evidence).”²⁰ Gutierrez from the Military Suicide Research Consortium in evaluating the existing

psychometric data on the C-SSRS concluded the “C-SSRS requires more study before it can be recommended for use[...].”¹⁹ This relatively light attention is worrisome since not all scales that are widely used or accorded gold standard status continue to live up to the level expected of them.²¹

MOUNTING CRITICISM

In the case of the C-SSRS, criticism of the instrument has been mounting.^{16,19,20} In particular, the third author of this article (D.S.) frequently hears from investigators—those who are now required to use the scale in sponsored studies—that the C-SSRS, because of its navigation flaws and other issues, under-identifies many cases of suicidal ideation (type II error), misclassifies and/or over-identifies others (type I error), and misses other suicidality combinations entirely. There is also growing concern that the scale in its current format does not map to the updated FDA classification algorithm outlined in the FDA’s 2012 Draft Guidance and is therefore becoming more of a liability than an asset.²²

PURPOSE OF THIS PAPER

This paper extends the debate on the C-SSRS. Guided by Guilford’s observation, articulated more than a half century ago, that rating categorizations need to be “well-defined, mutually exclusive, univocal and exhaustive,”²³ we evaluate whether the C-SSRS adequately meets widely accepted psychometric standards. In doing so, we ask the following questions: 1) Is the C-SSRS comprehensive? Does it address the *full* spectrum of suicidal ideation and behavior and avoid type II errors (i.e., the danger of missing true suicidal phenomena)? 2) Is the scale consistent? Does it provide consistent instructions, definitions, and probes, mitigating against misclassification, including type I errors (i.e., the danger of false positives)? 3) Is the scale unambiguous? Are questions worded in such a way that they are not susceptible to multiple

interpretations, mitigating against type I as well as type II errors? 4) To what extent does the C-SSRS map to the new 2012 FDA classification algorithm for suicide assessment (henceforth referred to as FDA-CASA 2012)? and 5) Is there consistency across all versions of the C-SSRS (a critical component to mapping to the FDA-CASA 2012 Draft Guidelines)?

FINDINGS

1. Is the C-SSRS comprehensive?

Does it address the full range of suicidal ideation and behavior? The first task in developing any scale is the definition of the construct. One part of this task is deciding how broadly the construct needs to be defined. Another part is how finely the construct should be divided. In principle, parsimony is preferable but not at the expense of ignoring the full range of phenomena that make up the construct.²⁴ In fact, generating a representative and comprehensive pool of items is considered a critical step in scale construction since, as Clark and Watson observe, “No existing data-analytic technique can remedy serious deficiencies in an item pool.”²⁵ In the case of multidimensional phenomena, such as suicidal ideation and behavior, particular care needs to be taken to ensure that the scale encompass *all* dimensions to avoid the risk of under-identification of phenomena that could pose a deadly risk. Ideally a construct-valid measure needs to tap into all the dimensions of the construct without surplus characteristics that might contaminate it and should provide comprehensive coverage of the range of the construct.²⁶

Our most serious criticism of the C-SSRS in terms of scale acceptance is that it does *not* cover the full spectrum of suicidal ideation. As shown in Table 1, there are as many as 16 possible combinations of active suicidal ideation (defined in terms of the presence or absence of method, intent, and plan) and 32 combinations if passive ideation is

considered as an additional factor. The C-SSRS, however, reduces this total number to four categories of active ideation and one category of passive ideation and the null of all these (possible combination number 1), i.e., a total of six categories. This means that as many as 26 categories of suicidal ideation are overlooked. While the overlooked combinations may constitute 20 percent of all combinations of suicide ideation event phenomena, they can constitute 60 percent of a patient’s time spent in suicidality, and often pose serious safety issues.^{27,28}

The S-STS/ISST-Plus/C-SSRS validation study, which mapped all three scales to the FDA-CASA 2012 categories, examined the extent of these missing combinations.²⁹ The results found that the C-SSRS categories did not capture combinations of suicidal ideation, method, intent, and plan that were detected in 67 percent of the subjects using the clinician-rated S-STS, 80 percent of the subjects using the patient-rated S-STS, and 76 percent of the subjects using the ISST-Plus.²⁹ Unlike the C-SSRS, the S-STS and the ISST-Plus map all 32 combinations as shown in Table 1.²⁹

Consider the patient who does not have current passive or active suicidal ideation but made a suicide plan some time ago and intends to carry out this plan at some point in the future. The patient has a plan with intent but without active or passive ideation (combination #11). Assume as well that the patient has never made a suicide attempt. This patient is likely to be at much more risk than someone without intent or plan, but there is no way of recognizing this risk on the C-SSRS because the patient’s experience doesn’t fit into any category.

What about the person who has passive ideation, method, and intent but no plan (combination #23)? This patient’s experience would not be recognized either.

Even more serious perhaps is the patient who presents to the doctor with a command hallucination to

TABLE 1. Possible combinations of suicidal ideation and missing combinations in C–SSRS based on category titles

Possible Combination Number	Passive Ideation	Active Ideation	Method	Intent	Plan	Which C–SSRS Category Title (1–5) Captures This Combination?		Combinations Not Captured by C–SSRS Titles
1	0	0	0	0	0	0	Covered by default	
2	0	1	0	0	0	2	“Non-Specific Active Suicidal Thoughts”	
3	0	0	0	1	0			✓
4	0	1	0	1	0	4	“Active Suicidal Ideation With Some Intent to Act, Without Specific Plan”	
5	0	0	1	0	0			✓
6	0	1	1	0	0	3	“Active Suicidal Ideation with Any Methods (Not Plan) Without Intent to Act”	
7	0	0	1	1	0			✓
8	0	1	1	1	0			✓
9	0	0	0	0	1			✓
10	0	1	0	0	1			✓
11	0	0	0	1	1			✓
12	0	1	0	1	1	5	“Active Suicidal Ideation with Specific Plan and Intent”	
13	0	0	1	0	1			✓
14	0	1	1	0	1			✓
15	0	0	1	1	1			✓
16	0	1	1	1	1			✓
17	1	0	0	0	0	1	“Wish to be Dead”	
18	1	1	0	0	0			✓
19	1	0	0	1	0			✓
20	1	1	0	1	0			✓
21	1	0	1	0	0			✓
22	1	1	1	0	0			✓
23	1	0	1	1	0			✓
24	1	1	1	1	0			✓
25	1	0	0	0	1			✓
26	1	1	0	0	1			✓
27	1	0	0	1	1			✓
28	1	1	0	1	1			✓
29	1	0	1	0	1			✓
30	1	1	1	0	1			✓
31	1	0	1	1	1			✓
32	1	1	1	1	1			✓
Total	32					6		26

C–SSRS: Columbia–Suicide Severity Rating Scale

commit suicide with a gun that same day at home. The patient does not fit the criteria for passive ideation. A “wish to be dead,”³¹ after all, implies a degree of willfulness. The patient cannot be said to have “active non-specific ideation.”³¹ In fact, the patient’s ideation is highly specific. On the other hand, you could say the patient has a deadly method in mind and, insofar as several of the “details are worked out,”³¹ the patient has a plan (combination #13), but this case too will not be recognized on the C–SSRS. This is an issue with the navigation instructions for suicidal ideation in the C–SSRS (see Section 2 on flawed navigation instructions).

