Summer 7-29-2021

# Internal consistency and concurrent validity of self-report components of a new instrument for the assessment of suicidality, the Suicide Ideation and Behavior Assessment Tool (SIBAT)

Eric C. Chan

Keanna Wallace

Esther H. Yang

Leslie Roper

Garima Aryal

*See next page for additional authors*

Psychiatry Research 304 (2021) 114128

Contents lists available at ScienceDirect

Psychiatry Research

journal homepage: www.elsevier.com/locate/psychres

# Internal consistency and concurrent validity of self-report components of a new instrument for the assessment of suicidality, the Suicide Ideation and Behavior Assessment Tool (SIBAT)

Eric C. Chan [a,b,*], Keanna Wallace [c], Esther H. Yang [c], Leslie Roper [c], Garima Aryal [c], Rohit J. Lodhi [d], Andrius Baskys [a,e,f], Richard Isenberg [g], Patrick Carnes [a], Bradley Green [h], Katherine J. Aitchison [a,c]

[a] Department of Psychiatry, University of Alberta, Edmonton, Alberta, Canada
[b] Department of Psychiatry, University of Calgary, Calgary, Alberta, Canada
[c] Department of Medical Genetics, University of Alberta, Edmonton, Alberta, Canada
[d] Department of Psychiatry, University of Saskatchewan, Saskatoon, SK, Canada
[e] Graduate College of Biomedical Sciences and University Medical Center, Western University of Health Sciences, Pomona, California, United States
[f] Memory Disorders and Genomic Medicine Clinic, Riverside, California, United States
[g] American Foundation for Addiction Research, Psychological Counseling Services, Scottsdale, Arizona, United States
[h] Department of Psychology, University of Texas at Tyler, Tyler, United States

ARTICLE INFO

*Keywords:*
E-mental health
Psychometric
Psychiatric status rating scale
Risk assessment
Self harm
Validation
Mobile applications

ABSTRACT

This study aimed to assess the internal consistency of self-report components of the Suicide Ideation and Behavior Assessment Tool (SIBAT) and validate it with relevant elements of the Mini International Neuropsychiatric Interview (MINI). The SIBAT is a newly developed instrument for the evaluation of suicidality. In this study, we invited university students and trainees participating in a study of addictions to complete the self-report component of the SIBAT as an add-on study. We evaluated the internal consistency of the self-report component of the SIBAT and validated it against the suicidality component of the MINI. Data were analysed using both complete case analysis and multiple imputation. SIBAT data were collected for 394 participants, 314 of whom had also completed the MINI. The internal consistency of modules 2, 3, and 5 of the SIBAT was high. Each item from module 5 had a statistically significant association with the corresponding item from the MINI. The sum of scores from modules 2 and 3 had a moderate correlation with the assessment of suicide risk determined by the MINI, and a strong correlation with the total score of SIBAT module 5. The completion median time of modules 2, 3 and 5 was 14.3 min.

## 1. Introduction

The evaluation of suicide risk is an important aspect of psychiatric care. With a rate of 11.0 per 100 000 population in 2016 (Statistics Canada, 2021a), suicide was the ninth leading cause of death in Canada across all ages and was the second leading cause of death in those aged 15 to 24 (Statistics Canada, 2021b). Data from the Public Health Agency of Canada indicate that 11.8% of Canadians report thoughts of suicide in their lifetime and 2.5% of Canadians report having thoughts of suicide in the past year (Public Health Agency of Canada, 2019). Many completed suicides are preceded by contact with health care services. A 2011 study of suicide in Alberta reports that 58% of suicides were preceded by an emergency department visit and 28% were preceded by an inpatient hospital discharge in the preceding year (Morrison and Laing, 2011). Given the frequency of contact with health services prior to suicide, accurate identification and risk stratification is an important step in ensuring that appropriate interventions are provided to those at elevated risk.

