Abstract

A Suicide Monitoring Board (SMB) evaluated putative events of suicidal behavior to determine whether they should be counted as outcome data (endpoints) in the International Suicide Prevention Trial (InterSePT) study. The InterSePT study compared clozapine and olanzapine prospectively, on reduction of suicidality in schizophrenia and schizoaffective disorder. SMB members, blinded to patient identification, evaluated packages of clinical information and the forms used to summarize the putative events and their context. The SMB members met regularly by international teleconference to discuss ratings they found problematic and to achieve consensus. Exactly 243 of the 980 patients enrolled (24.8%) were considered to have experienced 577 suicidal events. Of these events, 483 (83.7%) were confirmed by the SMB as real or valid endpoints. Before the consensus discussions, acceptable levels of agreement (κ = 0.52) were reached among the SMB members regarding ratings considered as valid suicidal events; agreement increased (κ = 0.64) when consensus ratings and those of on-site psychiatrists also blinded to treatment were also included. This article discusses our experience of evaluating schizophrenia patients for suicidality in general and in particular from reviewing chart information. The process of making clinical judgments regarding the seriousness of suicidal behavior in schizophrenia patients and their suicidal risk warrants further study.

Keywords: Schizophrenia, suicide risk, InterSePT, Clozaril, clozapine, Zyprexa, olanzapine, clinical trial.


Suicide is recognized as a significant concern among patients diagnosed with schizophrenia. Suicide deaths are conservatively estimated as occurring in a range of 4 percent (Inskip et al. 1998) to 10 to 13 percent (Caldwell and Gottesman 1990) of schizophrenia patients. Suicide attempt prevalence is about 30 percent (Gupta et al. 1998), and in schizoaffective disorders the lifetime attempt rate is close to 50 percent (Radomsky et al. 1999). Psychiatric case register studies found that suicides occurred 14 to 20 times more frequently among patients with schizophrenia (Newman and Bland 1991; Baxter and Appleby 1999), placing it in the highest risk category along with affective disorders, personality disorders, and (in males) substance abuse disorders.

The International Suicide Prevention Trial (InterSePT) study was a prospective, randomized, international, multicenter, 24-month study. The trial was designed to evaluate the effects of open-label clozapine and olanzapine on reduction of suicidality in schizophrenia or schizoaffective patients, deemed to be at high risk for committing suicide. The patients met DSM–IV criteria for schizophrenia or schizoaffective disorder. Patients had either attempted suicide or been hospitalized within the previous 3 years for management of suicidality, or had demonstrated within the previous week moderate to severe suicidal ideation with a depressive component, with or without command hallucinations to harm themselves. The detailed methodology has been described elsewhere (Meltzer 1999; Meltzer et al. 2000; Alphs et al. 2004; Lindenmayer et al. 2003), and the main findings are beginning to appear in the literature (Meltzer et al. 2003).

In this article we describe methodological issues relating to estimating suicidality as an outcome measure of interest in the trial, problems we encountered, our levels of interrater agreement, and our agreement with secondary outcome parameters. We also evaluate feasibility and examine the validity of retrospective researcher reviews of patient clinical charts to categorize levels of suicidality in these trials.

There is a great deal of literature exploring factors accompanying suicide risk in schizophrenia. Much of this

Send reprint requests to Dr. I. Sakinofsky, High Risk Consultation Clinic, Center for Addiction and Mental Health, 250 College Street, Toronto, ON, M5T 1R8, Canada; e-mail: isaac.sakinofsky@utoronto.ca.
information has been attained by retrospectively identifying risk factors among suicide victims, based on statistical analysis of cohorts of schizophrenia patients: almost none have been prospective. Some findings are controversial. For instance, conventional wisdom regards suicide risk as highest early in the course of the disorder (Mortensen and Juel 1993). However, major psychological autopsy studies in Finland (Heila et al. 1997) and Japan (Funahashi et al. 2000) show that suicide occurs at a high rate throughout the course of schizophrenia. Frequently, the risk peaks during an active phase of the disease (when two or more positive symptoms are present), with concomitant depression present in two-thirds of the patient population (Heila et al. 1997). This combination appears to be of even greater risk in women.

Paranoid symptoms, such as suspiciousness and delusions, tend to raise the risk of suicide, while negative symptoms tend to lower the risk (Fenton et al. 1997). Controversially, Allebeck et al. (1987) did not find a history of depression to be associated with increased suicide risk and therefore argued that suicide in schizophrenia was often impulsive and difficult to predict.

