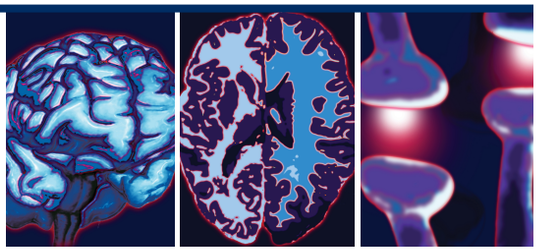

Suicide risk assessment tools, predictive validity findings and utility today: time for a revamp?

Leslie Roos[1,2], Jitender Sareen[2,3] & James M Bolton*[2,4]

Practice points

- When using a tool for 'risk assessment,' it is critical for the tool to be prospectively validated to predict the outcome of interest.
- The most commonly employed self-report suicide risk assessment tools to date have not been found to significantly predict suicide outcomes, in any adult sample, at the 1-year follow-up.
- Cognitive risk assessment tools offer a short-term risk assessment tool, with strong predictive validity that have outperformed other clinical indicators, including axis 1 diagnoses, patient prediction and clinician prediction.
- Researchers interested in suicide risk prediction should explore integrating cognitive risk assessment tools into ongoing or future studies to investigate how these tools compare with standard risk assessment tools.
- Hospitals and clinical settings pursuing suicide risk assessment tools should be cognizant of the findings that suggest a lack of predictive validity of standard assessment measures.
- Future research should seek to integrate a variety of risk assessment tools to determine if there might be a specific tool of clinical utility or if a composite scores assessment of different dimensions of risk can more accurately capture more at-risk individuals.

SUMMARY Suicide is an internationally recognized burden of health, but little progress has been made towards creating effective risk assessment tools. In order to be used in clinical settings, these tools must prospectively differentiate between future attempters and nonattempters, do so with adequate sensitivity and specificity, and do so in a clinically useful time frame. Given these criteria, we review the state of classic suicide risk assessment tools, which rely on self-report of suicidal symptoms or clinical risk factors (i.e., hopelessness). In summary, there are substantial limitations to such paper-based tools given the incentives to deny suicidal thoughts, a lack of replication and the lengthy follow-up time frames identified by most studies. Next, we review the evidence for a new type of computer-based

[1]Department of Psychology, University of Oregon, Eugene, OR, USA
[2]Department of Psychiatry, University of Manitoba, Winnipeg, MB, Canada
[3]Department of Psychology, University of Manitoba, Winnipeg, MB, Canada
[4]Department of Community Health Sciences, University of Manitoba, Winnipeg, MB, Canada
*Author for correspondence: jbolton@exchange.hsc.mb.ca

Future Medicine part of fsg

risk assessment tool that utilizes implicit cognitive associations with suicide as an indicator of implicit biases. By comparing the classic self-report versus cognitive risk assessment techniques, substantial advantages have emerged regarding predictive validity when using this cognitive test approach. Although these tools may take additional efforts to integrate into clinical assessments, they offer substantial advantages through their ability to predict short-term (~6 month) suicidal behavior. It is not suggested that such tools replace the utility and importance of clinical interviews or expertise, but that they could rather provide a valid tool to inform clinician decisions for acute care.

Substantial progress has been made in the past decade towards identification of risk factors for suicide and suicidal behaviors. Psychological autopsy studies have now established that psychiatric disorders account for 47–74% of suicides at a population level [1]. Observational studies have also consistently corroborated this finding [2,3]. Other suicidal behaviors, such as past attempt, or psychiatric hospitalization denote particularly high risk and occur in an estimated 25–65% of suicide attempts internationally [4,5]. While at the outset this news may seem to be excellent progress towards minimizing the burden of this public health issue, suicide remains a leading cause of death in both the USA and world [6].

The critical variable here is the art of predicting future suicides as opposed to describing correlational risk factors of the past. Without clear evidence-based tools to guide clinical provider decisions, this public health issue remains seriously unaddressed. A theoretically promising approach towards effective suicide risk management is the development of a standardized tool to gauge risk in a clinical setting. Much like family practitioners or surgeons need tests to identify patients as high or low risk for a certain outcome, psychiatrists and others in the mental health professions could use these tools with confidence in their predictive validity. Without such a tool, the science of suicide risk assessment is more of an art, based on the expertise of individual clinicians, which may vary widely [7].

Multiple factors, such as discrete risk profiles, specificity versus sensitivity and immediacy of risk, make the development of an effective tool inherently difficult, but no less critical. Over 20 of these tools have been developed by experts in the field, but effective implementation of these tools to reduce suicidal behavior has proven extremely difficult. Furthermore, many of these tools may be particularly poor at predicting completed suicide given the incentives acutely suicidal individuals have to deny psychiatric distress. Many studies report promising associations with past suicidal behavior, physician-reported self-efficacy, hospital admission or reliability, but very few have shown substantiative predictive validity [8]. Revisions to the DSM have further identified the importance of utilizing a clinically valid tool as top priority [9]. The challenges of predictive validity have been a forefront concern for the past 30 years, but predicting this human behavior has ben proven to be substantially challenging [10,11].

Here, we consider the current state of risk assessment tools by specifically detailing each scale that has been examined for predicting future suicidal outcomes in adults **(Table 1)**. To our knowledge, the dozens of other suicidal scales have not been prospectively examined in regards to suicidal outcomes. Some scales, such as the Columbia Suicide Severity Rating Scale, whose use is compulsory in US FDA clinical trials, have been assessed for predictive validity in adolescents; however, no published studies have examined such predictive validity in adults. Accordingly, this scale is not included in this review [12]. Of additional note is that we are not reviewing the predicative validity of these scales for specific subpopulations at elevated risk (adolescents or the elderly) as these age groups have their own well-developed literature and risk factor considerations [13,14]. Although prior reviews have examined the evidence for individual scales or in specific clinical settings, to our knowledge, none have considered the broad range of scales with research on predictive efficacy [15,16]. Additionally, multiple new scales have emerged in recent years.

We describe multiple challenges to the useful clinical implementation posed by classical risk assessment tools. After reporting on the evidence base for these standardized tools, we move on to consider an emergent field: cognitive risk assessment tools. This field presents substantial challenges for implementation, but shows great promise for predicting outcomes. It should be seriously considered if predicting suicidal outcomes is to be a health priority. Explicit acknowledgement of the mental health professional's inability to effectively predict the future with standardized

Table 1. Studies prospectively examining suicidal outcomes.

