

Columbia Classification Algorithm of Suicide Assessment (C-CASA): Classification of Suicidal Events in the FDA's Pediatric Suicidal Risk Analysis of Antidepressants

Kelly Posner, Ph.D.

Maria A. Oquendo, M.D.

Madelyn Gould, Ph.D., M.P.H.

Barbara Stanley, Ph.D.

Mark Davies, M.P.H.

Objective: To evaluate the link between antidepressants and suicidal behavior and ideation (suicidality) in youth, adverse events from pediatric clinical trials were classified in order to identify suicidal events. The authors describe the Columbia Classification Algorithm for Suicide Assessment (C-CASA), a standardized suicidal rating system that provided data for the pediatric suicidal risk analysis of antidepressants conducted by the Food and Drug Administration (FDA).

Method: Adverse events (N=427) from 25 pediatric antidepressant clinical trials were systematically identified by pharmaceutical companies. Randomly assigned adverse events were evaluated by three of nine independent expert suicidologists using the Columbia classification algorithm. Reliability of the C-CASA ratings and agreement with pharmaceutical company classification were estimated.

Results: Twenty-six new, possibly suicidal events (behavior and ideation) that were not originally identified by pharmaceutical companies were identified in the C-CASA, and 12 events originally labeled as

suicidal by pharmaceutical companies were eliminated, which resulted in a total of 38 discrepant ratings. For the specific label of "suicide attempt," a relatively low level of agreement was observed between the C-CASA and pharmaceutical company ratings, with the C-CASA reporting a 50% reduction in ratings. Thus, although the C-CASA resulted in the identification of more suicidal events overall, fewer events were classified as suicide attempts. Additionally, the C-CASA ratings were highly reliable (intraclass correlation coefficient [ICC]=0.89).

Conclusions: Utilizing a methodical, anchored approach to categorizing suicidality provides an accurate and comprehensive identification of suicidal events. The FDA's audit of the C-CASA demonstrated excellent transportability of this approach. The Columbia algorithm was used to classify suicidal adverse events in the recent FDA adult antidepressant safety analyses and has also been mandated to be applied to all anticonvulsant trials and other centrally acting agents and nonpsychotropic drugs.

(*Am J Psychiatry* 2007; 164:1035–1043)

Antidepressant use by children and adolescents dramatically increased in recent decades (1, 2), with up to 8 million prescriptions written annually in the United States (3). However, the use of antidepressant drug treatment has been fraught with controversy because of questions regarding both efficacy and safety. Efficacy results from pediatric trials are mixed and difficult to interpret, largely because of methodological limitations and regulatory idiosyncrasies in determining what is an "effective" study (4–6). Furthermore, regulatory agencies in the United States and the United Kingdom raised concerns in 2003 about the emergence of suicidal thoughts or behaviors during antidepressant treatment in pediatric populations, which may have led to a recent decline in prescription rates (7, 8), rendering risk-benefit analyses even more challenging.

To evaluate the potential association between suicidality and antidepressants, the Food and Drug Administration (FDA) decided to undertake a meta-analysis to examine suicidal events from 24 randomized placebo-controlled pediatric trials of selective serotonin reuptake inhibitors (SSRIs) and other newer generation antidepressants. However, inconsistent labeling of potentially suicidal events was identified as a significant threat to accurate risk-assessment analyses. This concern first arose during an FDA review of one pediatric SSRI study, in which events suggestive of suicidality were labeled "emotional lability." Subsequent examination of suicidality data from the other eight pediatric antidepressant studies underscored the problem, with a notable example being a subject who slapped herself in the face and was deemed as having made a suicide attempt (Table 1). The FDA deter-

This article is featured in this month's AJP Audio, is the subject of a CME course, and is discussed in an editorial by Dr. Brent on p. 989.

TABLE 1. Examples of Difficulties in Adverse Event Labeling^a

Original Label	Original Investigator Text From Adverse Event Report
Personality disorder	[A] 10-year-old male exhibited symptoms of personality disorder of moderate severity and was discontinued. One day later, [the patient] attempted to hang himself with a rope after [a] dispute with his father. [The] investigator did not consider this a serious adverse event but rather part of the personality disorder.
Accidental overdose and neurosis	The overdose of six capsules of study medication was in fact intentional and in response to an argument with the subject's mother.
Medication error	Age 14: The patient took 11 tablets impulsively and then went to school...the patient denied that it was a suicide attempt.
Suicide attempt	[The patient] had thoughts of killing self but had no intention of acting on them.
Hostility	Age 10: Before his mother's call to the site and again after arguing with his stepfather, he wrapped a cord from the miniblinds around his neck, threatening to kill himself.
Emotional lability/suicide attempt	Age 14: The patient is reported to have engaged in an episode of "automutilation," where she slapped herself in the face.

^a These labels were given by the study clinicians in the pharmaceutical company trials. They were given prior to the implementation of C-CASA and reflect why reclassification was necessary. Some labels are more severe than they should be, and other labels are less severe than warranted.

mined that conclusions based on these data would be unreliable and might produce either a false signal that would result in unwarranted restriction of useful medications or an underestimation of risk and subsequent danger to the general public.

The problem of inconsistent nomenclature of suicidal ideation and behavior (suicidality) encountered in this data set is not unique. Indeed, the ongoing debate concerning nomenclature has perpetuated the use of multiple terms to refer to the same behavior, frequently with pejorative connotations (e.g., threat, gesture) and descriptors (e.g., "manipulative," "hostile," "nonserious") (9–12). Such variability in terminology has consequences that extend beyond imprecise communication, limiting comparison of epidemiological prevalence rates and hampering prevention efforts (13). Additionally, it undermines the validity of risk-benefit analyses.