The above cases are not unique, nor are they trivial. Indeed, the consequences of not considering them are broad. One consequence is the real possibility that patients who are at risk for suicide are included in studies when, per exclusion criteria, they should be excluded. An example is someone who decided a year ago that he intends to kill himself when his parents die, but not sooner. He has not thought about suicide in the past week or month, but this intent is unchanged and is only at the back of his mind, not overtly thought about during the investigation timeframe. What if his parents are killed in a car crash next week? This would then pose an immediate safety threat not readily detected by the C–SSRS. Another concern is that worsening suicidality associated with a study treatment can be missed by the C–SSRS. Consider the patient who is entered into a study and has some suicidal ideation at baseline. However, at Week 4 he reports an increased need to act on the suicidal thoughts sooner rather than later (increased urgency). Such a treatment-emergent change goes easily undetected by the C–SSRS.

The C–SSRS misses some types of passive suicidal ideation. For example, it does not detect “the thought that you would be better off dead.”³⁰ The importance of including a question about this type of passive

suicidal ideation is supported by the findings of Preti et al,³¹ who found that the question “Did you think you would be better off dead or wish you were dead?” had a 0.774 loading on a unidimensional model fit for the S-STS. Furthermore, thoughts of being better off dead can be an immediate antecedent to impulsive suicidality and is associated with functional impairment in work, social life, family life, and quality of life impacted by suicidality.³² Suicidal patients have conveyed to us that this is a suicidal phenomenon worthy of note.

2. Are instructions, definitions, and probes well-defined and clear?

Another aspect of scale development and acceptance is that instructions need to be clear and should not pose a burden on the reader. Here we will examine whether the navigation instructions in the C–SSRS meet these criteria.

“*Suicidal Ideation.*” Consider the instructions under “Suicidal Ideation.”³¹ The rater is told to ask questions 1 and 2: if the patient has a “wish to be dead” and if the patient has “non-specific active suicidal thoughts.”³¹ The rater is then instructed as follows: “If both are negative, proceed to ‘Suicidal Behavior’ section. If the answer to question 2 is ‘yes,’ ask questions 3, 4 and 5. If the answer to question 1 and/or 2 is ‘yes,’ [sic] complete ‘Intensity of Ideation’ section below.”³¹ This was extracted from the Lifetime/Recent Version 1/14/09 and Baseline/Screening Version of the C–SSRS—is the same on most versions of C–SSRS, except Screen Version that is now filed under Scales for Clinical Practice.

There are no directions on what to do if the answer to question 1 is yes and to question 2 is no. Given this ambiguity, different raters could handle this in different ways, leading to inter-rater unreliability.

Apart from the unclear nature of this instruction (e.g., there are 3 different “ifs” and 3 different paths to consider), the rater confronting a suicidal patient is faced with an immediate dilemma. Let’s say the

response to question 2 (non-specific active suicidal thoughts) is no, precisely because the patient has very specific active suicidal ideation. In fact, at this moment, the patient has a specific method, plan, and intent, but no non-specific active suicidal ideation. Specifically, the patient plans to take a fatal overdose at home this evening after work. The rater, however, is instructed to only ask about method, plan, and intent if the patient endorses *non-specific active* ideation.

In effect, the rater has to choose between two undesirable options. The rater can opt to follow the instructions explicitly and skip over the questions about method, intent, and plan. In this scenario, ideation that includes method, intent, or plan will be missed (type II error). Alternately, to err on the side of safety and be able to document method, intent, and/or plan, the rater can violate the instructions by incorrectly responding yes to question 2, thereby endorsing a nonexistent “non-specific active suicidal thought” (type I error).

In either scenario, the resulting data will be incorrect. That is, in the second scenario, there will be an invalid inflation (over-identification) of non-specific active ideation. We have made several patient-rater videos associated with this flawed navigation instruction that document precisely how this plays out in practice. Additionally, the rater may be forced once again to violate the navigation instructions by indicating the patient experienced one of the combinations in C–SSRS probe questions 3, 4, or 5 when the patient really experienced another combination (e.g., combination #10). The FDA-CASA 2012 validation study²⁹ identifies and highlights this over-inflation of endorsements by the C–SSRS on non-specific suicidal ideation compared to the comparable items on both the S-STS and the ISST-Plus. The S-STS and the ISST-Plus are concordant with each other on this finding and discordant with the C–SSRS.²⁹

The intent and directive in implementing this navigation instruction is explicit in a paper by Mundt et al¹⁸ on the computer automated C-SSRS. It is also captured at 48:06 minutes in a video that was made of an October 2011 Grand Rounds at the Child Center of New York University and was online and accessed by the authors on November 30, 2012.³³ The experience of the second two authors of this paper (K.S. and D.S.) is that the SOP (standard operating procedures) of pharmaceutical companies, clinical research organizations, and rater training companies conducting clinical trials reflect this C-SSRS instruction and they implement it assertively in monitoring clinical trials.

“Intensity of Ideation.” In the Lifetime/Baseline Version of C-SSRS, under “Intensity of Ideation,” the rater is told the following: “The following features should be rated with respect to the most severe type of ideation (i.e., 1–5 from above, with 1 being the least severe and 5 being the most severe). Ask about time he/she was feeling the most suicidal.”¹

The second part of this instruction indicates that the rater should ask about the time the patient felt “most suicidal.” However, the first part indicates that the rater should only provide intensity ratings for the “most severe” type of ideation, defined in terms of the five categories above the instruction.

What if a patient reports feeling “most suicidal” when only having passive ideation (C-SSRS category #1)? After all, this ideation went on for hours and it was frightening to the patient. What if, in addition, the patient reports that the experienced active ideation with a plan and intent (C-SSRS category #5) passed quickly and troubled the patient less? That is, from the patient’s viewpoint, it was much less severe and much less troublesome.