Study (year)	Population sample (n)	Predictive of suicidal behavior?		Relevant statistics	Time of follow-up (years)	Ref.
		Suicide attempt	Suicide completion			
BHS						
Beck and Steer (1985)	Psychiatric inpatients with suicidal ideation (207)	–	Yes	Sensitivity: 0.91, specificity: 0.51, cutoff: >9	10	[19]
Beck *et al.* (1990)	Psychiatric outpatients (1958)	–	Yes	Sensitivity: 0.94, specificity: 0.41, cutoff: >8	7.5	[20]
Brown *et al.* (2000)	Psychiatric outpatients (6891)	–	Yes	PPV: 0.01, NPV: 1.00, hazard ratio: 4.46, cutoff: >8	20	[21]
Beck *et al.* (1989)	Hospitalized suicide attempters (413)	–	No	–	5–10	[22]
Oquendo *et al.* (2004)	Psychiatric outpatients with mood disorders (308)	No	No	–	2	[23]
Stefansson *et al.* (2012)	Psychiatric outpatients with suicide attempts (81)	–	No	–	10–15	[27]
BDI						
Beck and Steer (1985)	Psychiatric inpatients with suicidal ideation (207)	–	No	–	10	[19]
Beck *et al.* (1990)	Psychiatric clinical sample, across disorders (1958)	–	Yes	Sensitivity: 0.77, specificity: 0.64, cutoff: >22	7.5	[20]
Brown *et al.* (2000)	Psychiatric outpatients (6891)	–	Yes	PPV: 0.02, NPV: 1.00, hazard ratio: 3.55, cutoff: >22	20	[21]
Beck *et al.* (1989)	Hospitalized suicide attempters (413)	–	No	–	5–10	[22]
Tejedor *et al.* (1999)	Psychiatric inpatients with suicide attempt (150)	No	No	–	10	[28]
Oquendo *et al.* (2004)	Psychiatric outpatients with mood disorders (308)	Yes	Yes	Hazard ratio: 2.35	2	[23]
Hartl *et al.* (2005)	Veterans with post-traumatic stress disorder Study used two samples: exploratory (409) and confirmatory (221) Statistics include model with both BDI and recent attempt as predictors†	Exploratory: Yes Confirmatory: No	–	Exploratory: sensitivity: 0.63, specificity: 0.80, Confirmatory: sensitivity: 0.00, specificity: 1.00, cutoff: >45	4	[24]
Beck's SSI						
Beck and Steer (1985)	Psychiatric inpatients with suicidal ideation (207)	–	No	–	10	[19]
Brown *et al.* (2000)	Psychiatric outpatients (6891)	–	Yes	PPV: 0.03, NPV: 1.00, hazard ratio: 6.56, cutoff: >2	20	[21]

†Samples in the aforementioned studies included patients from a wide range of demographics (age, gender and race/ethnicity) and psychiatric diagnoses, which may influence both the predictive validity and reliability of study findings.
‡See Freedenthal for a review of all predictive studies [**25**].
BDI: Beck's Depression Inventory; BHS: Beck's Hopelessness Scale; IAT: Implicit Association Task; KIVS: Karolinska Interpersonal Violence Scale; NPV: Negative predictive value; PPV: Positive predictive value; R^2: R-squared; SNAP-SH: Schedule of Nonadaptive and Adaptive Personality Self-Harm; SIS: Suicide Intent Scale; SSI: Scale for Suicidal Ideation.

Table 1. Studies prospectively examining suicidal outcomes (cont.).

Study (year)	Population sample (n)	Predictive of suicidal behavior?		Relevant statistics	Time of follow-up (years)	Ref.
		Suicide attempt	Suicide completion			
Beck's SSI (cont.)						
Oquendo *et al.* (2004)	Psychiatric outpatients with mood disorders (308)	No	No	–	2	[23]
Nock *et al.* (2010)	Emergency room psychiatric patients across disorders (157)	No	–	–	6	[29]
SIS[†‡]						
Beck *et al.* (1989)	Hospitalized suicide attempters (413)	–	No Subscale did[†]	–	5–10	[22]
Stefansson *et al.* (2012)	Psychiatric outpatients with suicide attempts (81)	–	Yes	Sensitivity: 1.00, specificity: 0.52, PPV: 0.17, cutoff: 16	10–15	[27]
Nimeus *et al.* (2000)	Previous suicide attempters (555)	–	Yes	PPV: 0.10, cutoff: 19	1	[39]
SAD PERSONS and modified SAD PERSONS						
Bolton *et al.* (2012)	Emergency room patients, psychiatric referral (4019)	No Subscale did[†]	–	–	2	[35]
MINI suicidal subscale						
Roaldset *et al.* (2013)	Psychiatric inpatients, following discharge (307)	Yes, including nonsuicidal self-injury, suicide threat and suicide attempts[†]	–	Sensitivity: 0.73, specificity: 0.75, PPV: 0.39, NPV: 0.88, cutoff: >5 Sensitivity: 0.61, specificity: 0.43, PPV: 0.43, NPV: 0.86, cutoff: >9	1	[37]
SNAP-SH subscale						
Yen *et al.* (2011)	Individuals with a personality disorder (733)	Yes	–	Sensitivity: 0.72, specificity: 0.85, PPV: 0.33, cutoff: 12	8	[38]
Suicide Assessment Scale						
Waern (2010)	Emergency room psychiatric patients following suicide attempt (165)	Yes (for combined attempters and completers)		Sensitivity: 0.61, specificity: 0.40, cutoff: 24	3	[40]
Nimeus *et al.* (2000)	Previous suicide attempters (191)	–	Yes	Sensitivity: 0.75, specificity: 0.86, cutoff: >38	1	[39]

[†]Samples in the aforementioned studies included patients from a wide range of demographics (age, gender and race/ethnicity) and psychiatric diagnoses, which may influence both the predictive validity and reliability of study findings.
[‡]See Freedenthal for a review of all predictive studies [25].
BDI: Beck's Depression Inventory; BHS: Beck's Hopelessness Scale; IAT: Implicit Association Task; KIVS: Karolinska Interpersonal Violence Scale; NPV: Negative predictive value; PPV: Positive predictive value; R^2: R-squared; SNAP-SH: Schedule of Nonadaptive and Adaptive Personality Self-Harm; SIS: Suicide Intent Scale; SSI: Scale for Suicidal Ideation.