To enhance interpretability of pediatric antidepressant trial data to be used in their risk analysis, the FDA commissioned a study by Columbia University/New York State Psychiatric Institute investigators to classify all events that could represent suicidality. The investigators developed a systematic approach to the categorization of potential suicidal adverse events covering the full spectrum of suicidality, rooted in consensus recommendations and empirical findings regarding suicide-related definitions (10, 12, 14–16).

The whole continuum of suicidality was included in the system, given evidence that manifestations along the spectrum are linked (17, 18). For example, evidence suggests that suicide attempts with intent to die are predictive of completed suicide (16, 18, 19), and individuals who engage in preparatory suicidal behaviors with intent to die are also at risk for future suicide attempts (20) and completion (21). Epidemiological and clinical studies of adolescents and adults have established that severe or pervasive suicidal ideation is a predictor of both future attempts (17, 22–25) and completed suicide (26). Moreover, Brown et al. identified passive thoughts about wanting to be dead as a risk factor for completed suicide (27). These studies provide

the links between manifestations of suicidal process despite well-documented differences between them (28).

In the present article, we describe the structure and reliability of the Columbia Classification Algorithm of Suicide Assessment (C-CASA), the classification system of suicidal adverse events that produced the data used by the FDA in their critical assessment of pharmacologic risk.

Method

C-CASA

The C-CASA is a classification system that utilizes definitions of suicidality derived from empirical findings on the phenomenology of suicidality and identified predictive and risk factors. The criteria for a suicide attempt include both self-injurious behavior and suicidal intent (at least some intention to commit suicide). Intent to die portends a risk for future suicide and repeated attempts (15, 18, 29, 30) and can be reliably obtained (27). Inclusion of intent in the definition of suicide allows a distinction between those who self-injure in an attempt to die and those who self-injure for purely other nonsuicidal reasons (e.g., to manage affect) (31). The C-CASA has eight categories that distinguish suicidal events from nonsuicidal events and indeterminate or potentially suicidal events (Table 2). C-CASA definitions and training examples are presented in Table 2. Figure 1 illustrates the boundaries between categories.

C-CASA Rating Guidelines

The C-CASA includes operationalized guidelines for inference of suicidal intent. "Clinically impressive" behavior or circumstances are used to infer suicidal intent when the stated intent is missing, unclear, or denied. For example, a highly lethal act that is clearly not an accident might mean that no other intent except suicide can be inferred (e.g., a gunshot to the head, jumping from a high-story building). An illustrative example was a case of self-immolation, which was a circumstance allowing inference of intent to classify the event a suicide attempt. Alternatively, inference of suicidal intent could also be based on two other pieces of data, including clinical circumstances such as the method used, number of pills ingested, and location of injury on the body. For example, cuts on the legs typically represent nonsuicidal self-injurious behavior. According to C-CASA guidelines, other relevant data that could be used included past history of suicide attempt, past history of self-injurious behavior/self-mutilation, and family history of suicide/suicide attempts.

TABLE 2. C-CASA Definitions and Training Examples

Classification/ Category	Definition	Training Examples
Suicidal events		
Completed suicide	A self-injurious behavior that resulted in fatality and was associated with at least some intent to die as a result of the act.	1) After a long argument with his girlfriend, which resulted in the end of their relationship, the patient collected a rope and rode his bike to an isolated area where he fatally hanged himself. A suicide note was later found. 2) After four documented attempts at suicide, the patient stole his uncle's gun and shot himself and was fatally injured.
Suicide attempt	A potentially self-injurious behavior, associated with at least some intent to die, as a result of the act. Evidence that the individual intended to kill him/herself, at least to some degree, can be explicit or inferred from the behavior or circumstance. A suicide attempt may or may not result in actual injury.	1) After a fight with her friends at school, in which they discontinued speaking with her, the patient ingested approximately 16 aspirin and eight other pills of different types on the school grounds. She said that she deserved to die, which was why she swallowed the pills. 2) The patient used a razor blade to lacerate his wrists, his antecubital fossae, and his back bilaterally. He told his therapist that the "the main objective was to stop feeling like that," and he knew that he could die but didn't care. According to the patient, he also ingested a bottle of rubbing alcohol because in his health class he heard "that the medulla will get more suppressed that way," thereby increasing the chances that he would be "successful" and die.
Preparatory acts toward imminent suicidal behavior	The individual takes steps to injure him- or herself, but is stopped by self or others from starting the self-injurious act before the potential for harm has begun.	1) The patient had run away from home overnight because his father had gone to school and retrieved a recent "bad" report card. He was fearful of his father's reaction. Upon his return home, a 5- to 6-hour argument with his parents ensued, and he took a vegetable (broad, sharp) knife and went to his room. He reported putting the knife to his wrist but never puncturing the skin. 2) The patient stated that he "couldn't stand being depressed anymore" and "wanted to die." He decided to hang himself. He tied a telephone cord to the door knob and placed the cord loosely around his neck. Then, he stopped himself and did not follow through with the attempt.
Suicidal ideation	Passive thoughts about wanting to be dead or active thoughts about killing oneself, not accompanied by preparatory behavior. ^a	1) Active: The patient reported to the doctor that he was thinking about hanging himself in the closet. He was taken to the hospital and admitted. 2) Passive: The patient reported ideas about wanting to be dead but denied acting on these feelings.
Nonsuicidal events		
Self-injurious behavior, no suicidal intent	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/banging, or burns) or to effect change in others or the environment.	1) The patient was feeling ignored. She went into the family kitchen where her mother and sister were talking. She took a knife out of the drawer and made a cut on her arm. She denied that she wanted to die at all ("not even a little"), but she just wanted them to pay attention to her. 2) The patient reported feeling agitated and anxious after a fight with her parents. She went into her room, locked the door, and made several superficial cuts on the inside of her arms. She stated that she felt relieved after cutting herself and that she did not want to die. She reported that she had done this before at times of distress and that it usually helped her feel better. 3) The patient was in class, where a test was about to begin, and stabbed himself with a pencil in order to be taken to the nurse's office. 4) A 14-year-old girl wrote her name on her arm with a penknife and said that she often does so in order to reduce her anxiety. 5) The patient was noted to have multiple superficial burns on his arms. Upon questioning, he denied trying to kill himself.
Other, no deliberate self-harm	No evidence of any suicidality or deliberate self-injurious behavior associated with the event. The event is characterized as an accidental injury, psychiatric or behavioral symptoms only, or medical symptoms or procedure only.	1) The patient had a cut on the neck from shaving. 2) The patient was hospitalized for worsening of OCD or depressive symptoms with no suicidal thoughts or actions or 3) aggressive behavior. 4) Hospitalization was because of an infection, rhinoplasty, or pregnancy.
Indeterminate or potentially suicidal events		
Self-injurious behavior, suicidal intent unknown	Self-injurious behavior where associated intent to die is unknown and cannot be inferred. The injury or potential for injury is clear, but why the individual engaged in that behavior is unclear.	1) The patient cut her wrists after an argument with her boyfriend. 2) The patient was angry at her husband. She took 10 to 15 diazepam tablets and flushed the rest down the toilet. Her husband called the police for help, and she was taken to the hospital. She was groggy and stayed overnight in the hospital. 3) A 9-year-old patient had spoken about suicide frequently. After learning that his baseball coach was retiring, he began scratching his arm with a pencil.