This type of scenario is not uncommon. In fact, the active type of ideation may be less frightening to

the patient than the passive ideation. At the least, the rater is faced with another dilemma. Should the ratings be provided for the time when the patient felt “most suicidal” (following the second part of the instruction) or should the ratings be provided for the time when the patient met criteria for the most severe category (following the first part of the instruction)? And what are the implications for the answers to specific questions regarding such things as frequency and duration if the rater makes one choice instead of the other? Clearly, if different raters make different choices, the integrity of the data will suffer. In our view, this issue creates inconsistencies across the different versions of C-SSRS.

“Suicidal Behavior.” In the “Suicidal Behavior” section of the C-SSRS, the rater is told the following: “Check all that apply, so long as these are separate events; must ask about all types.”

We find this contradictory. The first part (“check all that apply”) indicates that the rater should *only* rate those types of suicidal behavior that apply (i.e., check “yes” or “no”) if they are “separate events.” In other words, the rater can safely ignore questions about preparatory acts or behavior or other types of events if the patient reports a single “interrupted attempt.” On the other hand, the second part (“must ask about all types”) suggests that questions should be asked about all five types of suicidal behavior shown in this section. Does this mean that the rater should leave all the other the yes/no boxes blank? Or, following the second part of the instruction, should the rater place checkmarks in the “no” boxes for the other types? What if the patient did engage in preparatory acts or behavior *before* an interrupted or actual attempt? Which type of behavior trumps here? We may *assume* that attempts should trump preparatory acts, but the instructions are not explicit. In many cases, a confused rater, to be safe, may just check all types that

apply whether or not they are separate. This scenario undoubtedly leads to inflation of some types of behavior in the results (type I error).

In our experience, patients and raters are confused when the term “interrupted attempt” is used to describe an interrupted preparatory behavior and patients are told this attempt is not a suicide attempt. The use of this ambiguous term led to discrepancies between the C-SSRS on one side and the ISST-Plus and the S-STS on the other side in the validation study of all three.²⁹ The S-STS does not use the term “interrupted attempt,” thus avoiding this confusion. Use of this term also leads to many problems in translations to other languages thereby causing linguistic invalidation. In many languages, the term “interrupted attempt” cannot be translated in a way that makes sense in the end language without the addition of the word suicide (as in “interrupted suicide attempt”) [Personal communication to the third author (D.S.) in 2010, 2011, 2012, 2013, 2014 from MAPI Group, a leading international translation and linguistic validation agency involved in translating psychiatry scales and structured interviews]. When asked for clarification on what this term means, the scale author is required to explain that an “interrupted attempt” is not a suicide attempt at all (per C-SSRS definition), but is classified as a preparatory behavior. The third author of this paper (D.S.) has repeatedly faced this challenge in the translation of the S-STS and the MINI suicidality module into many languages with MAPI. The only satisfactory way to avoid this confusion in all languages, including English, is to avoid using the word *attempt* for either “interrupted attempt” or “aborted attempt.”³⁴

Some of the navigation problems may relate to the use of a Guttman Scale-like procedure in the design of the C-SSRS. Guttman scaling is used when the responses can be ranked in an order so that agreement with one item implies agreement with a lower

order item.³⁵ There is an implicit assumption in the C-SSRS that there is a severity hierarchy of suicidality going from passive to active ideation to method to intent to plan to preparatory behavior to attempt—much like going up the steps of stairs. However, according to some of our patients who have chronic suicidal ideation, this is a flawed and potentially dangerous assumption, with many exceptions, though we could not find any reference to this concern by others in the literature. Some of our patients told us that this assumption contributes to the poor ability to predict suicidal behavior. Guttman scaling is not appropriate for the assessment of suicidality. The flawed navigation instruction on C-SSRS item 2 compounds this problem further by violating established Guttman scaling procedures.

Mismatches in titles, definitions, and probes. Complicating the above issues is the potential for further classification error because of mismatches in the C-SSRS's titles, definitions, and probes. Ideally, categories should be well defined to avoid overlap. But as shown in Table 2, probe questions for suicidal ideation do not fully align with their corresponding titles and definitions for *any* of the five ideation categories, while those for suicidal behavior only align for two of the four categories with definitions. These mismatches have the potential to create type I and II errors.

For example, for the category, “Non-specific Active Suicidal Thoughts,” the probe, “Have you actually had any thoughts of killing yourself?”³¹ could elicit either specific or non-specific thoughts or both. If the probe generates a “yes” because of the presence of *specific* thoughts and the rater opts to use the probe rather than the definition (referring to non-specific thoughts), the result will be a type I error (over-identification of this C-SSRS category).

As another example, while the title and probe for category #3,

“Active Suicidal Ideation With Any Methods (Not Plan) Without Intent to Act,” requires thought associated with a method, the example within the definition appears to exclude thoughts of method—“I never made a specific plan as to when, where or how I would do it.”³¹ At the very least, the patient is likely to be confused by the nuances of this example. We, the authors of this paper, regularly encounter this confusion in clinical settings.

For the category, “Preparatory Acts or Behavior,” the following probe, “Have you taken any steps towards making a suicide attempt or preparing to kill yourself (such as collecting pills, getting a gun, giving valuables away or writing a suicide note)?”³¹ could similarly elicit a “yes” that conflicts with the definition, one that requires “imminence.” Depending on whether the probe or definition or title is used (and it isn't clear *which* should be used), a type I or a type II error could ensue.

The horizontal alignment of the yes/no response check boxes with the probe questions in most (but not all) versions of the C-SSRS suggests that the response options relate more to the probe question than to the title or definition. On the interactive voice recognition software (IVRS) version of C-SSRS (the eC-SSRS³⁶), the IVRS response is mapped directly to the C-SSRS probe question. To the extent that the response check boxes on the paper version might be mapped to the title or the definition or the probe question, this makes the paper and the IVRS versions not infrequently inconsistent.

3. Is the instrument's wording unambiguous? As Guilford reminds us, categorizations should be “univocal” (i.e., unambiguous).²³ The possibility of misclassification (type I or II error) is enhanced on the C-SSRS by unclear and imprecise wording and the frequent presence of words, phrases, and sentences that are ambiguous (have more than one meaning). Examples follow.