Table 1. Studies prospectively examining suicidal outcomes (cont.).

Study (year)	Population sample (n)	Predictive of suicidal behavior?		Relevant statistics	Time of follow-up (years)	Ref.
		Suicide attempt	Suicide completion			
KIVS						
Jokinen *et al.* (2010)	Previous Suicide Attempters (161)	–	Yes (childhood exposure to violent behavior)	Sensitivity: 0.65, specificity: 0.80, cutoff: 3, PPV: 0.14	4	[42]
			Yes (adult expressed violence)	Sensitivity: 0.88, specificity: 0.60, cutoff: 3, PPV: 0.14		
Death/suicide IAT						
Nock *et al.* (2010)	Emergency room psychiatric patients across disorders (157)	Yes	–	Sensitivity: 0.50, specificity: 0.81, PPV: 0.32, NPV: 0.90, cutoff >0, suicide/death bias	0.5	[29]
Suicide Stroop task						
Cha *et al.* (2010)	Emergency room psychiatric patients across disorders (60)	Yes	–	R^2 change: 0.18	0.5	[56]

†Samples in the aforementioned studies included patients from a wide range of demographics (age, gender and race/ethnicity) and psychiatric diagnoses, which may influence both the predictive validity and reliability of study findings.
‡See Freedenthal for a review of all predictive studies [25].
BDI: Beck's Depression Inventory; BHS: Beck's Hopelessness Scale; IAT: Implicit Association Task; KIVS: Karolinska Interpersonal Violence Scale; NPV: Negative predictive value; PPV: Positive predictive value; R^2: R-squared; SNAP-SH: Schedule of Nonadaptive and Adaptive Personality Self-Harm; SIS: Suicide Intent Scale; SSI: Scale for Suicidal Ideation.

effectiveness is essential for opening a frank dialog on what can be done to better meet the grave needs of highly at-risk individuals.

Scales investigated for predictive validity

Much of the early work on the prediction of suicidal outcomes was performed by Beck and colleagues in the 1970s and 1980s. Following a call from a National Institute of Mental Health task force, this research sought to address the utility and predictive validity of constructs, such as 'intent' and 'lethality', in standardized form [17,18]. Briefly, Beck's Depression Inventory (BDI) aims to investigate 21 self-reported depressive symptoms based on severity (0–3), Beck's Hopelessness Scale (BHS) investigates pessimism through 20 true or false questions, and the Scale for Suicidal Ideation (SSI) investigates the severity of suicidal ideas and wishes with 19 questions based on severity (0–3). The Suicide Intent Scale (SIS) aims to measure intent to die among individuals with a history of past suicide attempts. These self-report measures are reviewed together as they are generally investigated in the same studies. These tools are the most widely studied and have the strongest evidence base.

Studies from Beck's group have investigated inpatients (n = 206) and outpatients (n = 1958) with admission for suicidal ideation (and no recent attempt) using BHS, BDI and SSI [19,20]. Over a 5–10-year follow-up period of inpatients and outpatients, 6.9 and 0.88% of patients, respectively, died by suicide. In both inpatients and outpatients, the BHS significantly differentiated between completers and noncompleters, while the BDI was also significantly related to outpatient completers. In this research, the authors suggest different cutoff scores for inpatients (10) and outpatients (9) with roughly 0.45 sensitivity and 0.90 specificity. Of particular note, the inpatient study investigated SSI, but did not find predictive validity, although a separate 20-year follow-up (n = 6891) by Beck's group did find differences on the SSI between completers and noncompleters [19,21]. Initial research on the SIS by Beck's group for predictive validity found that, although the whole scale did

not predict future suicide among prior attempters with alcohol abuse, the precautions factor subscale did [22].

While the aforementioned studies are incredibly impressive for their long-term follow-up and investigation of suicide completion as an outcome, research on the clinical utility of the scales has appreciable limitations. First, studies investigating the predictive validity of the BHS or the BDI for future attempts or completion have shown highly mixed results. A meta-analysis of the BHS found much lower predictive validity for suicide across studies than that reported in the validation studies (sensitivity 0.29–0.54/specificity 0.60–0.84) [16]. One study of depressed individuals found that severity of depression (BDI) predicted future attempt, but not severity of hopelessness (BHS) at 1-year follow-up [23]. Research in a military sample with posttraumatic stress disorder additionally highlighted the challenge of replication with 0.63 sensitivity and 0.84 specificity for future suicide attempt in an exploratory model (n = 409) containing a BDI cutoff of 46 and recent suicide attempt. However, the replication model (n = 221) had no predictive power with 0.00 sensitivity and 1.00 specificity. [24]. These findings point out challenges of scales in regards to cross-population validity and variable cutoff deviations validity, in addition to replication difficulties.

The inconsistency in predictive utility of the SIS among individuals with prior attempts was further identified by a review that found evidence for positive associations in five studies (follow-up of 10 months to 20 years), but no evidence in an additional seven studies (follow-up of 113 days to 10 years) [25]. Of the studies that did find SIS predictive power, multiple studies used different ideal cutoff scores [26,27]. Additionally, prospective studies examining BHS, BDI and SSI with more immediate follow-up periods (6 months to 2 years) often report marginal to null associations of the aforementioned scales with suicidal behavior [5,27–29]. While both long-term and shorter follow-up studies are of interest in suicide research, suicide risk in the short term (within the next 6–12 months) has perhaps greater clinical importance given the decisions clinicians face when determining disposition in treatment settings. In summary, the multiple null findings of Beck's scales in acute timelines combined with the relatively low sensitivity and inconsistent cutoff scores limit the practical utility of these scales in informing clinical care.