(continued)

TABLE 2. C-CASA Definitions and Training Examples (*continued*)

Classification/ Category	Definition	Training Examples
Not enough information	Insufficient information to determine whether the event involved deliberate suicidal behavior or ideation. There is reason to suspect the possibility of suicidality but not enough to be confident that the event was not something other, such as an accident or psychiatric symptom. An injury sustained on a place on the body consistent with deliberate self-harm or suicidal behavior (e.g., wrists), without any information as to how the injury was received, would warrant placement in this category.	1) A child who “stabbed himself in [the] neck with a pencil.” The event may have been deliberate as opposed to accidental, as suggested by “stabbed,” but not enough information was provided to determine whether the event was deliberate. 2) A cut on the neck.

^a If ideation is deemed inherently related to a behavioral act, a separate rating is not given. However, if there is no clear relationship to a behavioral event, a separate classification of ideation is warranted.

Data

Adverse event reports from 25 trials of antidepressant medications with a combined sample of 4,562 pediatric patients were included. Reports were provided by the FDA. Twenty-four trials were sponsored by pharmaceutical companies, and one was funded by the National Institute of Mental Health (NIMH) (32); however, data from that particular trial was subsequently utilized for a pediatric indication by a pharmaceutical company. Twenty-three trials were randomized controlled trials, and two were nonrandomized controlled trials. Participants were pediatric patients, ages 6 to 17 years, and clinical trials were conducted between 1983 and 2004. The treatment duration, across nine medications, ranged between 4 and 16 weeks. Among SSRI-medication trials, two were on citalopram, three on fluoxetine, one on fluvoxamine, six on paroxetine, and three on sertraline. Other newer generation antidepressants studies were three bupropion trials, one mirtazapine study, two nefazodone trials, and four venlafaxine trials. Psychiatric diagnoses treated were major depressive disorder (15 trials), obsessive-compulsive disorder (OCD) five trials), generalized anxiety disorder (two trials), social phobia (one trial), and attention deficit hyperactivity disorder (ADHD) two trials). Fifteen of the trials were conducted exclusively in the United States. The two nonrandomized controlled trials were 1) an open-label trial of bupropion for ADHD (N=17) and 2) a randomized withdrawal study of paroxetine for OCD (N=194). The FDA analysis (33) used a subset of events, classified by the C-CASA, from the 23 randomized controlled trials described previously and events from an additional federally funded trial (Treatment for Adolescent Depression Study). Events from the Treatment for Adolescent Depression Study were classified using the C-CASA but were not included in the present reliability study, since a different pool of raters was used and it was sponsored by NIMH.

Adverse Events

Pharmaceutical company identification of “possibly suicidal” events. The FDA requested that manufacturers of all nine antidepressants identify adverse events that could represent “possibly suicidal” events. Events were identified using an electronic text-string search of trial databases of patient data recorded by local study clinicians. Pharmaceutical companies were asked to search for any adverse events report that included the terms “suic overdos attempt,” “cut,” “gas,” “hang,” “hung,” “jump,” “mutilate,” “overdos,” “self-damage,” “self-harm,” “self-inflict,” “self-injur,” “shoot,” “slash” in the labeling of an event. The FDA permitted exclusion of obvious false positives (e.g., “gas” in “gastrointestinal”). The pharmaceutical companies were also asked to select a subset of events that were considered suicide attempts. No definitional criteria were given to categorize possibly

suicidal events and suicide attempts. The string search identified 114 possibly suicidal events; of these, 87 (76.3%) were considered suicide attempts by pharmaceutical companies.