“Active Suicidal Ideation With Any Methods (Not Plan) Without

Intent to Act.” The rater has to keep in mind two positives (active suicidal ideation + any method) along with two negatives (no plan or intent). But then the rater is told in the definition, “This is different than a specific plan with time, place or method details worked out (e.g. [sic] thought of method to kill self but not a specific plan).”³¹ To what does “this” refer? To what does the example refer? And how does the rater differentiate between “thought of method” and “method details worked out?” Presumably, there is a fine distinction here, but the wording is susceptible to different interpretations. To further compound this ambiguity is the probe, “Have you been thinking about how you might do this?”³¹ Again, “this” has no referent and is susceptible to more than one interpretation. Depending on whatever interpretation the rater uses, there could be over-identification or under-identification of this category.

“Active Suicidal Ideation With Some Intent to Act, Without Specific Plan.” Here the definition requires “active suicidal thoughts” with “some intent to act on such thoughts,” but the last phrase is qualified with the phrase “as opposed to ‘I have thoughts but I definitely will not do anything about them’.”³¹ This phrase is ambiguous because it can have very different meanings. One interpretation of “will not do anything about them”³¹ is that the patient will not attempt suicide. Another way to interpret it, however, is that patient will do nothing in the way of getting any assistance in coping with these thoughts *or* seek treatment for the thoughts. In the latter instance, someone who *does* want to get help could be improperly identified as having suicidal intent and classified in this category (type I error) when the opposite is the case (i.e., the person intends to act on the thoughts by getting help). The clinician- and patient-rated versions C-SSRS could easily provide opposite ratings on this point, depending on

TABLE 2. Mismatches in probes, definitions, and titles in C–SSRS

Categories	Title and Definition Match	Title and Probe Match	Definition and Probe Match	Title, Definition, and Probe Match
<i>SUICIDAL IDEATION</i>				
“Wish to be Dead”	Yes	Yes	No ¹	No ¹
“Non-Specific Active Suicidal Thoughts”	No ²	No ²	No ²	No ²
“Active Suicidal Ideation with Any Methods (Not Plan) Without Intent to Act”	No	No ³	No ³	No ³
“Active Suicidal Ideation With Some Intent to Act, Without Specific Plan”	No ⁴	No ⁴	No	No ⁴
“Active Suicidal Ideation with Specific Plan and Intent”	Yes	No ⁵	No	No ⁵
<i>SUICIDAL BEHAVIOR</i>				
“Actual Attempt”	No ⁶	No	No ⁶	No ⁶
“Interrupted Attempt”	Yes	Yes	Yes	Yes
“Aborted Attempt”	Yes	Yes	Yes	Yes
“Preparatory Acts or Behavior”	No ⁷	Yes	No ⁷	No ⁷
“Suicide”	No definition ⁸	Yes	No definition ⁸	No definition ⁸
“Non-Suicidal Self-Injurious Behavior”	No definition	No probe	N/A	N/A

¹ Probe narrows the definition by not including a wish to be “not alive anymore.”*

² Definition expands the title by specifying no method, intent or plan. Probe expands the title by not clarifying that thoughts need to be non-specific. Probe expands the definition by not excluding thoughts associated with method, intent or plan.*

³ Probe expands title and definition by not specifying that thoughts need to be active or non-specific. Probe further expands or narrows title and definition depending on how rater interprets “this” in “How have you been thinking about how you might do this?”*

⁴ Definition expands the title by not clarifying that the category is precluded if there is a specific plan. Probe similarly expands the title by not specifying that this category cannot be endorsed if there is a plan.*

⁵ Definition expands the title by not specifying that thoughts need to be active (cannot be passive). Probe expands the title by not specifying that plan and intent need to be accompanied by active ideation.*

⁶ Definition limits the range of suicide attempts captured. For example, a child eats an apple her brother tells her has been poisoned and will kill her if she eats it. She has intent to die and eats the apple. But there is nothing wrong with the apple and he is teasing her. She believes him. This should be classified as a suicide attempt even though eating a harmless apple has no “potential for injury or harm” as the definition requires.*

⁷ Definition narrows the title by indicating that preparatory acts or behavior need to be “imminent.” Probe, however, expands the definition by not making this requirement.*

⁸ This category is tucked into a probe under “Suicidal Behavior.” The probe matches the title but there is no definition beyond what is in the probe.*

NOTE: The 3 authors disagreed among themselves on several of the decisions in this table. The authors also had different interpretations of these mismatches at different points in time. Others disagreed with these decisions, reflecting very poor inter-rater reliability on these points.

* 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.

C–SSRS: Columbia–Suicide Severity Rating Scale

the rater's interpretation of this question.

"Preparatory Acts or Behavior." According to the definition, preparatory acts or behavior are "acts or preparation towards imminently making a suicide attempt."¹ Use of the qualifier "imminently" is ambiguous because it lends itself to multiple interpretations. One rater may take it to mean "in the next 24 hours," another may take it to mean "in the next week," while a third may feel it can be interpreted as "in the next month." Without an explicit definition of "imminent," individual raters may select very different time frames—causing many type I and type II errors, depending on one's own time frame reference, within a trial. As a further complication, neither the title or the probe use the word "imminent." As a result some raters may simply ignore it.

"Interrupted Attempt." This type of attempt is defined as follows: "When the person is interrupted (by an outside circumstance) from starting the potentially self-injurious act (if not for that, actual attempt would have occurred)."¹ The wording here can be confusing to raters and patients. What is "a circumstance" in this context? And what is meant by "if not for that?" Does "that" refer to the intervening circumstance or the self-injurious act? The point here is that the rater has to stop because the wording is ambiguous. Some of the illustrations under this category are also problematic. For example, a person with a noose around his neck could be engaged in autoerotic asphyxiation with no intent to die. Classifying this case as an interrupted attempt would be an error.

"Aborted" or "Self-Interrupted Attempt." This category is defined as one in which the patient "begins to take steps toward making a suicide attempt, but stops themselves before they actually have engaged in any self-destructive behavior."¹ Use of the plural (e.g., they and themselves) is not simply poor grammar; it suggests that more than one person has to be

involved—as in a suicide pact. For an obsessional rater, it could be interpreted to *exclude* situations when only one person makes an attempt, posing yet another classification error. We, the authors of this paper, have seen examples of this interpretation by patients.

Use of "he/she." On one version of the C-SSRS, the timeframe for the questions references the time "he/she felt most suicidal." Use of gender-specific language such as this can cause problems if the person is intersex (e.g., Klinefelter's or Turner's syndrome) or gender neutral. The standard today is gender neutrality in rating scales, a standard the S-STS and ISST-Plus have both consciously adopted.