However, the importance of depression as a risk factor is supported in other studies (Bartels et al. 1992; Jones et al. 1994; Rossau and Mortensen 1997; Addington et al. 1998; Saarinen et al. 1999; Fenton 2000), as is hopelessness (Drake and Cotton 1986; Aguilar et al. 1997; Kaplan and Harrow 1999). Depression in schizophrenia patients can be independent of adverse life events, which tend to be documented less frequently in schizophrenia suicides (Heila et al. 1999).

Impulsivity is often found in association with substance misuse or dependency in schizophrenia patients (Gut-Fayand et al. 2001). However, prior communication of suicidal intent in the period leading up to the suicide is as frequent with suicide victims with schizophrenia (in half of cases) as it is in other suicides. In both groups, warnings are absent immediately preceding the suicide (Heila et al. 1998). Previous suicide attempts occurred in over 80 percent of schizophrenia patient suicides. Hallucinations commanding suicide were present in 10 percent of the cases of Heila et al. (1997).

Given the high prevalence of depression as a symptom in schizophrenia patients and its disconnection from acute life events, it is reasonable to suspect that the dysphoria felt by schizophrenia patients could relate to the realization of suffering from a chronic, socially impairing disease. Schwartz (2001), using research instruments developed by Amador et al. (1996) and Ward (1995), associated depressed mood with insight into the need for treatment of the disorder, and to the realization of its social consequences. He found that depression and patient self-awareness were positively correlated. This led to Schwartz' suggestion of a linear relationship, the "insight—demoralization/hopelessness/helplessness—depression—suicidality syndrome" (p. 403), in many schizophrenia patients.

In a large case-control investigation, Danish researchers matched 508 suicides among 9,156 patients (diagnosed with first episode schizophrenia), with 10 controls each (Rossau and Mortensen 1997). Males proved more at risk than females, but age was not a factor after controlling for duration of illness. Suicide risk was highest in the 6 months immediately following the first admission for schizophrenia. The relative risk (RR) of suicide rose with the number of admissions during the year preceding suicide, to an RR of 11 with eight or more admissions. During the first 5 days after discharge, the risk doubled that estimated during the first 5 days of the admission and declined after 28 days postdischarge. The authors suggest that the current admissions policy of shorter and more frequent admissions, the "revolving door," might be having a negative effect, predisposing patients to suicide risk. Belgian researchers conducted a case control study of 63 schizophrenia suicide victims under the age of 30 at admission (De Hert et al. 2001) and also found that male gender and frequent short hospitalizations were risk factors. Other important factors associated with suicide risk (with their odds ratios) were depression (36), a history of a highly lethal attempt (11), noncompliance with treatment (7), impulsive behavior (6), deliberate self-harm or parasuicide (5), and high premorbid intelligence (4). Protective factors were early onset of negative symptomatology (6) and a daily activity (4).

The above background risk factors for suicide in schizophrenia are helpful clinically and should be weighed in every case. No simple instrument existed during the study (1998–2001) that could be used with reliability and accuracy to specifically screen behavior in schizophrenia patients for suicide risk, so clinical judgment remained the gold standard. Estimating the seriousness of suicidal ideation essentially centers on clinically judging the level of patients' intent to harm themselves and their strength and will to control its expression. Similarly, suicidal behavior must be judged by estimating the force and immediacy of patients' intent to harm themselves and the degree to which it is driven by mental pain, determination (adamance), aggressiveness, lethality, and perturbation (Shneidman 1993; Kral and Sakinofsky 1994; Nasser and Overholser 1999; Cochrane-Brink et al. 2000). At the same time, buffering or mitigating factors should be considered, such as the degree of ambivalence, alliance with positive supports, and positive reasons for living (Linehan et al. 1983).

In the present study, the Suicide Monitoring Board (SMB) members were required to judge imminent suicide
risk secondhand, not from their own interviews with the patients but from clinical records and forms prepared by the patients' clinicians and research investigators.

Method

The three authors of this article constituted an SMB that rated the outcomes of interest for the study; they were selected because of their extensive clinical and research experience in assessing suicidal patients. Packages of anonymous material for each candidate event as it occurred were faxed for evaluation to each of the SMB members on an ongoing basis throughout the study. These consisted of the patient's clinical outpatient and hospital records, any coroner's reports, and an account of the event and the context in which it had transpired. A Suicide Attempt Form (SAF) was sent containing the details of all suspected attempts and/or, if indicated, a form for Imminent Risk of Suicide Requiring Hospitalization (IRH). The Calgary Depression Scale (Addington et al. 1998), InterSePT Scale for Suicidal Thinking (Lindenmayer et al. 2003), and Rescue Intervention Form and Prior History of Suicide Attempts (a form detailing the number, periodicity, and seriousness of any past attempts) were included.