Broadening of event search. To insure that all potentially suicidal events were identified, the scope of the search was broadened beyond those events originally identified by pharmaceutical companies to include all accidental injuries, overdoses, and serious adverse events, such as life-threatening events and hospitalizations. Inclusion of these additional events enabled a blinded review, since both suicidal and other adverse events were included. For classification, 427 potentially suicidal adverse events were included. Among these events, 114 were originally rated by pharmaceutical companies as possibly suicidal.

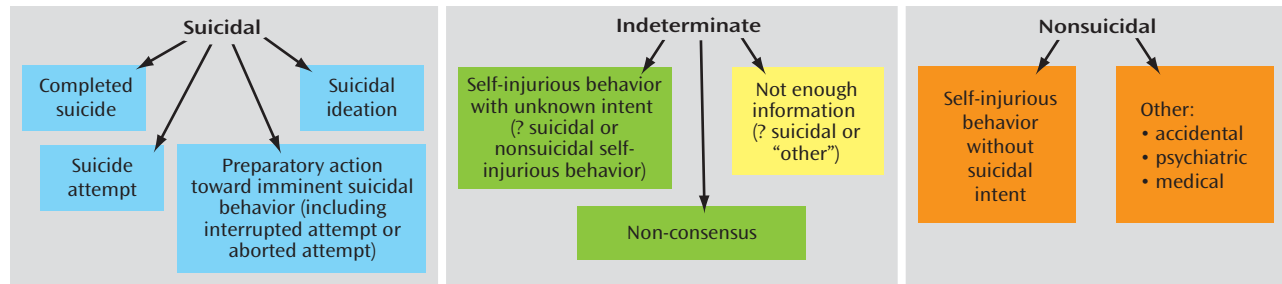
Adverse event narrative construction. Once adverse events were flagged by the string search, pharmaceutical companies composed narratives for each adverse event using data from case report forms, recorded by local study investigators during the course of the trials, and other sources, such as hospital records. When available, narratives included age, sex, history of suicidality, hospitalization status, current psychosocial stressors, and family history of suicide.

Blinding

Columbia University investigators developed comprehensive blinding procedures that removed information from all narratives that might have biased a classification decision. The FDA then implemented these procedures, removing all potential drug-identifying information, including the drug name, company/sponsor name, patient identification numbers, primary diagnosis, active or placebo arm, and all medication names and types, since treatment with other medications may be associated with a particular antidepressant side-effect profile. Case numbers that had no link to patient identifying information were randomly assigned to narratives by the FDA. Columbia University investigators further removed all original labels given by the pharmaceutical companies to categorize events (“preferred terms”) as well as adverse event labels given by participating investigators, including “serious” and “nonserious” determinations.

Expert Raters

Nine internationally recognized experts in suicide and suicide assessment were recruited as “raters.” Expert review of cases was needed for inference of suicidal intent based on the details of behaviors and related clinical data, since many narratives lacked stated suicidal intent. Expertise in suicidality was determined by relevant experience and publications. Panel members neither were involved in these industry trials nor were employed by Columbia University.

FIGURE 1. Suicidality Classification Scheme^a

^a Blue boxes=FDA “primary analysis” (includes events deemed suicidal). Blue+green boxes=FDA “sensitivity analysis” (includes any event that could possibly be suicidal).

Randomization and Expert Review Procedures

Event narratives were randomly distributed among raters using a balanced incomplete block design. Each event was classified by three raters; each triad of raters shared five cases. This randomization approach reduces rater burden without sacrificing precision in variance estimates (34).

Raters participated in a training teleconference to review classification parameters (categories, associated definitions, and case examples), followed by training reliability exercises prior to receiving narratives. Training exercises of each rater were reviewed for agreement with C-CASA definitions, and disagreements were discussed with the individual rater.

Each rater classified approximately 125 events. Raters could consult with a Columbia University trainer regarding the application of classification processes but were restricted from discussing specific events. Cases with discordant ratings were identified, and corresponding narratives were resent to raters. If ratings did not result in a unanimous agreement, a consensus discussion including the three raters assigned to assess the event was held and was led by another rater. The goal was to reach 100% agreement; otherwise, the event was classified as “indeterminate.” Final consensus classification determinations were provided to the FDA.

FDA Independent Audit of the C-CASA

To assess the reproducibility and reliability of the C-CASA methodology, four independent, nonsuicidologist FDA clinical reviewers were selected, including two pediatricians, one pharmacist, and one psychiatrist. Fifteen percent of the 427 event narratives were selected for review, with oversampling of “difficult-to-classify” cases. Raters received the same training and procedures as the expert panel. Audit results showed 89% agreement ($\kappa=0.84$) between audit ratings and expert ratings (35).

Statistical Analysis

Reliability coefficients were estimated with a random-effects linear model using the restricted maximum likelihood algorithm in SPSS 12.0 for Windows. Random effects modeled event-to-event, rater-to-rater, and error variation. Intraclass correlation coefficients (ICC) were estimated by the ratio of the variance because of the event divided by the total variance (sum of event-to-event, rater-to-rater, and error variation) (34). ICCs were estimated for each category.

Cohen's kappa was used to evaluate the agreement between pharmaceutical companies and C-CASA classifications. These analyses were conducted with only one event per subject. For subjects with multiple events, statistical calculations used the most severe event, which was chosen according to the severity hierarchy employed by the FDA for their unblinded analyses. This severity hierarchy was as follows: suicide attempt>preparatory behavior>suicidal ideation>self-injurious behavior intent un-

known>not enough information>self-injurious behavior, no suicidal intent. This approach identified 377 individual subjects, all of whom experienced one or more relevant adverse event. Only 50 individuals had more than one event, and most of those were accidental injuries.