4. Does the C-SSRS map to the FDA 2012 classification algorithm?

Since the FDA-CASA 2012 adopted category titles and definitions that were similar to the C-SSRS, one would expect to see compatibility going forward of data collected through years of use of the C-SSRS. Unfortunately, this is not the case since the FDA Draft Guidance document, in its update of 2012, actually made the C-SSRS incompatible with several of the 2012 FDA classification categories. The FDA 2012 Draft Guidance document states on page 6, lines 217–218, "Direct classification into the 11 preferred terms (see Appendix A): Use of the C-SSRS instrument accomplishes this goal directly" and on page 5, lines 177–179, "The direct classification of information collected in the C-SSRS interview into these 11 categories, along with integration of information about the event from other sources, renders it unnecessary to conduct any other classification step."⁶ Unfortunately these statements are factually inaccurate in substantial measure.

As shown in Table 3, there are numerous inconsistencies between the C-SSRS and the FDA-CASA 2012. Category titles are consistent across the C-SSRS and FDA-CASA 2012 for only four of the 11 categories (36%). Definitions are

consistent for eight of the 10 categories where the C-SSRS provides definitions (80%). However, probes tend to be inconsistent, matching FDA-CASA 2012 definitions for only two of the 10 categories where probes are used (20%).

When the mismatches in Table 3 are applied to the possible combinations in Table 1, the result is Table 4. The last column of Table 4 illustrates where the FDA-CASA 2012 titles and definitions and the C-SSRS titles, definitions, and questions all capture the same combination number. Four (12.5%) were captured, 25 (78.125%) were not captured, and three (9.375%) were overlooked and not considered as options.

Specific incompatibilities. Some of the specific incompatibilities and how they have the potential to increase error in classification, depending on whether the rater uses the C-SSRS title, definition, or probe or some combination of the three, are further detailed below.

"Wish to Be Dead." This C-SSRS category should align with the FDA-CASA 2012 category "Passive suicidal ideation: wish to be dead."⁶ The C-SSRS and FDA definitions are similar, both addressing wishes to be dead or not alive anymore or to fall asleep and not wake up.¹⁶ The C-SSRS probe question, however, does not inquire into the potential of a patient experiencing "a wish to be [...] not alive anymore."¹ This omission might seem like a small detail, but it has the potential to narrow the application of this category resulting in under-identification (type II error) on the C-SSRS.

"Non-Specific Active Suicidal Thoughts." This C-SSRS category should match the FDA-CASA 2012 category "Active suicidal ideation: non-specific (no method, intent or plan)."⁶ The C-SSRS probe, however, is not limited to the kind of suicidal phenomena described in the FDA definition. In fact, the probe, "Have you actually had any thoughts of killing yourself?"¹¹ could refer to *specific* as well as *non-specific*

TABLE 3. Mismatches between C–SSRS and FDA-CASA 2012

C–SSRS Title	FDA-CASA 2012 Title	C–SSRS and FDA-CASA 2012 Titles Match?	C–SSRS and FDA-CASA 2012 Definitions Match?	C–SSRS Probe Consistent?
<i>SUICIDAL IDEATION</i>				
“Wish to be Dead”	“Passive suicidal ideation: wish to be dead”	Partially ¹	Yes	Partially ¹
“Non-Specific Active Suicidal Thoughts”	“Active suicidal ideation: non-specific (no method, intent: or plan)”	Partially ²	Yes	No ²
“Active Suicidal Ideation with Any Methods (Not Plan) Without Intent to Act”	“Active suicidal ideation: method but no intent or plan”	Yes	Yes	No ³
“Active Suicidal Ideation With Some Intent to Act, Without Specific Plan”	“Active suicidal ideation: method and intent, but no plan”	No ⁴	Yes	No ⁴
“Active Suicidal Ideation with Specific Plan and Intent”	“Active suicidal ideation: method, intent, and plan”	No ⁵	Yes	No ⁵
<i>SUICIDAL BEHAVIOR</i>				
“Actual Attempt”	“Suicide attempt”	Yes	Yes ⁶	No ⁶
“Interrupted Attempt”	“Interrupted suicide attempt”	Yes	Yes	Yes
“Aborted Attempt”	“Aborted suicide attempt”	Mostly ⁷	Yes	Yes
“Preparatory Acts or Behavior”	“Preparatory acts toward imminent suicidal behaviors”	No ⁸	No ⁸	No
“Suicide”	“Completed suicide”	Yes ⁹	No definition	No probe
“Non-Suicidal Self-Injurious Behavior”	“Self-injurious behavior without suicidal intent”	No	No	No

¹ The C–SSRS title does not contain the FDA concept of “passive ideation.” The probe does not have any reference to a wish to not be alive. *,**

² The C–SSRS title omits the FDA reference to “no method, intent or plan.” While the definition is consistent, the probe is ambiguous. It could refer to specific or non-specific thoughts. *,**

³ The C–SSRS probe only asks a “how” (method) question. It could elicit an answer that is inconsistent with the FDA category since it does not specify no intent or plan. *,**

⁴ The C–SSRS title is inconsistent with the FDA title in that it does not mention the requirement of a “method.” The C–SSRS probe also fails to address the method requirement. In addition, instead of requiring there be “no plan,” the C–SSRS requires no “specific” plan. This narrows the concept of a “plan.”**,**

⁵ Again, the C–SSRS title is inconsistent with the FDA title in that it does not mention the requirement of a “method.” It also differs from the FDA-CASA 2012 in requiring a “specific” plan. Further the probe suggests that it is enough for the patient to have “started” to work out details of a plan, a broader window than that conveyed in the C–SSRS and FDA CASA definitions, which require details “fully or partially” worked out. *,**

⁶ Definitions are generally similar although C–SSRS definition is much wordier. Probe questions with open-ended answers could elicit a suicide attempt or non-suicidal self-injurious behavior. *,**

⁷ The C–SSRS uses “Self-Interrupted” as well as “Aborted” title whereas the FDA-CASA 2012 only uses “Aborted.”**,**

⁸ The C–SSRS differs from the FDA-CASA 2012 in not adding “imminent” as a qualifier in the title although the C–SSRS does add this qualifier in its definition. The C–SSRS definition is more expansive than that provided by the FDA-CASA 2012 since it does not indicate that such behaviors need to stop short of a suicide attempt, an interrupted suicide attempt, or an aborted suicide attempt. *,**

⁹ This category is only present on the C–SSRS “Since Last Visit” Version. The title “Suicide” is not consistent with the FDA-CASA 2012 title “Completed suicide” in that the word on its own could be interpreted to mean completed or not completed. Unlike the FDA-CASA 2012, the C–SSRS does not provide a definition for this category. *,**

* 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.