The clinical trial medical monitors rigorously scrutinized the material prior to transmission to SMB members and removed all identifying information (patient names, identifying references to the clinical settings and to the trial drugs), ensuring blindness by the SMB. Clinical information not already in English was translated, and all handwritten information was retyped for legibility (both the typescript as well as copies of the originals were submitted). SMB members (in Helsinki, Durham, and Toronto) first rated each case independently and submitted a completed Suicide Event Form to the medical monitors. They then met by international teleconference at least once monthly to discuss cases that warranted further discussion because of conflicting interpretations. At these hour-long teleconferences, conflicts were thoroughly discussed and resolved by consensus. For several cases additional information was requested through the medical monitors in order to achieve consensus. The first 3 months were regarded as a training period, during which the investigators debated each other's understanding of critical definitions and concepts in estimating suicidality, such as "lethality" and "suicide intent."

The main outcome variable for the investigation was designed to be the time elapsing from entry into the study to the occurrence of an "endpoint," that is, a potentially lethal suicide attempt, or a rescue intervention (hospital admission or act of increased surveillance considered necessary to protect the patient from the genuine risk of a serious suicide attempt). Obviously, every death by suicide (i.e., a successful attempt) also counted as an endpoint. So-called "empty gestures," self-mutilation without intending serious self-injury, and expressed suicidal ideation without real intent to commit suicide, were not intended to be classified as endpoints. It was the task of the SMB—blinded to which trial drug a patient had been taking and at arm's length from the clinical investigators—to determine whether a given candidate event (including a death considered a possible suicide) qualified as an endpoint. Adjudication of each event was made by the SMB shortly after the fact, making such judgments necessarily retrospective. It is an inherent weakness of retrospective investigations that the data of interest are not necessarily recorded at the time of an event and thus are, afterward, unavailable for analysis. However, the InterSePT study was designed as a prospective drug trial with suicidality intended as the chief outcome from the beginning; hence, data collection instruments were incorporated that captured essential features required to help make the determinations on suicidality.

Although the decisions of the SMB provided the primary outcome data, the design of the multicenter study also called for a staff psychiatrist at each center who would serve as a blinded evaluator (the "blinded psychiatrist") of the clinical status of patients participating in the study and would review all suicidal events. The blinded psychiatrist interviewed study patients at 8-week intervals (after 48 weeks the intervals changed to between 4 and 12 weeks) and administered the InterSePT Scale for Suicidal Thinking—Blinded Psychiatrist (ISST–BP) (Lindenmayer et al. 2003). The ISST–BP was included among the clinical records faxed to the SMB whenever a putative event took place, but the intervals when it was recorded did not, of course, necessarily coincide with the date and time of the possible events. They were used as auxiliary material in the context of the clinical records and other information provided. The rationale for choosing ratings by the SMB as the main outcome measures, rather than those by the blinded psychiatrists, is detailed in the article discussing the design for the InterSePT study (Alphs et al. 2004).

Results

Of the 980 patients enrolled in the study internationally, 243 (24.8%) were considered to have experienced one or more potential endpoint events. In total 577 events were submitted to the SMB for review; of these, 483 (83.7%) were considered by consensus to be endpoints. Table 1 shows the initial levels of agreement among SMB raters before any discussions leading to consensus. These levels of agreement were statistically significant for each category of outcome event. Overall agreement among the
Table 1. Agreement before consensus discussion among Suicide Monitoring Board members

<table>
<thead>
<tr>
<th>Response</th>
<th>Kappa</th>
<th>Standard error</th>
<th>z</th>
<th>p &gt; z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-endpoint</td>
<td>0.38</td>
<td>0.024</td>
<td>15.69</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Valid rescue</td>
<td>0.51</td>
<td>0.024</td>
<td>21.05</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Valid attempt</td>
<td>0.68</td>
<td>0.024</td>
<td>27.76</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Overall</td>
<td>0.52</td>
<td>0.018</td>
<td>28.85</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Three members yielded a kappa coefficient (κ) of 0.52; this falls within a range of moderate agreement (Seigel et al. 1992; figure 1). Substantial agreement was reached in deciding whether a self-harm event constituted a potentially serious suicide attempt (κ = 0.68), but initial agreement over what represented a non-event (no endpoint) was no more than fair (κ = 0.38). Table 2 shows substantial levels of agreement achieved among the initial ratings of the SMB members, the overall SMB consensus, and ratings by the blinded psychiatrists at the clinical-investigative centers, with an overall kappa coefficient (0.64) higher than that in the initial SMB ratings. Agreement also improved in the other categories, for instance, that for classifying a non-event improved from κ = 0.38 to 0.45 (moderate).