Blinded examination of de-identified case records was considered exempt from review by the institutional review board of the New York State Psychiatric Institute and the Columbia University Department of Psychiatry.

Results

Frequencies of the 427 events according to C-CASA classifications are presented in Table 3. Completed suicides are not included, since none occurred in the pediatric trials.

Reliability of C-CASA

Excellent overall reliability (median ICC=0.89) was demonstrated among independent ratings of nine experts using the C-CASA. ICCs for the seven categories are presented in Table 3.

Of the 427 events, 366 (85.7%) had unanimous agreement among the three raters. Fifty-nine events (13.8%) had agreement between two of three raters, while two (0.47%) events had no agreement. Consensus discussions were held via teleconference whereby agreement was reached for all cases that were not unanimous.

Comparison With Pharmaceutical Companies

Discrepant cases. Thirty-eight discrepant cases were identified when comparing C-CASA with pharmaceutical company ratings (Table 4). Of these, 26 were new, possibly suicidal cases that were originally labeled by pharmaceutical companies as something other than suicidal (e.g., accidental injury). These cases were as follows: one suicide attempt, one suicidal preparatory act, 13 suicidal ideation events, four self-injurious behaviors with unknown intent, and seven cases without enough information but reason to suspect suicidality. The following is an example of a newly identified suicidal event: “The patient, age 11, held a knife to his wrist and threatened to harm himself. The patient was hospitalized with an acute exacerbation of major depressive disorder.” The original adverse event label was “exacerbation of major depres-

TABLE 3. Frequency and Reliability Results

Classification/Category	Frequency (N=427)	Percent (N=427)	Reliability of C-CASA Ratings ICCs (Mean=0.89)
Suicide attempt	36	8.4	0.81
Preparatory acts toward imminent suicidal behavior	8	1.9	0.89
Suicidal ideation	62	14.5	0.97
Self-injurious behavior, suicidal intent unknown	35	8.2	0.67
Not enough information	9	2.1	0.47
Self-injurious behavior, no suicidal intent	17	4.0	0.59
Other, no deliberate self-harm	260	60.9	0.93

TABLE 4. Agreement Between C-CASA and Pharmaceutical Company Ratings of Possible Suicidal Events and Suicide Attempts

Pharmaceutical Company Ratings	C-CASA Ratings		
	Yes	No	Total
Possibly suicidal events			
Yes	102	12	114
No	26	237	263
Total	128	249	377
Suicide attempts			
Yes	33	45	78
No	1	298	299
Total	34	343	377

sive disorder,” without an indication of suicidality from either the site investigator or pharmaceutical company. The new label was preparatory suicidal behavior. This event was discovered only because it was within a serious adverse event report of a hospitalization.

Twelve cases that were originally identified as potentially suicidal by pharmaceutical companies were classified as *not* potentially suicidal by C-CASA raters. These events were reclassified as psychiatric, involving no suicidality (N=2), accidental injury (N=1), and self-injurious behavior without suicidal intent (N=9).

Agreement on suicide attempts. Modest agreement was found between pharmaceutical company and C-CASA raters' classification of suicide attempts ($\kappa=0.53$ [SE=0.06]) (Table 4). Of their 114 possibly suicide-related events, pharmaceutical companies rated 78 (68.4%) as attempts, versus the C-CASA raters identifying 34 out of 128 (26.6%) as attempts. Forty-five of the 78 (57.7%) events classified as suicide attempts by the pharmaceutical company raters were not classified by C-CASA raters as suicide attempts. One suicide attempt was identified by C-CASA raters that had not been identified by pharmaceutical companies. Although the C-CASA identified more potentially suicidal cases overall, the rate of specific suicide attempts was lower.

Agreement on definitely suicidal cases. Agreement between C-CASA and pharmaceutical company ratings increased when comparing the broader C-CASA categorization of definitely suicidal events (attempts, preparatory acts, and suicidal ideation) with the pharmaceutical company rating of possibly suicidal cases ($\kappa=0.69$ [SE=0.04]). Thirty-two events identified as possibly suicidal by pharmaceutical companies were not classified as definitely suicidal by the C-CASA. Conversely, 15 newly identified definitely suicidal cases were identified by the C-

CASA. This C-CASA grouping was used by the FDA in their primary analysis (33).

Agreement on possibly suicidal cases. When comparing the broad nonspecific pooling of all categories that could possibly represent suicidality, there was good agreement between C-CASA (suicide attempts, preparatory behaviors, suicidal ideation, self-injurious behavior with unknown intent, and not enough information) and pharmaceutical company identification of possibly suicidal events ($\kappa=0.77$ [SE=0.04]) (Table 4). This C-CASA grouping was used in the FDA's "sensitivity analysis" to conservatively examine results that included anything that could have possibly represented suicidality (i.e., "worst case") (33). Thus, the C-CASA identified an increased number of possibly suicidal events in the data set overall.