** 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.

C–SSRS: Columbia–Suicide Severity Rating Scale; FDA-CASA 2012: United States Food and Drug Administration–Classification Algorithm for Suicide Assessment

TABLE 4. Possible combinations of suicidal Ideation and missing combinations between the C-SSRS and the FDA-CASA 2012 based on category titles, definitions, and probe questions

Possible Combination Number	Passive Ideation	Active Ideation	Method	Intent	Plan	Combinations Not Captured By FDA-CASA 2012 Titles	Combinations Not Captured By FDA-CASA 2012 Definitions	Combinations Not Captured By C-SSRS Titles	Combinations Not Captured By C-SSRS Definitions	Combinations Not Captured By C-SSRS Probe Questions	Do the FDA-CASA 2012 Titles & Definitions and the C-SSRS Titles, Definitions, & Probe Questions ALL Capture This Combination?
1	0	0	0	0	0						Yes
2	0	1	0	0	0						Yes
3	0	0	0	1	0	✓	✓	✓	✓	✓	Overlooked
4	0	1	0	1	0	✓					No
5	0	0	1	0	0	✓	✓	✓	✓		No
6	0	1	1	0	0						Yes
7	0	0	1	1	0	✓	✓	✓	✓		No
8	0	1	1	1	0		✓	✓	✓		No
9	0	0	0	0	1	✓	✓	✓	✓	✓	Overlooked
10	0	1	0	0	1	✓	✓	✓	✓		No
11	0	0	0	1	1	✓	✓	✓	✓	✓	Overlooked
12	0	1	0	1	1	✓					No
13	0	0	1	0	1	✓	✓	✓	✓		No
14	0	1	1	0	1	✓	✓	✓	✓		No
15	0	0	1	1	1	✓	✓	✓	✓		No
16	0	1	1	1	1		✓	✓	✓		No
17	1	0	0	0	0						Yes
18	1	1	0	0	0	✓	✓	✓	✓		No
19	1	0	0	1	0	✓	✓	✓	✓		No
20	1	1	0	1	0	✓	✓	✓	✓		No
21	1	0	1	0	0	✓	✓	✓	✓		No
22	1	1	1	0	0	✓	✓	✓	✓		No
23	1	0	1	1	0	✓	✓	✓	✓		No
24	1	1	1	1	0	✓	✓	✓	✓		No
25	1	0	0	0	1	✓	✓	✓	✓		No
26	1	1	0	0	1	✓	✓	✓	✓		No
27	1	0	0	1	1	✓	✓	✓	✓		No
28	1	1	0	1	1	✓	✓	✓	✓		No
29	1	0	1	0	1	✓	✓	✓	✓		No
30	1	1	1	0	1	✓	✓	✓	✓		No
31	1	0	1	1	1	✓	✓	✓	✓		No
32	1	1	1	1	1	✓	✓	✓	✓		No
Total	32					26	26	26	26	3	Yes: 4 of 32; No: 25 of 32; Overlooked: 3 of 32

Note 1: This table summarizes most of the other tables into one. If the titles, definitions, and probe questions were consistent within the C-SSRS titles, definitions, and probe questions and consistent within the FDA-CASA 2012 titles and definitions, then the check marks in the columns titled “Combinations Not Captured...” should be consistent across every row. But they are not.

Note 2: There are ambiguities within the components of the FDA-CASA 2012 title, within the components of the FDA-CASA 2012 definition, within the components of the C-SSRS title, within the components of the C-SSRS definition, and within the components of the C-SSRS probe questions, and further ambiguities between the wording across the titles, definitions, and probe questions. This table ignores the ambiguities within the components of the titles, definitions, and probe questions and uses the best case scenario for interpretation.

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

thoughts, thereby undermining the meaning of the category and potentially inflating the rate intended for capture in this category (type I error). This C-SSRS flaw is exposed in the cross-validation study of three suicidality scales by Sheehan et al.²⁹

“Active Suicidal Ideation With Any Methods (Not Plan) Without Intent to Act.” This C-SSRS category should match the FDA-CASA 2012 category, “Active suicidal ideation: method, but no intent or plan.”⁷⁶ The definitions and examples appear to correspond. However, the C-SSRS probe, “Have you been thinking about how you might do this?”⁷¹ has the potential to elicit responses that do not directly relate the FDA category and could go beyond it. The problem here is the use of the word “this,” a word that in the context is ambiguous and could refer to any number of phenomena, causing either type I or type II errors.

“Active Suicidal Ideation With Some Intent to Act, Without Specific Plan.” This C-SSRS category differs from the FDA-CASA 2012 category, “Active suicidal ideation: method and intent, but no plan.”⁷⁶ in three ways. First, the C-SSRS title does not indicate the need for a method, a requirement that is clearly made on the FDA-CASA 2012. Second, the title requires the absence of a “specific” plan, whereas the corresponding FDA-CASA 2012 title and definition only require the absence of a plan. This difference could lead to under-identification of this category, a type II error. Third, the C-SSRS probe question is so broad that it has the potential to capture suicidal phenomena well beyond that covered by the definition, leading to over-identification of this category, a type I error.

“Active Suicidal Ideation With Specific Plan and Intent.” The title for this C-SSRS category should match the title for the FDA-CASA 2012 category, “Active suicidal ideation: method, intent, and plan.”⁷⁶ However, in contrast to the FDA-

CASA 2012 title, it does *not* indicate the need for a method. Additionally, the C-SSRS probe questions, “Have you started to work out or worked out the details of how to kill yourself?” and “Do you intend to carry out this plan?”⁷¹ do not correspond with the FDA-CASA 2012 definition, which indicates that the patient would need to have gone beyond merely starting to work out the details and “have details of [a] plan fully or partially worked out.”⁷⁶ This difference could lead to over-identification of cases in this category.

“Actual Attempt.” The definition for this C-SSRS category should match that for the FDA category “Suicide attempt.” However, the C-SSRS definition requires the behavior to be “thought of as a method to kill oneself”⁷¹ while the FDA definition does not include this requirement. Further potential incompatibility is introduced by the C-SSRS qualification that the behavior be “in part” thought of as a method to kill oneself.¹ This could mean at least in part or wholly in part. It should specifically have stated “at least in part” to avoid to this ambiguity. This qualification not only affects the interpretation of the C-SSRS category, but it also distances it further from the FDA-CASA 2012.