Table 3 demonstrates that SMB ratings provide some validation for the ISST-BP as a measuring instrument, given that increasing mean and median scores for the ISST match progressively more serious event ratings.

The following five examples illustrate the variations among cases and events even in the same patient in their level of difficulty for outcome categorizations.

Case A. A married white female, 35 years old, at week 30, was proposed as a hospitalization for imminent risk of suicide on account of exacerbated psychosis with suicidal ideation. She had eloped from another facility and been brought in by the police. She refused to answer questions about whether she was thinking of suicide but had previously spoken of hopelessness and of severe paranoid delusions of harm to herself and her family from Satan and other evildoers. She admitted that she heard voices commanding her to run away from the other facility and that the television and radio were talking about her. She was aware that she had "lost her mind." Her past history included overdosing on tricyclics and abusing cocaine and amphetamines. She concomitantly suffered from Crohn's disease. The diagnosis of record was chronic paranoid schizophrenia. Since 2 years ago she had made four suicide attempts, most recently by an overdose, intending to kill herself. The ISST-BP, recorded 2 days after this...
been five lifetime hospitalizations to prevent suicide years before, regarded as a gesture. However, there had been two suicide attempts over his lifetime, one of them an overdose.

He was 18 at his first attempt and had made no steps to implement his plan to overdose and made no threats (suggesting a desire to conceal intent). Increased severity of auditory hallucinations led to increasing feelings of hopelessness and increased agitation. He was not using substances and was not intoxicated. Most of the time, he rejected suicide but could be impulsive and explosive. There were recent major life stressors, including adverse family dynamics and social isolation. A personal history of medically serious suicide attempts or family history of suicide was absent. He revealed suicidal ideation when specifically questioned by his psychotherapist, and he was not sure of his self-control. He was 18 at his first attempt and had made two over his lifetime, one of them an overdose 4 years before, regarded as a gesture. However, there had been five lifetime hospitalizations to prevent suicide attempts. The patient was admitted to a step-down facility to increase his level of support. The following day, under 24-hour supervision, he was feeling less suicidal and the day after that was discharged home.

The Calgary Depression Scale a month earlier showed mild depression and hopelessness and no suicidal feelings. However, he was moderately nervous, shaky, and jumpy. The clinical record stated that he had been experiencing auditory hallucinations for about a week, with passive suicidal ideation; he was feeling lonely, being alone at home during the day, and complained of lack of structure to his day. Two days following the intervention, he said that the auditory hallucinations were milder and he no longer had suicidal ideation. However, the auditory hallucinations he experienced were derogatory. His antidepressant (doxepin) was increased to 200 mg. There was no ISST-BP score.

SMB Conclusion. Initially, one SMB member voted against this being an endpoint, but on discussion all agreed that indeed it was a reasonable hospitalization according to the criteria. This was a male patient having some form of psychotherapy. His current age was not available, and there was little background history. The Calgary Depression Scale was unhelpful because it was a month old, but agitation was present even then. Although one of his two previous suicide attempts was regarded as a gesture (an overdose in 1995), clinicians saw fit to admit him on five occasions to prevent a suicide attempt. He was impulsive and explosive, and he had become more depressed and hopeless, and his derogatory auditory hallucinations had intensified. Although he lived with his mother and his sister, he complained of loneliness during the day making him depressed. His intention was to take an overdose, but he did not tell anyone; had he wanted to be stopped from carrying out such a plan, he probably would have told someone. He had opportunity because he was left alone. He told his psychotherapist on direct questioning, and this fortuitously led to his admission. Although this was only a short-term intervention of 24 hours, there was a very good short-term result, in that the following day his auditory hallucinations were lessened and he was no longer suicidal. The crisis admission was considered justified because it reduced the risk of self-harm in this suicidal, impulsive, agitated, depressed, hopeless, psychotic individual who was suffering from humiliating voices in his head.