■ Suicide probability scale

The suicide probability scale (SPS) was developed by Cull and Gill and includes 36 self-report items relating to hopelessness, suicide ideation, negative self-evaluation and hostility [30]. These items are derived from factors observed in a large clinical sample, which distinguished between adults and adolescents who had and had not completed suicide. While the authors did not investigate the predictive validity, it has been suggested for use in this context and has widespread use across clinical settings [31]. Only one study has investigated the predictive validity of this scale in adults, specifically an adult inmate population (n = 1047) over a mean follow-up of 46 months (maximum 10 years) while in prison [32]. This study found that, compared with individuals with any suicidal behavior (nonsuicidal self-injury, suicidal ideation, suicide attempt, and completed suicide) scores on the SPS significantly discriminated between groups at both the short-term (2 years) and long-term (up to 10 years) follow-up time points. Regarding specific suicidal behaviors, the SPS distinguished between suicide outcomes of multiple attempts and suicidal ideation, but not of completed suicide. Another limitation of note is the relatively poor sensitivity (0.36) and specificity (0.85) of SPS. Interestingly, this scale has shown positive predictive power among at-risk adolescents in previous studies, but reviewing the differential risk patterns of adolescent suicide is outside the scope of this article [33,34]. Overall, this scale has promising use for clinical settings, however, there is only one study investigating its predictive validity and no evidence base for usage in nonincarcerated settings.

■ SAD PERSONS & modified SAD PERSONS scale

Developed in 1983, the SAD PERSONS scale attempts to assess suicide risk by ten factors identified through a literature review. While widely internationally implemented across clinical professions, research on predictive validity was not reported until recently. In this study, our group assessed the ability of SAD PERSONS to predict future suicide attempts among all psychiatric service referrals (n = 4019) [35]. This study found that medium or high risk of SAD PERSONS and modified SAD PERSONS were significantly associated with suicidal presentation compared with low risk. However, half of suicidal presentations had low risk scores, and both scales had large false-positive rates, indicating both low

sensitivity and specificity for acute psychiatric settings. While specific items from the scale were highly correlated with future attempts, findings suggested that the SAD PERSONS was no better at predicting suicide attempts than chance.

■ The Mini-International Neuropsychiatric Interview suicidal subscale

The Mini-International Neuropsychiatric Interview (MINI) is designed as a short neuropsychiatric interview tool to assess a range of axis I disorders. While it has not been considered in predicting suicide risk until recently, it is well validated against longer clinical interviews, such as the Structured Clinical Interview for the DSM, and has a minimum of 0.70 specificity and 0.85 sensitivity across disorders [36]. The first predictive suicidal research on this scale was reported in 2013, in a sample of patients (n = 307 at follow-up) tracked for 1 year after discharge from an acute psychiatric ward [37]. Outcome measures included both suicidal behavior (imminent threat or attempt) and nonsuicidal self-injury (NSSI). Patients were categorized in the following groups: no symptoms, suicidal behavior, suicidal behavior and NSSI, and NSSI only. Controlled for age, gender and psychiatric diagnoses, MINI suicidal subscale scores were significantly related to each of the self-harm groups. Using cutoffs of six points (or more than two positive items) and ten points (or more than three positive items) for moderate- and high-risk groups, respectively, produced good sensitivity (0.61–0.75) and specificity (0.61–0.75) for patients with a history of self-harm acts. Considering the three groups separately, it was determined that the MINI suicidal subscale score was a good predictor of both suicidal behavior groups, but not the NSSI group alone. This study offers an exciting potential risk assessment tool, with particular use given its ease and relatively short time period of study. However, replication studies and studies examining its predictive ability for future suicide are needed prior to its implementation as a clinical tool for suicide prediction.

■ Schedule of Nonadaptive & Adaptive Personality self-harm subscale

The self-harm subscale of the Schedule of Nonadaptive and Adaptive Personality assesses nine items investigating suicide proneness and seven items related to self-esteem. The larger Schedule of Nonadaptive and Adaptive Personality scale investigates the presence of both normal and abnormal personality characteristics. This scale has also been newly introduced for prospective suicidal literature and studied over 8 years of prospective outcomes in participants (n = 701) with axis II disorders in the Collaborative Longitudinal Study of Personality Disorders [38]. Using a cutoff score of 10, the authors report a sensitivity of 0.84 and specificity of 0.70, or a sensitivity of 0.72 and specificity of 0.85 with a cutoff score of 12. In a population with axis II disorders, these findings suggest that the self-harm subscale of the Schedule of Nonadaptive and Adaptive Personality scale may provide clinical utility in assessing the risk of suicide attempts among individuals with axis II disorders. This narrow population, however, would be of limited use in a clinical setting where a full assessment of axis I and II disorders may or may not be needed. Substantially more research would be needed to establish a relationship between score and likelihood of a suicide attempt across populations.

■ Suicide Assessment Scale

The Suicide Assessment Scale (SUAS) attempts to create a composite of suicide risk by assessing five areas relating to affect, control and coping, bodily states, emotional reactivity, and suicidal thoughts and behaviors. Each of these 20 questions is rated on a five-point severity scale from 0–4. Evidence for its utility in identifying individuals at risk for suicide was first reported in research investigating predictors of completed suicide in a sample of previous suicide attempters (n = 191) [39]. This study found that a cutoff score of 39 significantly differentiated between completers and noncompleters with a sensitivity/specificity of 0.72/0.86. A similar study examined the SUAS in a sample of emergency room psychiatric patients presenting with suicidal behavior, using a 3-year follow-up from hospital records. The SUAS version used was slightly updated from the one used in the aforementioned study and measurements were taken at both baseline and 2-month follow-up. This study found that the SUAS significantly predicted repeated suicidal behavior, using an optimal cutoff score of 24. However, this score yielded relatively poor sensitivity/specificity of 0.61/0.40 [40]. Additionally, the SUAS significantly differentiated between attempters and nonattempters when controlling for age and depression, but did not when controlling for anxiety. Of note when considering the results of theses studies, is that different clinically optimal cutoff scores were identified in the

two studies, which make a useful cutoff guideline challenging to consider.

■ Karolinska Interpersonal Violence Scale

Suicide research using the Karolinska Interpersonal Violence Scale (KIVS) assesses exposure to violence and expressed violence as predictors of completed suicide in previous suicide attempters.