The National Action Alliance for Suicide Prevention in the United States has recommended that all patients identified as being at risk of suicide be assessed using a standardized instrument or scale (National Action Alliance for Suicide Prevention: Transforming Health Systems

* Corresponding author.
*E-mail addresses:* echantai@ualberta.ca (E.C. Chan), kaitchis@ualberta.ca (K.J. Aitchison).

https://doi.org/10.1016/j.psychres.2021.114128
Received 17 March 2021; Received in revised form 14 July 2021; Accepted 15 July 2021
Available online 29 July 2021

Initiative Work Group 2018). There is no consensus, however, on the best screening and assessment tools for use in clinical settings (National Action Alliance for Suicide Prevention: Transforming Health Systems Initiative Work Group 2018). Multiple recent studies evaluating the utility of currently available instruments indicate that none predicted suicide or suicidal behavior with sufficient accuracy to be relied on in clinical settings (Carter et al., 2017; Chan et al., 2016; Runeson et al., 2017; Steeg et al., 2018). Similarly, the third edition of the American Psychiatric Association Practice Guidelines note that, while suicide assessment instruments may have clinical utility in assisting the clinician to develop a thorough line of questioning, no scale has been shown to produce a clinically useful score for suicide prediction (American Psychiatric Association, 2016).

One of the major factors limiting the predictive validity of tools for suicide risk assessment is the rarity of suicide as an event. As Pokorny (1983) and Rosen (1954) have previously discussed, even a tool with high sensitivity and specificity can have a low positive predictive value given the infrequency of suicide in the general population. Estimation of suicide risk is further complicated by the multifactorial nature of suicidality. Suicide is influenced by a wide variety of biological, psychological and social factors (Mental Health Commission of Canada, 2018; (Hawgood and De, 2016)). Some of these factors, such as physical health and family connectedness, can change drastically over short periods of time. Furthermore, recent data suggest that suicidality itself can fluctuate over the course of days and possibly even hours (Giddens and Sheehan, 2014; Kleiman et al., 2017). Given the limitations of currently