Case C. This was a single woman of 26, with events occurring in weeks 84 and 86. She was said to be in intermittent conflict with her mother, at which times she would become aggressive and break a window of the rehabilitation unit when her mother refused to take her home. The window

| Table 3. Suicide Monitoring Board-rated categories by mean InterSePT scale scores⁠¹ |
|-----------------|------------------|------------------|
| Final (consensus) rating | n   | Mean (SD) | Median |
| No event         | 90   | 4.2 (5.1) | 2     |
| Valid rescue     | 64   | 5.8 (5.7) | 4     |
| Valid attempt    | 111  | 7.3 (6.0) | 7     |

Note.—InterSePT = International Suicide Prevention Trial. ⁠¹ Kruskal-Wallis test (c² = 13.69, df = 2, p = 0.001)
and one act of self-mutilation, called an attention-seeking gesture. He had had six hospitalizations to prevent a suicide attempt. He was an adopted child. He was said to be mildly depressed and moderately psychotic but not suicidal. The Calgary Depression Scale had been done 2 months earlier and showed him to be moderately depressed, hopeless, agitated, and mildly suicidal. The original clinical notes were in French and stated that he appeared to have no intention or plan. His revelation of the ingestion evidently surprised the clinician. He expressed delusional thoughts concerning the Royal Canadian Mounted Police and believed he had discovered a medical treatment for herpes and AIDS.

The ISST–BP was done 2 days after the event and reported no suicide wish or ambivalence and no expectation of an attempt. He was seen again 1 week after the ingestion and appeared stable but, although denying suicidal intentions, reported vague suicidal ideas.

**SMB Comment.** The SMB saw this as a suicide attempt. This was a male patient who ingested seven sleeping pills on an impulse at a time when he was discouraged and thought he had nothing much to live for. The impulse was brief, and the suicidality dissipated. Three of his four lifetime suicide attempts since the age of 19 were considered significant, and he had been hospitalized six times to prevent an attempt. It was noteworthy that his clinician regarded him as fragile less than a week before the attempt. He was floridly psychotic with paranoid delusions. Clearly, in spite of the low lethality of the overdose and his subsequent denials of intent to kill himself, this was, at the time that it occurred, a deliberate act intended to harm himself, motivated by depression and impulsivity and based on the conviction that his life was worthless.

**Case E.** This was a 47-year-old female chronically hospitalized on a psychiatric unit. She presented with nine events between weeks 64 and 98, sometimes occurring close together. There was almost no personal background history available, consistent with her chronic hospitalization and loss of personal identity and contact with her family. She was suffering auditory hallucinations, behavioral disorganization, and depressed mood, with irritability and insomnia. Over 18 years she had made at least a half-dozen suicide attempts, including taking four overdoses, jumping in front of a moving truck, and attempting a hanging. However, these acts were poorly documented.
Table 4 shows the individual categorizations of the three raters and the final consensus rating following discussion for each of the nine events. For the first event, at week 64, the story was that she was verbally aggressive over some days and impulsively mutilated herself after threatening suicide. She regularly experienced suicidal ideation, said she was ambivalent about death, and consistently spoke about her despair. She was described as a sad and hopeless woman who suffered from anxiety, command hallucinations, and insomnia. The last Calgary Depression Scale was done a month before this event and supported the view of her as being depressed, anxious, and hopeless. There was no recent ISST-BP. At consensus the SMB members felt that although she was chronically at high risk for suicide, on this particular occasion the self-harm was not intended to be fatal.

Several more events were not considered to be endpoints. However, in week 84 she again mutilated herself on the ward, with people around her and able to intervene, apparently doing so with mixed motivation, to escape from her loneliness and auditory hallucinations as well as to get attention. She was depicted as ambivalent about death, with poor self-control and more severe hopelessness than usual. As always, she was verbally aggressive with the nurses and other patients. On that day she scratched her neck with her razor but took no other action, and the following day she herself reported during the scheduled InterSePT study visit by the investigator that she had intended suicide because she could no longer endure her auditory hallucinations. Her act might even have gone unnoticed by the ward staff, she said. This time the unanimous conclusion was that this was a suicide attempt in spite of the low lethality of her self-harm.

In week 86 she again broke a glass and scratched her neck after making verbal threats of suicide. It was now apparent that she was experiencing considerable hopelessness and depression. The Calgary Depression Scale 2 months before showed severe hopelessness, self-depreciation, and observed depression. When asked why she had harmed herself she gave her reasons as being in the hospital for many years and not being able to live in society.