This work is based off of research suggesting atypical impulsivity and violence found in individuals who attempt suicide [41]. The KIVS has been previously validated against biomarker correlates of suicidality and violent acts, such as the serotonin metabolite 5-hydroxyindoleacetic acid, but, as the authors note, these biomarkers are impractical for clinical use, given their influence by previous psychotropic medication [42,43]. Suicide attempters (n = 161) and healthy controls (n = 95) were assessed along subscales for exposure to violence and expressed violence in both childhood and adulthood [42]. Differences emerged at baseline on the KIVS subscales between suicide attempters and healthy controls. Importantly, significant differences also occurred when prospectively assessing completed suicide (n = 5) within 4 years, as assessed through the National Death Registry of Sweden. Specifically, exposure to violence in childhood and expressed violence in adulthood significantly predicted completed suicide among suicide attempters. In multipredictor regression models, including comorbid substance use, personality disorder, and gender, only the model with expressed violence as an adult was significant. While the KIVS shows promise as a predictor of completed suicide in a high-risk population, there are substantial limitations of its use as an acute clinical tool, including a relatively low positive predictive value (0.07 and 0.14 for exposure to violence as a child and expressed violence as an adult) and long-term follow-up time. Furthermore, these findings were in a sample with a small incidence of completed suicide, so it would be necessary to replicate these findings in a larger sample.

Summary of current standardized tools

Overall, there is a paucity of evidence to support the use of traditional risk assessment tools to predict future suicidal behavior or prevent future attempts. Across these scales, it is evident that predicting future suicidal behavior based on standardized scales produces, at best, inconsistent results. Even if the inconsistency of results in the aforementioned studies were acceptable, additional important caveats would have to be reconciled. First, many of the aforementioned scales demonstrated predictive validity only among groups with prior suicide attempts, so widescale implementation in acute care settings for individuals with suicidal ideation or plans would be unsupported. Additionally, many scales have been validated with suicide attempts only, but not completed suicides. Furthermore, multiple scales rely on self-reported current or future suicidal thoughts, which are often denied by individuals with high intent-to-die suicide attempts [44]. Multiple articles also report on 'some' predictive validity of scales by referencing singular items with predictive validity, but again, this is of marginal use (and cannot be replicated) unless the aforementioned subscales are revised for clinical use [5,27,35].

Designers of these scales should not be faulted given the well-documented heterogeneity of individuals with suicidal behaviors across gender, age, culture and constructs such as temperament, impulsivity, mental health and sociological variables [1,45,46]. Most scales are based off of risk factors compilations, but as noted in a recent meta-analysis by Large and colleagues, while 3% of high-risk patients will die by suicide within 1 year of hospital discharge, 60% of suicide completers will be classified as 'low risk.' As noted in an option article on the state of the risk assessment tools, 'risk categorization is of no value in attempts to decrease the numbers of patients who will commit suicide after discharge' [47].

Cognitive risk assessment tools: a new frontier?

Given the aforementioned limitations of clinical risk assessment tools, we believe the medical profession should widen its field of view in considering clinically valid predictors of suicidal behavior. A particularly promising route to consider is the use of cognitive tests investigating the implicit associations of patients presenting with suicide. Since patients who are acutely suicidal with high intent to die actually have motivation to deny suicidal thoughts or behavior, any measure that relies on the patient as an informant is likely to miss some of the most at-risk individuals [44].

Researchers postulate that differential coping mechanisms may underlie why risk factor

approaches are relatively ineffective at predicting suicidal behavior, particularly in the short-term. While risk factors may put individuals at risk for psychiatric distress, individual differences in coping mechanisms or cognitive schemas may provide a more proximal predictor. The role of implicit associations in guiding both behavior and beliefs is present across psychological domains, such as prejudices, and pathologies, including anxiety, phobia, and depression [48]. Theoretically, these factors may account for why, under psychiatric distress, some may use adaptive approaches (i.e., treatment seeking), while others may use maladaptive coping strategies and escape tendencies (i.e., suicidal behavior).

Investigating novel uses of cognitive tools in clinical settings is also well matched to address the specific 2009 National Institute Of Mental Health plan towards the identification of behavioral and biological markers associated with mental disorders. The 2012 USA Surgeon General also elicits a call for the research on the adoption, dissemination and implementation of guidelines for screening tools to help identify individuals at risk with use across community provider, primary care and emergency services. Compared with the substantial literature on traditional tools, there is a relative lack of research on cognitive behavioral measures, allowing for promising work to grow with room for discovery. In reviewing evidence for two of the most promising cognitive risk assessment tools, in light of the substantial limitations of traditional tools, it become evident that problem solving to determine how these tools can be integrated into some of the implementation challenges should be at the forefront of suicidal research policy.

Much like traditional tools, prospectively predicting suicide from cognitive tools originated from research on current suicidal populations, which found cognitive processing abnormalities among individuals with acutely suicidal behavior. These cognitive abnormalities are widespread and include different behavioral patterns in implicit task associations, including decision-making in a gambling task and attempted violence, implicit anxiety/depression with self-associations and suicidal ideation, and neuropsychological performance across a battery of measures [49–51]. Neurocognitive conceptions of acute suicidality was also supported by functional imaging studies and post-mortem abnormalities in regions specifically associated with neurocognitive–behavioral tasks [52–54].

Neurocognitive risk assessment tools

Suicidal Stroop task

Nock's group at Harvard University (MA, USA) has almost exclusively reported on research on the prospective predictive validity of suicidal tasks. The first report examining the role of implicit associations and suicidal outcomes used the a modified Stroop task to examine how much suicide-related words interfered with color-word naming compared with neutral or negatively valenced words [55,56]. Cognitive theories of emotional disorders suggest distinct attentional biases (and schemas influenced from information processing) increase vulnerability towards particular disorders. A sample of adults (n = 124) presenting to a psychiatric emergency department served as participants across a range of psychiatric emergencies. This research found that in addition to predicting the recency of past suicidal behavior, increased interference for suicide-related words predicted suicide attempt within a 6-month follow-up among the individuals who were able to be contacted (n = 60). This relationship was present after stringently controlling for other suicide risk factors (depressed mood and multiple attempts) and common clinical predictors (SSI, clinician prediction and patient prediction). While the small sample at follow-up is a substantial limitation, it is noteworthy that the Stroop interference effect was the only one out of all the aforementioned risk factors or predictors to significantly associate with future suicide attempts. The authors suggest that this bias in attention may reflect a high accessibility of suicide-related thoughts.