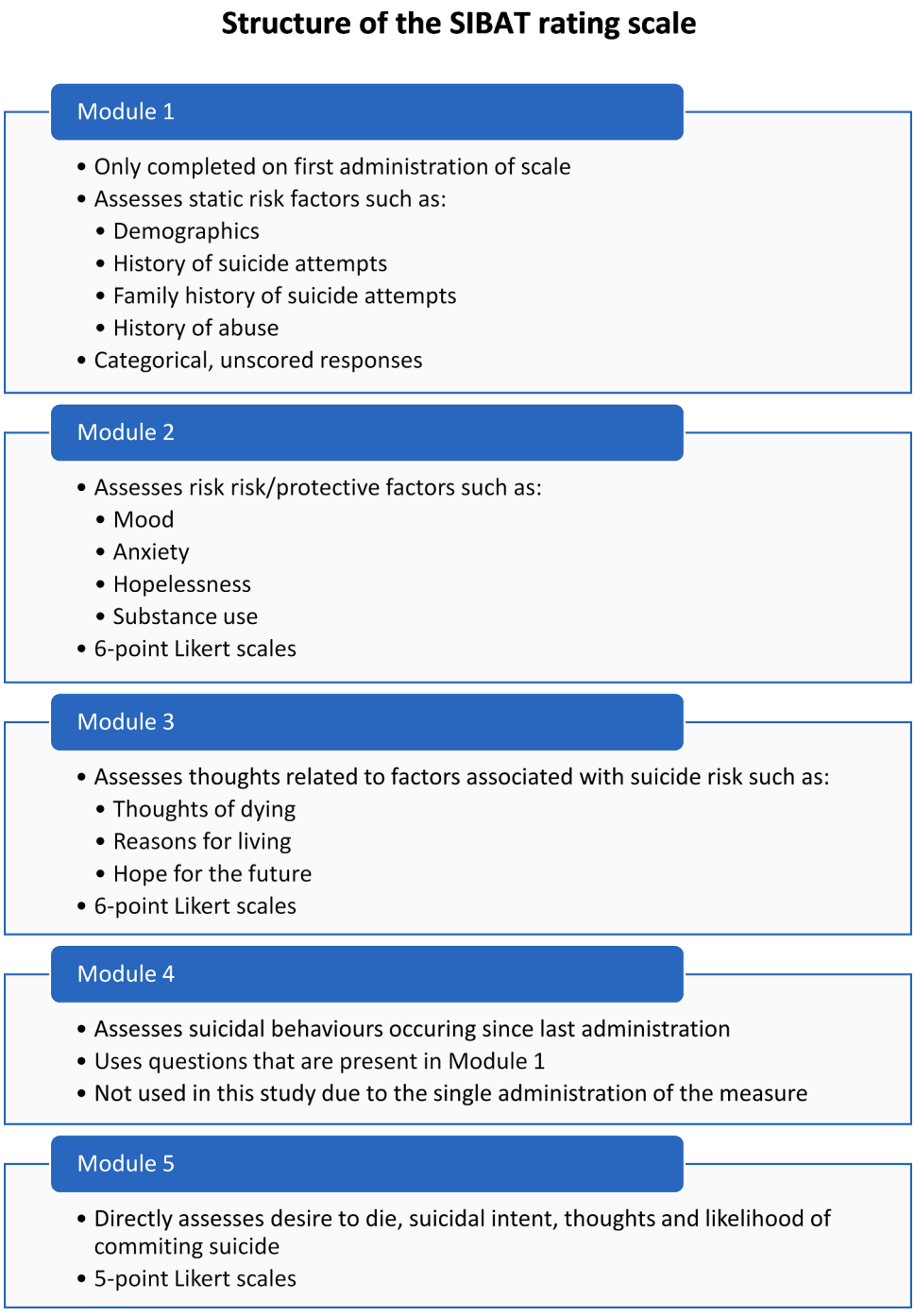


**Fig. 1.** Structure of the SIBAT rating scale.

available instruments for suicide risk prediction, the National Action Alliance for Suicide Prevention in the United States identified the "development of validated procedures that can determine the degree of suicide risk" as an aspirational goal in the prevention of suicide (Boudreaux and Horowitz, 2014).

The Suicide Ideation and Behavior Assessment Tool (SIBAT) is a new, comprehensive rating scale developed for the assessment and monitoring of suicidality by a group of clinical trial and academic experts in scale development. It is a comprehensive tool made up of both clinician-assessed and self-report components and previous data indicate that it is sensitive to changes over time (Alphs et al., 2015, 2018a, 2018b; Alphs et al., 2018c; Turkoz et al., 2018; Williamson et al., 2017). Furthermore, the SIBAT has shown inter- and intra-rater reliability and adequate mapping to the Columbia Classification Algorithm of Suicide Assessment in a recently published study (Alphs et al., 2020). The structure of the SIBAT has been detailed elsewhere and the SIBAT has been used in a phase 3 program for evaluating the efficacy and safety of a rapid acting antidepressant (Fu et al., 2020; Ionescu et al., 2021). Our goal in this study was to assess the internal consistency of the SIBAT and cross-validate it with the Mini International Neuropsychiatric Interview (MINI), in order to evaluate the concurrent validity of responses obtained.

## 2. Methods

### 2.1. Assessment tools

#### 2.1.1. Suicide ideation and behavior assessment tool (SIBAT)

The self-report component of the SIBAT consists of five modules that are described in Fig. 1. Examples of question stems from the SIBAT are included in Appendix 1. As items in modules 1 are categorical responses and not scored, they were excluded in the comparison of SIBAT and MINI responses. Module 2 consists of 21 items assessing risk and protective factors such as emotional state, interpersonal relationships, and other psychiatric symptoms scored on a 6-point Likert scale and 10 unscored categorical (binary and ordinal) items assessing substance use and history of incarceration. Module 3 consists of 48 items assessing thoughts related to risk factors including thoughts of death/dying, hopelessness, helplessness, guilt, stressors, and self-esteem scored on a 6-point Likert scale. Module 5 consists of 4 items assessing active suicidal ideation/intent scored on a 5-point Likert scale. Items scored on Likert scales were converted to scores of 0–6 in modules 2 and 3 and 0–5 in module 5. In total, modules 2, 3 and 5 have a total of 73 scored items that were used in total score calculation. [Insert Fig. 1 near here]