SMB Conclusion. The SMB unanimously voted to call this a suicide attempt because of the severity of her disturbed mood and her obvious despair. It is interesting, in retrospect, that although this patient may have made some serious past attempts, the significance of the current one had less to do with the low lethality of the attempt than with her mental state of severe desperation. Moreover, she was employing the restricted methods of self-harm available to her. Her previous history showed that when living in the community she resorted to the lethal methods then available (e.g., jumping in front of a moving truck).

There appears to have been a gap of about 12 weeks before the next event (although others might have occurred and gone unreported to the study personnel). In week 98 she again scratched herself with broken glass—as with the other similar attempts, an injury that required little or no medical attention. Four days later, at her week 98 appointment, she told the investigator that she had not been suicidal on this last occasion but had scratched herself because she wanted to get the trial drug stopped and because she was fed up with her situation. The SMB concluded that this particular event was not an endpoint, that is, a bona fide suicide attempt. This case was a reminder that in schizophrenia patients the logic of suicidal intentionality is neither linear nor consistent from event to event, and the degree of lethality of an event does not necessarily parallel its intentionality.

Discussion

In this study, the SMB, a team of three clinician-researchers, was required to adjudicate behavior suggestive of imminent suicidality by schizophrenia/schizoaffective patients who had been deemed by their clinicians to

<table>
<thead>
<tr>
<th>Week</th>
<th>Rater 1</th>
<th>Rater 2</th>
<th>Rater 3</th>
<th>Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>N</td>
<td>N</td>
<td>A</td>
<td>N</td>
</tr>
<tr>
<td>67</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>68</td>
<td>H</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>69</td>
<td>H</td>
<td>N</td>
<td>A</td>
<td>N</td>
</tr>
<tr>
<td>83</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>84</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>86</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>98</td>
<td>N</td>
<td>N</td>
<td>A</td>
<td>N</td>
</tr>
</tbody>
</table>

Note.—A = valid suicide attempt; H = valid rescue event (hospitalization or other increased protection); N = no endpoint event.
be at high risk for suicide. Review of the literature showed that while it is possible to delineate historical and current risk factors for suicide within samples of patients, when extrapolated to the individual patient they are suggestive but not definitive, either yielding too many false positives or failing to identify too many of those who later turn out to have been at risk for suicide. At the time of this study no published instruments existed for measuring suicide risk in the short, medium, or long term among schizophrenia patients. The SMB fell back on the gold standard of careful, clinical judgment. SMB members' teleconferences tested any conflicting opinions and achieved consensus, sometimes after accessing additional information they requested about the patient. The difficult task before the SMB members was to make clinical judgments in patients they had not directly interviewed and with documentation that varied in quality. However, Appleby et al. (1999a, 1999b) made similar determinations of predictable risk of suicide from case notes in their case control study of the aftercare of patients who subsequently committed suicide. In the present study, the key to evaluating the significance of the behavioral phenomena turned on assessing the seriousness of the suicidal intent and the driving force behind it. This is always difficult, even when one is able to carry out an empathic interview with a nonpsychotic patient; it is even more difficult when the evaluation must be made in a psychotic patient from information put on a chart by other clinicians.

It is useful to distinguish between intent that is subjective (explicitly stated) and that which is objective (implicit in the circumstances) (Rudd and Joiner 1998). Evidence of preparation, choice of a highly lethal method, and steps taken to prevent the act from being thwarted by others are evaluated as objective intent. The clinician has to estimate the degree of trust that can be placed in a patient's statement of intent in both directions—that is, that the person will or will not kill him- or herself. Patients are well known to threaten self-harm for the sake of some gain such as admission to a hospital. On the other hand, a seriously suicidal person is likely to deny or conceal intent so that suicide will not be prevented. Furthermore, psychotic patients do not always follow conventionally logical pathways; for instance, the lethality of an attempt may not equate to the degree of intent. This may account for the frequent finding that suicide victims were not perceived as at risk for suicide at their last clinical appointment before suicide (Appleby et al. 1999a, 1999b).