Death/Suicide Implicit Association Task

The next tool reported on by Nock's group was a modified implicit association task, which investigated mental associations participants have between self ('me') versus other ('not me') and words associated with life and death/suicide [29]. In this study, participants were recruited from the psychiatric emergency room and followed for 6 months through psychiatric admittances and phone interviews. Using the standard algorithms for calculating death/suicide Implicit Association Task (IAT) scores, researchers found that 'positive scores' indicating higher death/suicide self-associations were associated with baseline measures of suicide attempt presentation, current depressive disorder and history of prior suicide attempts (but not nonsuicide self-injury) [29,57].

Controlling for these factors, the IAT prospectively predicted the occurrence of future suicide attempts above all reported explicit patient, clinical and self-report (SSI) factors. Finally, a split using a bias score of '0' (stronger associations with self and death/suicide vs life) was suggested with a 6-month sensitivity/specificity of 0.50/0.81.

Summary of neurocognitive risk assessment tools

While very much a nascent field, the clinical application of neurocognitive suicide risk assessment tools offers both exciting findings and advantages over traditional self-report measures. Most importantly, these tools seem to be able to identify individuals at imminent risk for suicidal behavior with reasonable sensitivity and specificity at 6-month follow-up, as opposed to traditional tools, which report predictive validity over much longer time scales, or with poor sensitivity/specificity indices. Additionally, these measures are sensitive to changes over time (i.e., recency of attempt), which suggests they could be readministered to track changes in suicidality and are less likely related to stable trait characteristics. Importantly, each of the aforementioned studies also determined that these tasks could be administered in emergency psychiatric settings (hospital beds, waiting areas and small offices) and could discriminate from a wide pool of patients instead of a more narrowly defined group such as those with prior suicide attempts. Of final note is that both the Suicide Stroop and Death/Suicide IAT provide better predictive validity than clinical prediction, patient prediction or scores on Beck's SSI, and remained a significant predictor when controlling for these factors in addition to any depressive disorders and multiple prior attempts [29,56].

Substantial limitations in researching neurocognitive risk assessment tools should also be mentioned, including a lack of validation in individuals with cognitive impairment, substance intoxication, violent or agitated behavior, or diverse populations. It is unknown what proportion of acute psychiatric individuals would be excluded for these reasons, or if the tests would have validity administered at later time points following the reduction of agitated behavior or intoxication impairment. Nock and colleagues' results could help address these important questions.

Conclusion

In sum, we believe that if suicide risk assessment tools are to be implemented in acute clinical settings to guide clinical decisions, it is necessary for them to have validated predictive utility. Given the lack of such evidence for many currently implemented measures, it may be misleading to label a suicidal assessment tool as 'validated' unless it has a consistent evidence base in relation to future suicidal behavior for the specifically assessed population. The current widespread implementation of several of these scales without such an evidence base should be reconsidered. Multiple experts have echoed these sentiments with reports that the traditional tools have ostensibly no practical use in for advising acute clinical behavior, despite that fact that this is how they are commonly used in clinical settings [15]. Some proponents may argue that an inconsistently validated scale is better than no scale, but such usage can afford care providers a perception of increased suicidal prediction competency resulting in less attentional resources or clinical judgment for at-risk individuals.

However, it is important to note that, if traditional risk assessment tools may not be valid as clinical indicators of acute suicide risk, they can have utility as lists of questions relevant for clinical interviews that may add to an attending clinician's expertise. Research does not suggest that clinicians should ignore information about a client's mood, hopelessness, past suicidal behavior, demographics or impulsivity, but rather that risk 'categorization' based on such factors is unlikely to be predicatively valid or useful [15,47].

Given the general lack of supportive evidence for suicide risk assessment tools to categorize high-risk patients, and the practical difficulty in obtaining such evidence, should the argument be made to abandon the use of risk assessment tools altogether? Suicide, although a leading cause of death worldwide, is a relatively low prevalence outcome. This is one of the primary reasons why little evidence exists for suicide prevention. Researchers have focused on higher prevalence events, such as suicidal ideation or attempts, and have favored cross-sectional investigations instead of the prospect of costly and difficult prospective studies. Large epidemiological studies, necessary for sufficient power to detect suicide outcomes, are lacking with regards to suicide assessment tools. However, despite these inherent challenges, we argue

that the investigation of suicide risk assessment tools should continue. The primary rationale for continuing this direction is the ongoing public health concern of suicide. Additionally, the development of a validated suicide assessment tool is extremely important from a standard-of-care perspective. At this point the assessment of suicide risk is highly variable both across and within care settings. A standardized tool would enhance clinical attention towards suicide risk, improve documentation, and promote the development of multicenter research endeavors that could generate the much-needed large scale studies capable of examining future suicide risk.

In conclusion, we propose that serious attentional and research resources be paid to experiment with new techniques for implementing and building the evidence base for suicide risk assessment tools. As clinicians and researchers, we need to consider alternatives to the conventional risk factor checklists. A good example and promising direction is cognitive risk assessment tools. This could be an invaluable opportunity for the psychiatric field to incorporate evidence-based theories of cognitive schemas, well validated across numerous clinical psychology populations, to directly improve medical outcomes and patient care. Before implementing these on a large scale more widespread confirmation is needed across populations with different proposals to maximize efficient and effective implementation and interpretation.

Future perspective

Considering the implementation of cognitive tools in medical settings, it is relevant to consider how they might be utilized. One option would be to include them as a standardized cognitive assessment tool for all patients presenting with psychiatric distress. The score could be automatically calculated and display results to the attending physician, much like a readout from standardized blood pressure instruments. This readout could then be utilized in the context of the clinical interview and presenting case, much like other standard medical assessment tools. A related future direction might include a composite battery of tasks, including cognitive risk assessment, biological measures or self-report indicators, which could more acurately identify individuals at risk along different dimensions [58]. At this point, it is unknown, among the clinically validated measures, if individuals not identified by one tool might be better identified by another tool. Once again, more coordinated research is needed in clinical settings to both assess patients and follow-up for outcome data. This is, no doubt, a daunting task that requires motivated stakeholders at every level of hospital staff, researchers, and administrations. However, we feel that it is essential if standardized suicide risk assessment is to be a priority for informing decisions about patient care.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

References

Papers of special note have been highlighted as:

- ▪ of interest
- ▪▪ of considerable interest

1 Cavanagh J, Carson A, Sharpe M, Lawrie S. Psychological autopsy studies of suicide: a systematic review. *Psychol. Med.* 33(3), 395–405 (2003).