The SIBAT was administered using either an application (app) developed using the AppSheets platform (Fig. 2) or the Qualtrics interface. [Insert Fig. 2 near here]

#### 2.1.2. Suicidality component of the mini international neuropsychiatric interview

The suicidality component of the Mini International Neuropsychiatric Interview (MINI) consists of 19 Yes/No items and categorizes the participant into low, medium and high-risk categories based on their responses. It has been shown to be a significant predictor of suicidal behavior (Roaldset et al., 2012). This scale was administered to university students and trainees participating in a study of addictions and it was completed separately from the SIBAT. Depending on the timing of the participant's entry into the study, the suicidality component of the MINI may have been completed before or after the SIBAT.

### 2.2. Participant recruitment

This study was approved by the Health Research Ethics Board at the University of Alberta. Participants were recruited from University of Alberta students and trainees participating in a study of addictions which includes a first phase of data entry via the Qualtrics platform, and a second phase, in which participants were invited to complete the Mini International Neuropsychiatric Interview (MINI) 7.0.2 through a link to the questionnaire provided online. These participants were enrolled as the control cohort of the study of addictions. The SIBAT was added partway into control recruitment for the study of addictions as an optional additional measure for participants to complete. Participants were aware that the SIBAT was optional and no remuneration was offered for this.


**Fig. 2.** Screenshot of the SIBAT application.

Invitations to complete the SIBAT were sent to 913 participants of the aforementioned addictions study. Of those invited, 411 began the SIBAT; however, 17 of these participants did not complete all required modules and were excluded from the study. Complete SIBAT data were collected for 394 participants, 204 of whom completed the SIBAT using the AppSheets platform and the other 180 completed the SIBAT using the Qualtrics interface. These data were used to evaluate the internal consistency of the SIBAT. At the time of data analysis, 314 of these participants had also completed the MINI and these data were used to compare responses to the SIBAT with the MINI. Further information on the process for participant recruitment is available in the supplementary. A recruitment flowchart is shown in Fig. 3. [Insert Fig. 3 near here]