Discussion

Classification of suicidal adverse events in 25 pediatric antidepressant trials with the C-CASA resulted in reliable classification of suicidal events. The C-CASA classification identified 38 discrepant cases, including events not previously deemed potentially suicidal (N=26) and those changed from suicidal to nonsuicidal (N=12). Furthermore, while C-CASA classification found more suicidal events, estimates of suicide attempts were significantly reduced. The new potentially suicidal events identified involved both suicidal ideation and behavior, across a range of classifications. Thus, when we expanded the search, many new suicidal events were found that had been missed by the pharmaceutical companies. However, of the suicidal events that the pharmaceutical companies identified, C-CASA classification resulted in a 50% reduction in the rate of suicide attempts. This reflects a tendency of the pharmaceutical companies to label any potentially suicidal event or self-injurious behavior as a suicide attempt

(e.g., suicidal ideation or a “slap in the face” labeled suicide attempt). These findings underscore the need for a standardized assessment of suicidality. Additionally, the need to expand the search for suicidal events as evidenced by the 26 newly found cases suggests that approaches currently employed in clinical trials lack sensitivity.

When comparing the C-CASA ratings with pharmaceutical company ratings, a relatively low level of agreement was found with more specific identification of suicidal occurrences, namely suicide attempts. Only when identifying a “suicidal range” or a broad nonspecific category of “possibly suicidal” was there better agreement. Pharmaceutical companies rated 45 events as suicide attempts that C-CASA raters did not. Thus, with respect to suicide attempts, reclassification with C-CASA would yield less of a hazard from the medication than if the original pharmaceutical ratings were used. Indeed, the FDA safety analysis that used these C-CASA ratings (33) found reduced risk estimates of suicidality in a depressed pediatric sample when compared with earlier FDA estimates that relied on the pharmaceutical labels (36). Additionally, a more precise risk estimate resulted (i.e., tighter confidence interval) using the C-CASA. These findings support the notion that misclassification may lead to overestimation of true risk (37). Such a change in risk estimation has clinical implications and likely affects risk-benefit analyses. Furthermore, the final FDA data set with the C-CASA ratings (33) included one-third (38/114) of cases that were different compared with the original data set (36), a substantially different sample. The use of data sets with imprecisely classified suicidal events can result in misleading findings, such as inaccurate risk and protective factors for suicidality.

The reliability of this classification approach was confirmed by the FDA's independent audit, which concluded that the C-CASA was “robust and reproducible” (35). The reliable use of this classification schema by nonsuicidologists reflects the transportability of this methodology. Notably, the FDA has mandated application of C-CASA to classify suicidal adverse events in adult antidepressant trials, as well as nonpsychotropic drug classes, and other centrally acting agents, including all anticonvulsants, cannabinoid 1 receptor (CB1R) inverse agonists for the treatment of obesity and metabolic disease. C-CASA classified data were used in the recent FDA investigation of an association between antidepressants and suicidality in adults (38).

Limitations and Future Directions

The study findings are limited by the quality of the available data describing adverse events. Descriptions of suicidal occurrences were variable and limited, particularly regarding intent. Furthermore, the expanded search for unidentified occurrences elucidated the inadequate quality of the elicitation and description of suicidal adverse events.

Although neither the C-CASA raters nor Columbia University investigators were responsible for subsequent

analysis using C-CASA ratings—by Hammad et al. (33) in the FDA's safety analysis, for example—some discussion of the limitations of these subsequent analyses is warranted. Suicidal adverse events were not systematically elicited but were revealed spontaneously, allowing the possibility of ascertainment bias. Subjects receiving active medication may be more likely to report suicidal occurrences than those on placebo because of increased contact with providers, consequent to other side effects. Such ascertainment bias is an alternate explanation for differential rates among subjects receiving drug treatment versus those receiving placebo found in the FDA safety analysis (33). In addition, improvement from active medication may lead subjects to discuss suicidal thoughts with their clinician for the first time, as opposed to such thoughts being caused by the medication.

Future intervention trials that prospectively and systematically monitor occurrence and emergence of suicidality with consistent methods of ascertainment would be informative. Such investigations would more optimally delineate the relationship between suicidal adverse events and antidepressant treatments as well as for any other treatment risk analysis. Improved assessment of suicidal events is necessary both to better inform research-derived risk-benefit analyses and to foster improved clinical management and identification. Accordingly, a prospective counterpart to this system, the Columbia Suicide Severity Rating Scale (39), is being widely used and frequently recommended by the FDA. The Columbia Suicide Severity Rating Scale is a tool designed to systematically assess and track suicidal adverse events (behavior and ideation) throughout any clinical trial as well as other settings.

The strength of this suicide classification system is, perhaps, in its ability to comprehensively identify suicidal events while limiting the overidentification of suicidal behavior. This classification system is research-based and can be applied in both clinical and research settings. Its use might result in more accurate identification of suicidality and more precise communication among researchers and clinicians, which would ultimately benefit treatment of suicidal individuals. The incorporation of research-supported, standardized suicidality terminology into psychiatric diagnostic manuals could also promote greater accuracy in communication between clinicians, allowing dissemination to a broad audience. Such a common language of suicide classification could be used in the same way that diagnostic criteria are currently used to provide a method for precise, widely understood communication.

Received June 16, 2006; revision received Dec. 15, 2006; accepted Jan. 14, 2007. From the Division of Child and Adolescent Psychiatry and the Department of Neuroscience, Columbia University/New York State Psychiatric Institute, New York; Department of Psychology, John Jay College of Criminal Justice, City University of New York, New York; and the Department of Epidemiology, Mailman School of Public Health, Columbia University, New York. Address correspondence and reprint requests to Dr. Posner, Division of Child Psychiatry, Columbia University/New York State Psychiatric Institute, 1051 Riverside Dr.,

Unit 74, New York, NY 10032; posnerk@childpsych.columbia.edu (e-mail).

Supported in part by NIMH grant P20 AA015630-03 to Dr. Stanley.