2 Nordentoft M, Mortensen PB, Pedersen CB. Absolute risk of suicide after first hospital contact in mental disorder. *Arch. Gen. Psychiatry* 68(10), 1058–1064 (2011).

3 Tidemalm D, Långström N, Lichtenstein P, Runeson B. Risk of suicide after suicide attempt according to coexisting psychiatric disorder: Swedish cohort study with long term follow-up. *BMJ* 337(8), 2205 (2008).

4 Qin P, Nordentoft M. Suicide risk in relation to psychiatric hospitalization: evidence based on longitudinal registers. *Arch. Gen. Psychiatry* 62(4), 427–432 (2005).

5 Beck AT, Steer RA. Clinical predictors of eventual suicide: a 5- to 10-year prospective study of suicide attempters. *J. Affect. Disord.* 17(3), 203–209 (1989).

6 Nock MK, Hwang I, Sampson N *et al.* Cross-national analysis of the associations among mental disorders and suicidal behavior: findings from the WHO World Mental Health Surveys. *PLoS Med.* 6(8), e1000123 (2009).

▪▪ **Highlights the challenges of relying on clinical judgement for suicide risk, including the tendency of a clinician to rely on epidemiological risk factors (e.g., male gender) that are not predictively accurate. It further highlights inconsistencies in the judgement of absolute risk for suicide risk between clinicians.**

7 Paterson B, Dowding D, Harries C, Cassells C, Morrison R, Niven C. Managing the risk of suicide in acute psychiatric inpatients: a clinical judgement analysis of staff predictions of imminent suicide risk. *J. Ment. Health* 17(4), 410–423 (2008).

8 Cochrane-Brink KA, Lofchy JS, Sakinofsky I. Clinical rating scales in suicide risk assessment. *Gen. Hosp. Psychiatry* 22(6), 445–451 (2000).

9 Gever J. DSM-V draft promises big changes in some psychiatric diagnoses. *MedPage Today*, 10 February 2010.

10 Hughes CW. Objective assessment of suicide risk: significant improvements in assessment, classification, and prediction. *Am. J. Psychiatry* 168(12), 1233–1234 (2011).

11 Bürk F, Kurz A, Möller HJ. Suicide risk scales: do they help to predict suicidal behaviour? *Eur. Arch. Psychiatry Neurol. Sci.* 235(3), 153–157 (1985).

12 Posner K, Brown GK, Stanley B *et al.* The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *Am. J. Psychiatry* 168(12), 1266–1277 (2011).

13 Gould MS, Greenberg TE, Velting DM, Shaffer D. Youth suicide risk and preventive interventions: a review of the past 10 years. *J. Am. Acad. Child Adolesc. Psychiatry* 42(4), 386–405 (2003).

14 Waern M, Rubenowitz E, Wilhelmson K. Predictors of suicide in the old elderly. *Gerontology* 49(5), 328–334 (2003).

15 Randall JR, Colman I, Rowe BH. A systematic review of psychometric assessment of self-harm risk in the emergency department. *J. Affect. Disord.* 134(1), 348–355 (2011).

16 McMillan D, Gilbody S, Beresford E, Neilly L. Can we predict suicide and non-fatal self-harm with the Beck Hopelessness Scale? A meta-analysis. *Psychol. Med.* 37(6), 769–778 (2007).

17 Beck AT, Davis JH, Frederick C *et al. Suicide Prevention in the 70's.* Resnik H, Hathorne BC (Eds). Center for Studies of Suicide Prevention, National Institute of Mental Health, DC, USA (1973).

18 Beck AT, Beck R, Kovacs M. Classification of suicidal behaviors: I. Quantifying intent and medical lethality. *Am. J. Psychiatry* 132(3), 285–287 (1975).

19 Beck AT, Steer RA. Hopelessness and eventual suicide: a 10-year prospective study of patients hospitalized with suicidal ideation. *Am. J. Psychiatry* 142, 559–563 (1985).

20 Beck AT, Brown G, Berchick RJ, Stewart BL, Steer RA. Relationship between hopelessness and ultimate suicide: a replication with psychiatric outpatients. *Am. J. Psychiatry* 147(2), 190–195 (1990).

21 Brown GK, Beck AT, Steer RA. Risk factors for suicide in psychiatric outpatients: a 20-year prospective study. *Risk* 68(3), 371–377 (2000).

22 Beck AT, Steer RA, Trexler LD. Alcohol abuse and eventual suicide: a 5-to 10-year prospective study of alcohol-abusing suicide attempters. *J. Stud. Alcohol* 50(3), 202–209 (1989).

23 Oquendo MA, Galfalvy H, Russo S *et al.* 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* 161(8), 1433–1441 (2004).

■ **Higlights the inaccuracies of exploratory risk categorization cutoff scores using the Beck Depression Inventory. While in the exploratory model, an effective cutoff score was determined; this score was not effectively replicated in the confirmatory model.**

24 Hartl TL, Rosen C, Drescher K, Lee TT, Gusman F. Predicting high-risk behaviors in veterans with posttraumatic stress disorder. *J. Nerv. Ment. Dis.* 193(7), 464–472 (2005).

25 Freedenthal S. Assessing the wish to die: a 30-year review of the suicide intent scale. *Arch. Suicide Res.* 12(4), 277–298 (2008).

26 Niméus A, Én M, Träskman-Bendz L. High suicidal intent scores indicate future suicide. *Arch. Suicide Res.* 6(3), 211–219 (2002).

27 Stefansson J, Nordström P, Jokinen J. Suicide Intent Scale in the prediction of suicide. *J. Affect. Disord.* 136(1–2), 167–171 (2012).

28 Tejedor MC, Diaz A, Castillón JJ, Pericay JM. Attempted suicide: repetition and survival findings of a follow-up study. *Acta Psychiatr. Scand.* 100(3), 205–211 (1999).

■■ **Introduces the suicide/self-implicit association task as a promising risk assessment tool for acute (6-month) suicide risk.**

29 Nock MK, Park JM, Finn CT, Deliberto TL, Dour HJ, Banaji MR. Measuring the suicidal mind implicit cognition predicts suicidal behavior. *Psychol. Sci.* 21(4), 511–517 (2010).

30 Cull JG, Gill WS. *Suicide Probability Scale.* Western Psychological Services, CA, USA, (1982).

31 Brown GK. *A Review of Suicide Assessment Measures for Intervention Research with Adults and Older Adults*. University of Pennsylvania, PA, USA (2001).