“Self-Injurious Behavior Without Suicidal Intent.” This C-SSRS category is tucked into a probe, “Has subject engaged in non-suicidal self-injurious behavior?” under “Actual Attempt.”⁷¹ To find the definition, one needs to review another probe, one that may or may not be asked depending on the answer to the first question under the “Actual Attempt” category. This probe asks if the attempt was done “purely for other reasons / without ANY intention of killing yourself (like to relieve stress, feel better, get sympathy, or get something else to happen)?”⁷¹

Compared to the FDA definition (“Self-injurious behavior associated with no intent to die. The behavior is intended purely for other reasons,

either to relieve distress [often referred to as self-mutilation [e.g., superficial cuts or scratches, hitting or banging, or burns]] or to effect change in others or the environment.”),⁶ the C-SSRS definition is somewhat more expansive, adding behaviors “feel better,” “get sympathy,” and the rather ambiguous “to get something else to happen.”⁷¹

“Aborted” or “Self-Interrupted Attempt.” As described in above in section 3, use of the word “themselves” in the definition for the C-SSRS category, “Aborted Suicide Attempt,”⁷¹ is ambiguous and may be interpreted by some that more than one person has to be engaged for the act to be an aborted attempt. This is not the case with the corresponding FDA-CASA 2012 category, “Aborted Suicide Attempt,” where no ambiguity is present.⁶

“Preparatory Acts or Behavior.” The C-SSRS category, “Preparatory Acts or Behavior,”⁷¹ differs from the FDA-CASA 2012 category, “Preparatory acts toward imminent suicidal behaviors,”⁷⁶ in its placement and hence emphasis on the word “imminent.” Whereas the FDA-CASA 2012 category has this word in the title, the C-SSRS only has it in the definition. Furthermore, while the FDA-CASA 2012 definition clearly differentiates this category from other categories, indicating that the acts or behaviors stop short of a suicide attempt, an interrupted suicide attempt, or an aborted suicide attempt, the C-SSRS probe question does not make this distinction clearly. As a result patients who start an attempt, even if the attempt is interrupted or aborted, may be included in this category on the C-SSRS, thereby inflating its endorsement. This C-SSRS probe question wording conflicts with the definition wording.

“Completed Suicide” or “Suicide.” Although this category (labeled “Suicide”) is provided in the current version (downloaded from C-SSRS website on February 5, 2014¹⁵) of the

C-SSRS, it is not present in some other versions (C-SSRS Lifetime/Recent Version, C-SSRS Screen Version, C-SSRS Risk Assessment Version, and C-SSRS Baseline Version). Moreover, the category, when it is present, is not defined, leaving clinicians to assume but not be entirely sure that the FDA-CASA 2012 definition for a completed suicide is what the authors intended.

5. Are versions of the C-SSRS consistent? There are several inconsistencies between versions of the C-SSRS. Some versions, for example, do not include a “Completed Suicide” category. There are also version control problems. For example, during the months of December 2013 and January 2014, we carefully monitored the Columbia University website for different C-SSRS versions and found that the versions changed repeatedly (sometimes daily) and remained inconsistent with each other and with the earlier 1/14/09 version, and yet they all retained the 1/14/09 version date. We observed that problems with the 1/14/09 version we had identified a few months earlier were gone but new problems were introduced—all under the same version date. This lack of version control does not meet regulatory requirements for translation and linguistic validation.³⁶

Another inconsistency in versions is that the instruction, “Answer for actual attempts only,”³¹ placed above the “Actual Lethality/Medical Damage” section, is missing in the Lifetime/Recent Version and Risk Assessment Version of the scale. This omission could lead to inconsistencies in the way lethality/medical damage is rated. For the versions *without* the instruction, the rater might reasonably include damage from non-suicidal self-injurious behavior as well damage from actual attempts. After all, the probe for non-suicidal self-injurious behavior is placed in the “Suicidal Behavior” section just above the “Actual Lethality/Medical Damage”

section. For the versions with the instruction, however, damage from non-suicidal behavior would not be allowed. These differences have the potential to create inflation of lethality/medical damage in some versions but not others.

The C-SSRS dated June 17, 2007,³⁷ had a broad range of six passive suicidal ideation questions (including “Better Off Dead”), while the current versions have eliminated most of these questions and substituted others that were not in the 2007 version into this category. Is it justifiable to integrate as one, into the same FDA database, questions that are worded differently?

Problems going forward. The inconsistencies outlined above make the C-SSRS incompatible with the updated FDA-CASA 2012 and incompatible with the S-STS and the ISST-Plus. This situation poses a dilemma. Revising the C-SSRS to correct all the flaws identified here would make a new corrected C-SSRS incompatible with the current versions, rendering it impossible to merge current data for a particular drug or research study with new data going forward. Indeed, the argument could be made that any findings resulting from application of the current C-SSRS should be re-evaluated in light of the flaws identified above and the FDA’s 2012 revised, more exacting draft guidance standards. These incompatibilities cannot be remedied by retroactive reanalyses and correction in existing datasets. In addition, in accordance with regulatory standards, any newly corrected C-SSRS along the lines above would then need to be properly re-validated.

CONCLUSION

The C-SSRS has been endorsed by the FDA as the gold standard for assessment of suicidal ideation and behavior in clinical trials. While the instrument makes an effort to provide standardized categories and definitions, it falls short of widely

accepted standards. Its psychometric properties, moreover, have only been evaluated in a limited way to date. The most serious criticisms we have of the scale are as follows: 1) it is not comprehensive, which can lead to over-endorsements; 2) it does not cover the range of suicidal ideation or behavior; 3) the categories are not well defined, and wording in many cases is ambiguous and imprecise; and 4) many titles, definitions, and probes do not align with the updated FDA-CASA 2012 as documented in 2012 Draft Guidance. We see these flaws as having the potential to lead to endless arrays of type I and type II errors.

These findings suggest that C-SSRS, at the very least, requires substantial revision and reconstruction. What are internal review board members, data safety monitoring board members, and peer reviewers for journal articles and grants to do if they are aware of these flaws in the C-SSRS and they have to adjudicate approval for use for funding or for publication? In its current form, we believe the C-SSRS is a liability. In our opinion, the field of suicide assessment cannot scientifically move forward until these issues are resolved. We leave the reader with these unanswered questions and dilemmas. We believe we all have a responsibility to our suicidal patients to fix these problems and to resolve these dilemmas. How best to do this remains unresolved.