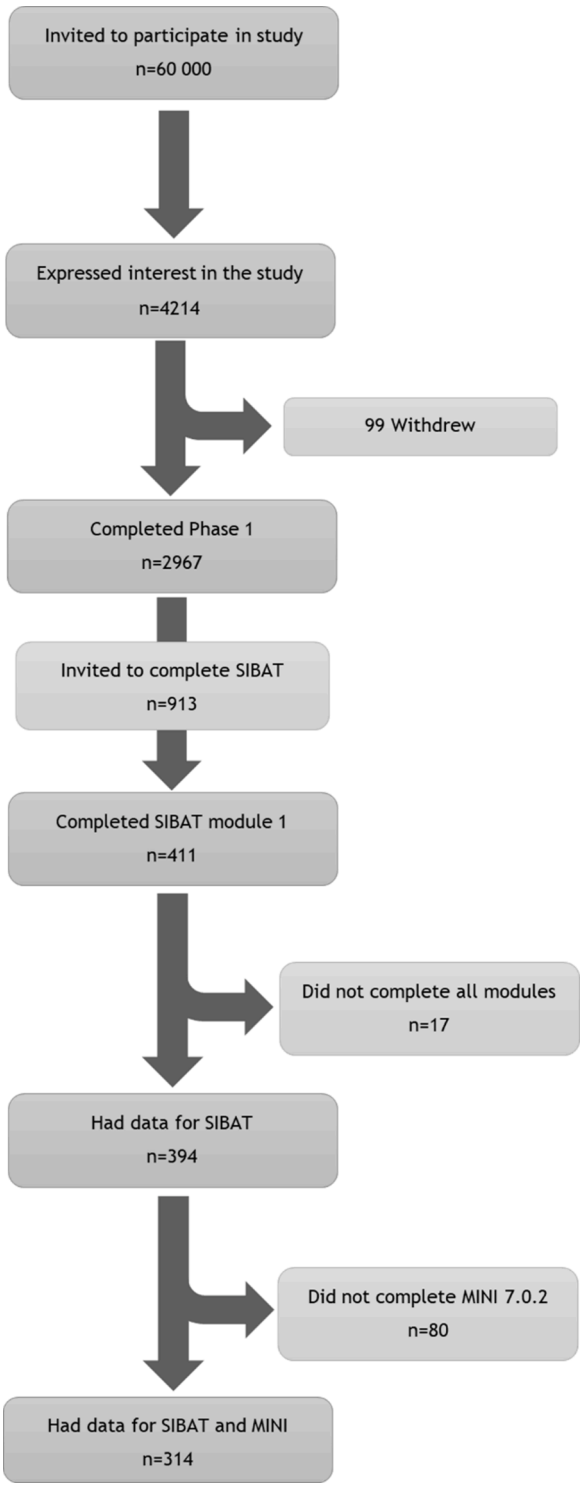


**Fig. 3.** Recruitment flowchart.

### 2.3. Statistical analyses

Statistical analysis was performed using IBM SPSS Statistics Version 25. Data were analysed using complete case analysis as the primary method of data analysis and multiple imputation as the secondary method of data analysis. Little's Missing Completely at Random (MCAR) test was conducted in order to test whether the data were missing completely at random. Multiple imputation of responses from modules 2, 3 and 5 (73 variables total) was performed by fully conditional specification with 8 imputations. It has previously been suggested that the number of imputations should be similar to the percentage of incomplete cases (Von Hippel, 2009). As 32/394 (8.1%) cases were missing responses, we used eight imputations for this analysis.

The internal consistency for module 5 was calculated using Cronbach's alpha coefficient. The number of items in module 2 and in module 3 were 21 and 48 respectively. As scales with more than 14 items have been demonstrated to have high Cronbach's alpha irrespective of the internal consistency of the scale (Streiner, 2003), we calculated the internal consistency of modules 2 and 3 using item-total correlations for these modules.

In order to assess the criterion validity of the SIBAT, individual items from module 5 of the SIBAT were compared to similarly worded items from the MINI. We then compared the total score of modules 2 and 3 combined to the severity rating of the suicidality component of the MINI and the total score of SIBAT module 5 (as questions in module 5 ask about suicidality directly).

An estimate of the time required to complete modules 2, 3 and 5 was determined using timestamps of data collected by the AppSheets app ($n$ = 204). The Qualtrics interface did not collect data on start or completion times and so these data were not included in this analysis. The timestamp of module 1 submission was used as the starting time of module 2 and the timestamp of module 5 submission was used as the completion time (data on the start time of module 1 was not recorded). The difference between these two times was then used as an estimate of the time to complete the modules studied. As users did not necessarily need to complete the measure all in one sitting, some results greatly overestimated the duration of the measure; therefore, the median time for scale completion is reported.

## 3. Results

Only 4/73 items had greater than 2/394 (0.5%) missing responses. These items were module 2, item 2, "Over the past 7 days I have felt agitated" (4/394, 1.0%), module 3, item 5, "My spiritual/religious beliefs prevent me from ending my life" (6/394, 1.5%), module 3, item 7, "My concern for others prevents me from ending my life" (10/394, 2.5%), and module 3, item 8, "If I developed a life-threatening illness, I would make every effort to overcome it" (3/394, 0.8%). No case had more than 4/73 (5.5%) missing responses. As such, we report our findings using complete case analysis. Little's MCAR test suggested that data was missing completely at random ($\chi^2$ (1855, 394)= 1932.7, $p$ = 0.102). As such, we performed secondary analysis using multiple imputation. Similar findings for all outcomes were obtained when using multiple imputation.

The internal consistency of module 5 of the SIBAT was high (Cronbach's $\alpha$ = 0.87, $n$ = 4 items). Item-total correlations for module 2 ranged from 0.22 – 0.79, with 19/21 items having item-total correlations > 0.4. The remaining two items related to aggressive impulses and command auditory hallucinations. Item-total correlations for module 3 ranged from −0.20 – 0.79, with 41/48 items having item-total correlations > 0.4. Of the remaining 7 items, two had negative item-total correlation values. One item with a negative item-total correlation value assessed concern for others and the other item assessed fear of dying. The other items with item-total correlations < 0.4 assessed the role of spiritual/religious beliefs related to thoughts of dying, the presence of severe physical pain, the desire to improve one's life, the benefit from helping others and the desire to spend time with others.