32 Naud H, Daigle MS. Predictive validity of the Suicide Probability Scale in a male inmate population. *J. Psychopathol. Behav. Assess.* 32(3), 333–342 (2010).

33 Larzelere RE, Smith GL, Batenhorst LM, Kelly DB. Predictive validity of the suicide probability scale among adolescents in group home treatment. *J. Am. Acad. Child Adolesc. Psychiatry* 35(2), 166–172 (1996).

34 Huth-Bocks AC, Kerr DC, Ivey AZ, Kramer AC, King CA. Assessment of psychiatrically hospitalized suicidal adolescents: self-report instruments as predictors of suicidal thoughts and behavior. *J. Am. Acad. Child Adolesc. Psychiatry* 46(3), 387–395 (2007).

35 Bolton J, Spiwak R, Sareen J. Predicting suicide attempts with the SAD PERSONS scale: a longitudinal analysis. *J. Clin. Psychiatry* 73(6), e735 (2012).

36 Lecrubier Y, Sheehan, DV, Weiller E *et al.* The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *European Psychiatry* 12, 224–231 (1997).

37 Roaldset JO, Linaker OM, Bjørkly S. Predictive validity of the MINI suicidal scale for self-harm in acute psychiatry: a prospective study of the first year after discharge. *Arch. Suicide Res.* 16(4), 287–302 (2012).

38 Yen S, Shea MT, Walsh Z *et al.* Self-harm subscale of the Schedule of Nonadaptive and Adaptive Personality (SNAP): predicting suicide attempts over 8 years of follow-up. *J. Clin. Psychiatry* 72(11), 1522–1128 (2011).

39 Nimeus A, Alsen M, Traeskman-Bendz L. The Suicide Assessment Scale: an instrument assessing suicide risk of suicide attempters. *Eur. Psychiatry* 15(7), 416–423 (2000).

40 Waern M, Sjöström N, Marlow T, Hetta J. Does the Suicide Assessment Scale predict risk of repetition? A prospective study of suicide attempters at a hospital emergency department. *Eur. Psychiatry* 25(7), 421–426 (2010).

41 Dumais A, Lesage A, Alda M *et al.* Risk factors for suicide completion in major depression: a case–control study of impulsive and aggressive behaviors in men. *Am. J. Psychiatry* 162(11), 2116–2124 (2005).

42 Jokinen J, Åsberg M, Nordström P. FC13–02-Karolinska interpersonal violence scale predicts suicide in suicide attempters. *J. Clin. Psychiatry* 71,(8) 1025–1032 (2010).

43 Moberg T, Nordström P, Forslund K, Kristiansson M, Åsberg M, Jokinen J. CSF 5-HIAA and exposure to and expression of interpersonal violence in suicide attempters. *J. Affect. Disord.* 132(1–2), 173–178 (2011).

44 Busch KA, Fawcett J, Jacobs DG. Clinical correlates of inpatient suicide. *J. Clin. Psychiatry* 64(1), 14–19 (2003).

45 Engström G, Alsén M, Gustavsson P, Schalling D, Träskman-Bendz L. Classification of suicide attempters by cluster analysis: a study of the temperamental heterogeneity in suicidal patients. *Pers. Individ. Dif.* 21(5), 687–695 (1996).

46 Mcgirr A, Seguin M, Renaud J, Benkelfat C, Alda M, Turecki G. Gender and risk factors for suicide: evidence for heterogeneity in predisposing mechanisms in a psychological autopsy study. *J. Clin. Psychiatry* 67(10), 1612–1617 (2006).

■■ **Identifies the challenges of risk categorization in a systematic meta-analysis of classical risk assessment scales.**

47 Large M, Sharma S, Cannon E, Ryan C, Nielssen O. Risk factors for suicide within a year of discharge from psychiatric hospital: a systematic meta-analysis. *Aust. NZ J. Psychiatry* 45(8), 619–628 (2011).

48 Houwer JD. The Implicit Association Test as a tool for studying dysfunctional associations in psychopathology: strengths and limitations. *J. Behav. Ther. Exp. Psychiatry.* 33(2), 115–133 (2002).

49 Jollant F, Bellivier F, Leboyer M *et al.* Impaired decision making in suicide attempters. *Am. J. Psychiatry* 162(2), 304–310 (2005).

50 Glashouwer KA, De Jong PJ, Penninx BW, Kerkhof AJ, Van Dyck R, Ormel J. Do automatic self-associations relate to suicidal ideation? *J. Psychopathol. Behav. Assess.* 32(3), 428–437 (2010).

51 Keilp JG, Sackeim HA, Brodsky BS, Oquendo MA, Malone KM, Mann JJ. Neuropsychological dysfunction in depressed suicide attempters. *Am. J. Psychiatry* 158(5), 735–741 (2001).

52 Arango V, Underwood MD, Mann J. Postmortem findings in suicide victims. implications for *in vivo* imaging studies. *Ann. NY Acad. Sci.* 836(1), 269–287 (1997).

53 Fukui H, Murai T, Fukuyama H, Hayashi T, Hanakawa T. Functional activity related to risk anticipation during performance of the Iowa Gambling Task. *Neuroimage* 24(1), 253–259 (2005).

54 Bechara A, Tranel D, Damasio H, Damasio AR. Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cereb. Cortex* 6(2), 215–225 (1996).

55 Stroop JR. Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18(6), 643–662 (1935).

■■ **Introduces the suicide–Stroop association task as a promising risk assessment tool for acute (6-month) suicide risk.**

56 Cha CB, Najmi S, Park JM, Finn CT, Nock MK. Attentional bias toward suicide-related stimuli predicts suicidal behavior. *J. Abnorm. Psychol.* 119(3), 616–622 (2010).

■ **Seminal implicit association task paper, important in understanding the rationale and computations behind implicit bias associations.**

57 Greenwald AG, Nosek BA, Banaji MR. Understanding and using the implicit association test: I. An improved scoring algorithm. *J. Pers. Soc. Psychol.* 85(2), 197–216 (2003).

58 Jokinen J, Nordström A-L, Nordström P. CSF 5-HIAA and DST non-suppression – orthogonal biologic risk factors for suicide in male mood disorder inpatients. *Psychiatry Res.* 165(1–2), 96–102 (2009).