The authors thank Thomas Laughren, M.D. (FDA) and Tarek Ham-mad, M.D., Ph.D., M.Sc., M.S. (FDA), J. John Mann, M.D. (Columbia University, New York), David Shaffer, M.D. (Columbia University, New York), Laurence Greenhill, M.D. (Columbia University, New York), and the Expert Suicidality Panel: Annette Beautrais, Ph.D.; David Brent, M.D.; Greg Brown, Ph.D.; Cheryl King, Ph.D.; Peter Marzuk, M.D.; Alec Miller, Psy.D.; Patrick O'Carroll, M.D., M.P.H.; David Rudd, Ph.D.; Anthony Spirito, Ph.D.; and Kees van Heeringen, M.D., Ph.D. The authors also thank Marjorie Kleinman, M.S., Glenn A. Melvin, Ph.D., and S. Sonia Gugga, M.S.

CME Disclosure

Dr. Posner has received funding from the FDA to develop the suicidal classification system used in their antidepressant safety analysis. Subsequently, as part of an effort to help execute the FDA adult suicidal classification mandates that followed this FDA-sponsored pediatric study, Dr. Posner has had research support from GlaxoSmith-Kline, Forest Laboratories, Eisai, Inc., AstraZeneca, Johnson and Johnson, Abbott Laboratories, Wyeth, Organon USA, Bristol-Myers Squibb, Sanofi-Aventis, Cephalon, Novartis, Shire Pharmaceuticals, and UCB Pharma. Dr. Oquendo has received funding from NIMH, National Institute on Alcohol Abuse and Alcoholism, Moody's Foundation, American Foundation for Suicide Prevention, and Eli Lilly (unrestricted educational grant); she has also served as a consultant to Pfizer. Dr. Stanley has received fluoxetine/placebo for a randomized controlled trial from Eli Lilly. Mr. Davies is a shareholder with Merck, Pfizer, Wyeth, GlaxoSmithKline, Johnson and Johnson, Amgen, and Bard. Dr. Gould reports no competing interests.

APA policy requires disclosure by CME authors of unapproved or investigational use of products discussed in CME programs. Off-label use of medications by individual physicians is permitted and common. Decisions about off-label use can be guided by scientific literature and clinical experience.

References

1. Olsson M, Marcus SC, Weissman MM, Jensen PS: National trends in the use of psychotropic medications by children. *J Am Acad Child Adolesc Psychiatry* 2002; 41:514–521
2. Vitiello B, Zuckerman SH, Norquist GS: National estimates of antidepressant medication use among US children, 1997–2002. *J Am Acad Child Adolesc Psychiatry* 2006; 43:271–279
3. IMS Health America: National Prescription Audit: Information, Plymouth Meeting, Plymouth, Pa, Information Marketing Systems
4. Whittington CJ, Kendall T, Fonagy P, Cottrell D, Cotgrove A, Boddington E: Selective serotonin reuptake inhibitors in childhood depression: systematic review of published versus unpublished data. *Lancet* 2004; 363:1341–1345
5. Cheung AH, Emslie GJ, Mayes, TL: Review of the efficacy and safety of antidepressants in youth depression. *J Child Psychol Psychiatry* 2005; 46:735–754
6. Mann JJ, Emslie G, Baldessarini RJ, Beardslee W, Fawcett JA, Goodwin FK, Leon AC, Meltzer HY, Ryan JD, Shaffer D, Wagner KD: ACNP Task Force Report on SSRIs and suicidal behavior in youth. *Neuropsychopharmacology* 2006; 31:473–492
7. Rosack, J: New data show declines in antidepressant prescribing. *Psychiatr News* 2005; 40:1
8. Gibbons RD: Efficacy and safety of antidepressants for depression and suicide risk in youth and adults: results of new analyses, in the American College of Neuropsychopharmacology Annual Meeting, Hollywood, Fla, American College of Neuropsychopharmacology, 2006
9. Bille-Brahe U, Kerkhof A, DeLeo D, Schmidtke A: Definitions and termination used in the World Health Organization/EURO Multicentre Study, in *Suicidal Behaviour in Europe*. Edited by Schmidtke A, Bille-Brahe U, DeLeo D, Kerkhof A. Gottingen, Germany, Hogrefe and Huber, 2004, pp 11–14
10. O'Carroll PW, Berman AL, Maris RW, Moscicki EK, Tanney BL, Silberman MM: Beyond the Tower of Babel: a nomenclature for suicidology. *Suicide Life Threat Behav* 1996; 26:237–252
11. Goldston DB: Measuring Suicidal Behavior and Risk in Children and Adolescents. Washington, DC, American Psychological Association, 2003
12. Beck AT, Beck R, Kovacs M: Classification of suicidal behaviors: I. quantifying intent and medical lethality. *Am J Psychiatry* 1975; 132:285–287
13. Joe S: Self-directed violence definitions: a review of the scientific literature, in the Uniform Definitions of Self-Injury Conference, the National Center for Injury Prevention and Control, Centers for Disease Control and Prevention. Miami, Fla, 2004
14. Brown GK, Jeglic E, Henriques GR, Beck AT: Cognitive Therapy, Cognition and Suicide Behavior, in *Cognition and Suicide: Theory, Research, and Therapy*. Edited by Ellis TE. Washington, DC, American Psychological Association, 2006, pp 53–74
15. Goldsmith SK, Pellmar TC, Kleinman AM, Bunney WE (eds): Reducing Suicide: A National Imperative (Committee on Pathophysiology and Prevention of Adolescent and Adult Suicide, Board on Neuroscience and Behavioral Health, Institute of Medicine of the National Academies). Washington, DC, The National Academies Press, 2002
16. Oquendo MA, Galfalvy H, Russo S, Ellis SP, Grunebaum MF, Burke A, Mann JJ: Prospective study of clinical predictors of suicidal acts after a major depressive episode in patients with major depressive disorder or bipolar disorder. *Am J Psychiatry* 2004; 161:1433–1441
17. Brent DA, Perper JA, Moritz G, Allman C, Friend A, Roth C, Schweers J, Balach L, Baugher M: Psychiatric risk factors for adolescent suicide: a case-control study. *J Am Acad Child Adolesc Psychiatry* 1993; 32:521–529
18. Dorpat TL, Ripley H: The relationship between attempted suicide and committed suicide. *Compr Psychiatry* 1967; 8:74–79
19. Nordstrom P, Samuelsson M, Asberg M: Survival analysis of suicide risk after attempted suicide. *Acta Psychiatr Scand* 1995; 91:336–340
20. Marzuk PM, Tardiff K, Leon AC, Portera L, Weiner C: The prevalence of aborted suicide attempts among psychiatric in-patients. *Acta Psychiatr Scand* 1997; 96:492–496
21. Steer RA, Beck AT, Garrison B, Lester D: Eventual suicide in interrupted and uninterrupted attempters: a challenge to the cry-for-help hypothesis. *Suicide Life Threat Behav* 1988; 18: 119–128
22. Fergusson D, Doucette S, Glass KC, Shapiro S, Healy D, Hebert P, Hutton B: Association between suicide attempts and selective serotonin reuptake inhibitors: systematic review of randomised controlled trials. *Br Med J* 2005; 330:396–399
23. Wichstrom L: Predictors of adolescent suicide attempts: a nationally representative longitudinal study of Norwegian adolescents. *J Am Acad Child Adolesc Psychiatry* 2000; 39:603–610
24. Pfeffer C, Klerman GL, Hurt SW, Kakuma T, Peskin JR, Siefer CA: suicidal children grow up: rates and psychosocial risk factors for suicide attempts during follow-up. *J Am Acad Child Adolesc Psychiatry* 1993; 32:106–113
25. Lewinsohn PM, Rohde P, Seeley JR: Psychosocial risk factors for future adolescent suicide attempts. *J Consult Clin Psychol* 1994; 62:297–305
26. Brown GK, Beck AT, Steer RA, Grisham JR: Risk factors for suicide in psychiatric outpatients: a 20-year prospective study. *J Consult Clin Psychol* 2000; 68:371–377
27. Brown GK, Steer RA, Henriques GR, Beck AT: The internal struggle between the wish to die and the wish to live: a risk factor for suicide. *J Consult Clin Psychol* 2005; 162:1977–1979