For the above reasons, the accurate prediction of suicidal risk in an individual with schizophrenia is extremely difficult. Even though suicide remains a relatively common problem in this disorder, its low base rate adds to the difficulty (Pokorny 1993). Efforts have been made to improve the quality of assessment. A largely Finnish group developed a 25-item Schizophrenia Suicide Risk Scale (SSRS) based on the literature on suicide in schizophrenia. Applying the instrument to data from psychological autopsies of 69 schizophrenia patients who committed suicide and from 69 living schizophrenia patients as controls, they found that significant predictors of suicide were suicide plans communicated to someone in the previous 3 months; one or more previous suicide attempts; loss of professional skills; observed depression; and suicide plans divulged during an interview (Taiminen et al. 2001). In spite of these findings, the authors found the scale to be either too insensitive or too nonspecific for general use as a screening device in schizophrenia patients.

Turner et al. (1998) developed another instrument. A semistructured Interview for Suicide in Schizophrenia (ISIS) was part of a comprehensive assessment that included chart review, staff reports, and collateral information from families. Psychometric properties of the ISIS-3 (a 140-item third revision of the ISIS) were tested on 270 schizophrenia patients at a state psychiatric hospital (Korslund 2001). Suicide Risk Index (SRI) scores were found to correlate significantly with independent clinical ratings ($r = 0.49, r^2 = 0.24$), hospital records of suicidal behavior ($r = 0.47, r^2 = 0.22$), and chart documentation of suicidal precautions ($r = 0.58, r^2 = 0.34$). Sensitivity and specificity of the ISIS-3 diagnosis of suicidality risk were satisfactory, as indexed by the clinical rater diagnosis (sensitivity 70%; specificity 85%) and chart reports of suicidal behavior events (sensitivity 64%; specificity 87%). The ISIS-3 was evaluated for validity not against fatal suicide events (as the SSRS was) but rather against nonfatal self-harm behavior.

The instructions for administering the ISIS prescribe that, "when considering a rating for a patient, the interviewer should consider all available information, including direct questioning of the patient, observation, chart review, and reports from family and treatment providers... as well as interviews with hospital staff and family members who have knowledge of the patient's psychiatric and suicide history... Level of unit precaution and documentation of recent suicide behavior (is) recorded at this time" (Korslund 2001, pp. 20 and 28). These directions imply that SRI scores may be contaminated by other sources of information and that they may not be independent.

In principle, however, a semistructured instrument such as the ISIS-3 ensures that all known vulnerability factors for suicide in schizophrenia are systematically explored at interview. Both the SSRS and the ISIS-3 are worthy steps in the effort to refine the identification of schizophrenia patients at risk for suicide. The ISIS-3 requires further evaluation that ensures blindness of the
Estimating Suicidality


Interviewer to the outcome variable of interest, and both instruments could benefit from external and prospective evaluation.

Presently, therefore, clinical judgment, guided by the facts of the case history and interviews, remains the gold standard for estimating degree of suicidality. Nonetheless, if a trial such as the InterSePT were to be designed today, the SSRS and ISIS—3 might well be incorporated as auxiliary instruments and evaluated. Using the gold standard of clinical judgment alone, the SMB, when making initial judgments, achieved acceptable levels of interrater agreement prior to consensus discussion (overall $\kappa = 0.52$; table 1). This degree of concordance is regarded by Landis and Koch (1977) as moderate and by Fleiss (1981) as fair to good (figure 1). However, the design of the trial called for the consensus ratings reached after discussion, and not the initial ratings, to constitute the outcome measures for this trial. When the consensus ratings and those of the blinded psychiatrist were included with the initial ratings (table 2), the overall kappa coefficient reached 0.64, which is considered to be substantial by Landis and Koch (1977). The kappa coefficients attained in the study need also to be considered in light of the fact that kappas tend to be lower when the population sample is homogeneous, as was the present sample of patients, constituted by a narrow range of diagnoses and selected for being suicidal.