REFERENCES

1. 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.
2. Brent, DA, Greenhill L, Compton S, et al. The Treatment of Adolescent Suicide Attempters study (TASA): predictors of suicidal events in an open treatment trial. *J Am Acad Child Adolesc Psychiatry*. 2009;

- 48:987–996.
3. Oquendo MA, Halberstam B, Mann, JJ. Columbia suicide history form. In: Rush J, First MB, Black D (eds). *Handbook of Psychiatric Measures, Vol. 2*. Washington, DC: American Psychiatric Press Inc.; 2008:241–242.
 4. Posner K, Oquendo MA, Gould M, et al. Columbia Classification 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.
 5. 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.
 6. 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.
 7. Roth M. Max Hamilton: a life devoted to psychiatry. In: Bech P, Coppen A (eds). *The Hamilton Scales*. Berlin: Springer-Verlag; 1990:1–9.
 8. Worboys M. The Hamilton Rating Scale for Depression: The making of a “gold standard” and the unmaking of a chronic illness, 1960–1980. *Chronic Illness*. 2013;9(3):202–219.
 9. Harkavy Friedman JM, Asnis GM. Assessment of suicidal behavior: a new instrument. *Psychiatr Ann*. 1989;19:382–387.
 10. Lindenmayer JP1, Czobor P, Alphas L, et al. InterSePT Study Group. The InterSePT scale for suicidal thinking reliability and validity. *Schizophr Res*. 2003;63:161–170.
 11. Osman A, Bagge CL, Gutierrez PM, et al. The Suicidal Behaviors Questionnaire-Revised (SBQ-R): validation with clinical and nonclinical samples. *Assessment*. 2001;8:443–454.
 12. Beck AT, Kovacs M, Weissman A. Assessment of suicidal intention: the Scale for Suicide Ideation. *J Consult Clin Psychol*. 1979;47(2):343–352.
 13. Coric V, Stock EG, Pultz J, et al. Sheehan Suicidality Tracking Scale (S-STSS): preliminary results from a multicenter clinical trial in generalized anxiety disorder. *Psychiatry (Edmont)*. 2009;6:26–31.
 14. Crosby A, Ortega L, Melanson C. *Self-Directed Violence Surveillance: Uniform Definitions and Recommended Data Elements, version 1.0*. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Injury Prevention and Control; 2011.
 15. Center for Suicide Risk Assessment. About the C–SSRS. Columbia University Medical Center Website. http://www.cssrs.columbia.edu/about_cssrs.html. Accessed October 1, 2014.
 16. Chappell P, Feltner DE, Makumi C, et al. Initial validity and reliability data on the Columbia–Suicide Severity Rating Scale. Letter to the Editor. *Am J Psychiatry*. 2012;169:662–663.
 17. Brent DA, Emslie GJ, Clarke GN, et al. Predictors of spontaneous and systematically assessed suicidal adverse events in the Treatment of SSRI-Resistant Depression in Adolescents (TORDIA) study. *Am J Psychiatry*. 2009;166:418–426.
 18. Mundt JC, Greist JH, Gelenberg AJ, et al. Feasibility and validation of a computer-automated Columbia–Suicide Severity Rating Scale using interactive voice response technology. *J Psychiatr Res*. 2010;44(16):1224–1228.
 19. 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.
 20. Vandeppeer M. Health Policy Advisory Committee on Technology. Technology Brief. Columbia Suicide Severity Rating Scale. HealthPact Wmerging Health Technology. State of Queensland, Australia. August 2012. <http://www.health.qld.gov.au/healthpact/docs/briefs/WP114.pdf>. Accessed October 1, 2014.
 21. Bagby RM, Ryder AG, Schuller DR, et al. The Hamilton Depression Rating Scale: Has the Gold Standard Become a Lead Weight? *Am J Psychiatr*. 2004;161:2163–2177.
 22. Alphas L. Assessment of suicidal ideation, behavior and risk. Presented to the Suicide Assessment Working Group at the 52nd American College of Neuropsychopharmacology Annual Meeting. Hollywood, FL: December 2013.
 23. Guilford JP. *Fundamental Statistics in Psychology and Education, Fourth Edition*. New York: McGraw-Hill; 1965.
 24. Spector PE. *Summated Rating Scale Construction: An Introduction*. Sage University Papers. Series Quantitative Applications in the Social Sciences, 82. Newbury Park CA: Sage Publications;1992.
 25. Clark LA, Watson D. Constructing validity: basic issues in objective scale development. *Psychologic Assess*. 1995;7:309–319.
 26. Haynes SN, Richard DCS, Kubany ES. Content validity in psychological assessment: a functional approach to concepts and methods. *Psychologic Assess*. 1995;7:238–247.
 27. Giddens JM, Sheehan DV. Do the five combinations of suicidal ideation in the FDA 2012 Draft

- Guidance document and the C-SSRS adequately cover all suicidal ideation combinations in practice? *Innov Clin Neurosci*. 2014;11(9-10):172-178.
28. 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.
 29. Sheehan DV, Alphas L, Mao L, et al. Comparative validation of the S-STS, the ISST-Plus, and the C-SSRS for assessing the suicidal thinking and behavior FDA 2012 suicidality categories. *Innov Clin Neurosci*. 2014;11(9-10):32-65.
 30. Columbia University Psychiatry Department. What we really know about suicide and medications [video]. March 28th, 2012; 54:32 minutes into the video. Available at: <http://www.veomed.com/va041172402012>. Accessed January 2014.
 31. 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.
 32. 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.
 33. Grand Rounds 2011 [video]. Child Center of New York University (NYULMC). October 28, 2011. See at 48:20 minutes. Accessed online November 20, 2012. Available on request from nicole.aujero@nyumc.org.
 34. Boudrot A, Sheehan DV, Acquardo C. Lost in translation: translatability of psychiatric terms—the example of the Mini-International Neuropsychiatric Interview (M.I.N.I.). Poster presentation. International Society for Quality of Life Research 20th Annual Conference. Miami, FL: October 9–12, 2013.
 35. Andrich D. An elaboration of Guttman scaling with Rasch models for measurement. In: Tuma NB (ed). *Sociological Methodology*. San Francisco, CA: Jossey-Bass; 1985:33–80.
 36. Wild D, Eremenco S, Mear I, et al. Multinational trials: recommendations on the translations required, approaches to using the same language in different countries, and the approaches to support pooling the data: the ISPOR Patient-Reported Outcomes Translation and Linguistic Validation Good Research Practices Task Force Report. *Value in Health*. 2009;12(4):430–440.
 37. Posner K. Suicidality issues in clinical trials: Columbia Suicidal Adverse Event Identification in FDA Safety Analyses. Division of Metabolism and Endocrinology Products Advisory Committee Meeting. June 13, 2007. ■