28. Young EA, Coryell W: Suicide and the hypothalamic-pituitary-adrenal axis. *Lancet* 2005; 366:959–960
29. Hawton K, Cole D, O'Grady J, Osborn M: Motivational aspects of deliberate self-poisoning in adolescents. *Br J Psychiatry* 1982; 141:286–291
30. Beck RW, Morris JB, Beck AT: Cross validation of the Suicidal Intent Scale. *Psychol Rep* 1974; 34:445–446
31. Stanley B, Gameroff MJ, Michalsen V, Mann JJ: Are suicide attempters who self-mutilate a unique population? *Am J Psychiatry* 2001; 158:427–432
32. Emslie GJ, Rush AJ, Weinberg WA, Kowatch RA, Hughes CW, Carmody T, Rintelmann J: A double-blind, randomized, placebo-controlled trial of fluoxetine in children and adolescents with depression. *Arch Gen Psychiatry* 1997; 54:1031–1037
33. Hammad TA, Laughren T, Racosin J: Suicidality in pediatric patients treated with antidepressant drugs. *Arch Gen Psychiatry* 2006; 63:332–339
34. Shrout PE, Fleiss JL: Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979; 86:420–428
35. Iyasu S: Audit of the Columbia Suicidality Classification Project, in FDA Psychopharmacologic Drugs Advisory Committee and the Pediatric Advisory Committee Meeting, Bethesda, Md, FDA, 2004 (http://www.fda.gov/ohrms/dockets/ac/04/slides/2004-4065S1_07_FDA-Iyasu.htm) (Accessed Aug. 2006)
36. Mosholder A: Comparison between original ODS and current DNDP analyses of pediatric suicidality data sets, in Psychopharmacologic Drugs Advisory Committee and the Pediatric Advisory Committee Meeting. Bethesda, Md, FDA, 2004 (http://www.fda.gov/ohrms/dockets/ac/04/slides/2004-4065S1_09_FDA-Mosholder.htm) (Accessed Aug. 2006)
37. Jurek AM, Greenland S, Maldonado G, Church TR: Proper interpretation of non-differential misclassification effects: expected vs observations. *Int J Epidemiol* 2005; 34:680–687
38. Laughren TP: Overview for Meeting of Psychopharmacologic Drugs Advisory Committee (PDAC) Memorandum Department of Health and Human Services, Public Health Service. Bethesda, Md, Food and Drug Administration, Center for Drug Evaluation and Research, 2006 (<http://www.fda.gov/ohrms/dockets/ac/06/briefing/2006-4272b1-01-FDA.pdf>)
39. Posner K, Melvin GA, Stanley B, Oquendo MA, Gould M: Factors in the assessment of suicidality in youth. *CNS Spectr* 2007; 12: 156–162