It is interesting to consider the types of situations the SMB found that led to difficulties in categorizing whether an event should be rated as an endpoint for purposes of the trial. Tables 1 and 2 show that agreement was highest in determining whether an event was a valid suicide attempt, followed by whether a “rescue intervention” was justifiable in terms of the risk for significant self-harm. Agreement was lowest in determining that a given event was not an endpoint, perhaps revealing a conservative stance of the SMB members. These differences reflect the difficulty that SMB members had in estimating suicidal intent when interpreting clinical records made by third-party interviewers. Psychotic patients were less accessible than other types of patients; their suicidal intentions were often covert and behaviorally explosive. SMB members could also not rely on the lethality of the putative suicidal behavior alone, because schizophrenia patients with cognitive disturbances do not reason like people not similarly affected. The example of Case E (the woman who scratched her neck superficially because she could no longer endure her auditory hallucinations) illustrates that there can be serious suicidal intent behind what appears to be a trivial act of self-harm, and this is probably true as well in the reverse (Beck et al. 1975). The SMB could not extend judgments made on earlier events to later ones; each event had to be independently categorized on its own contextual evidence. Suicidality, as in Cases C and E, is a “moving target” in vulnerable individuals driven by transitory exacerbations of suicidal ideation and intentionality (Kral and Sakinofsky 1994)—that is, it can be seen as a dynamically emergent process subject to the rules of both quantum and chaos theory (Korslund 2001). A contextual framework that ignites suicidal behavior at one time may not necessarily do so at another time. With schizophrenia patients who carried diagnoses of comorbid substance abuse disorder (alcohol or illicit drugs), the evaluation was complicated by the pharmacological effects of intoxication as well as by their impulsivity.

Overall, SMB members relied on evidence of subjective distress and/or concomitant depression combined with agitation, hopelessness, and impulsivity that might trigger a self-destructive event and, in the case of putative attempts, on objective circumstances indicating precautions taken by the patient to ensure the success of the attempt (Beck et al. 1974; Rudd and Joiner 1998). If command hallucinations were present, they were regarded as carrying significant weight, as was a history of previous suicidal behavior of potential lethality. In each case, the SMB members scrutinized the data for critical factors that might sway the balance of the decision (e.g., a patient taking an overdose when living alone—even one that would not be lethal—and telling nobody), but often the classification decision had to be made on an overall impression of the case history in its circumstantial context. Once the global perspective or empathic understanding of a case fell into place, SMB members might with confidence reassign different weights to the provisionally weighted, individual features of the case. Indeed, the process whereby clinicians estimate suicidality in their patients and assess the complex factors that go into decision making has not been well studied previously (McNiel and Binder 1997) and deserves further investigation.

The ISST–BP was completed regularly by the blinded psychiatrists at set timepoints for each patient in the trial and was included in the material sent to the SMB for consideration. ISST–BP scales constituted data that were weighed in the overall context of the clinical information available for each case. Based on the overall clinical pictures, SMB members sometimes classified events differently from what might be indicated by an ISST–BP that might have been completed at a different time. Table 3 demonstrates a significantly increased progression of ISST scores for each of the categories rated by the SMB (no event, valid rescue, and valid attempt). This finding supports both the validity of the ISST–BP and the consensus ratings by the SMB.

In summary, although the three SMB psychiatrists classified suicidal events from clinical records of schizophrenia or schizoaffective patients they had never directly examined, the findings show acceptable agreement...
among their classifications, both before and after reaching consensus, and those of the on-site psychiatrists who saw the patient but were blinded to the treatment. These outcome data have become the pivotal endpoints applied in the InterSePT study, showing that clozapine had a greater benefit than olanzapine in reducing suicidal behavior in high-risk patients (Meltzer et al. 2003). Finally, although information obtained by direct interview of schizophrenia patients is preferable, the findings show that where this cannot be done, experienced clinicians are able to come to useful conclusions from chart reviews. Given that the InterSePT study encompassed 67 sites in 11 countries, with patients coming from diverse subcultures, socioeconomic groups, and medical systems and speaking eight different languages (Alphs et al. 2004), the SMB provided an outcome measure for the investigation that brought to bear the benefits of guaranteed blindness, joint training, uniformity, and standardization of classificatory concepts. This would not have been practical to achieve with the individual psychiatrist raters based at each of the 67 sites.

References


**Acknowledgments**

The authors thank Drs. Larry Alphs, Ravi Anand, Kevin Cox, Franco DiCesare, and Jay Hsu for assistance. This study was made possible by financial support from Novartis Inc. Dr. Krishnan served as chair of the Suicide Monitoring Board during the InterSePT study.

**The Authors**

Isaac Sakinofsky, M.D., F.R.C.P.C., F.R.C.Psych., is Professor Emeritus, Department of Psychiatry and Public Health Sciences, University of Toronto, Toronto, Canada. Hannele Heila, M.D., Ph.D., is Senior Researcher, National Public Health Institute, Department of Mental Health and Alcohol Research, Helsinki, Finland. Ranga Krishnan, M.D., Ph.D., is Professor and Chair, Department of Psychiatry, Duke University Medical Center, Durham, NC.