The comparison of items from module 5 with items from the MINI is shown in Table 1. Each item from module 5 of the SIBAT had a statistically significant association with its corresponding MINI item. As three items were ordinal and 25% of cells had expected counts less than 5, they were compared using Fisher's Exact Test and validity coefficients could not be generated. The fourth item of module 5 had a moderate correlation with its corresponding item from the MINI. [Insert Table 1 near here]

The sum of scores from modules 2 and 3 had a moderate correlation

**Table 1**
Comparison of items from SIBAT module 5 to suicidality component of MINI 7.0.2.

| SIBAT | MINI (In the past month did you…) | Significance |
|---|---|---|
| Which of the following ratings best describes your desire to die in the past 7 days? | Think (even momentarily) that you would be better off dead or wish you were dead or needed to be dead? | $p < 0.001$[a] |
| Which of the following ratings best describes your thinking about suicide right now? | Think (even momentarily) about harming or of hurting or of injuring yourself with at least some intent or awareness that you might die as a result or think about suicide (i.e. about killing yourself)? | $p < 0.001$[a] |
| Which of the following ratings best describes your intent to end your life in the past 7 days? | Intend to act on thoughts of killing yourself? | $p = 0.029 – 0.032$[a] |
| Given your current thinking and past experience, which of the following best describes the likelihood that you attempt to end your life in the near future? | How likely are you to try to kill yourself within the next 3 months on a scale of 0–100%? | Spearman's rho $= 0.442$, $p < 0.001$ |

[a] Compared using Fisher's Exact Test.

with the assessment of suicide risk determined by the MINI (Spearman's rho = 0.44, $P < 0.001$), which assesses suicidality in the preceding month, and a strong correlation with the total score of SIBAT module 5 (Spearman's rho = 0.62, $P < 0.001$), which assesses suicidality in the preceding week.

The median time to completion of modules 2, 3 and 5 was 14.3 min. The majority of participants used in this analysis ($n = 175/204$, 85.8%) completed the measure within 30 min. Some users ($n = 7/204$, 3%) completed the modules on separate days.

## 4. Discussion

Overall, our findings indicate that the self-report component of the SIBAT has high internal consistency overall and items from module 5 have good concurrent validity with corresponding items from the suicidality component of the MINI. The total score of modules 2 and 3 was associated with the severity of suicide risk determined by the MINI and with the score from SIBAT module 5. Most participants completed the measure within 30 min.

Some specific items demonstrated poor item-total correlation. Some items focused on factors related to aggressive impulses, psychotic symptoms and somatic symptoms. These factors are less likely to be present in the university population compared to other populations such as patients with an established history of psychiatric illness and may contribute less to suicide risk in the participant population studied than in other populations. Two items, "My spiritual/religious beliefs prevent me from ending my life" and "My concern for others prevents me from ending my life", had both the highest rates of non-response and low item-total correlations.

The sum of scores from SIBAT modules 2 and 3 had a stronger correlation with module 5 than with the suicidality component of the MINI. There are a few possible reasons for this. SIBAT module 5 would have been completed around the same time as modules 2 and 3; whereas the MINI was completed at a different time, either before or after SIBAT completion. In addition, questions from the MINI asked about symptoms occurring in the preceding month, whereas SIBAT module 5 asked about symptoms occurring in the preceding week. The difference in timing of scale completion and in the time period assessed could both contribute to the increased correlation with SIBAT module 5. In addition, SIBAT module 5 responses were scored on a 5-point Likert scale whereas the MINI had yes/no responses. As such, the differences observed may be due to dichotomizing data resulting in lower correlations in general (Streiner, 2002). Furthermore, the increased granularity afforded by the greater range of possible responses may also explain the difference in correlations, as participants may have been hesitant to provide a positive response if they felt their symptoms were quite mild.

Even though modules 2, 3 and 5 combined had 74 items, the median time to completion was less than 15 min. This suggests that implementation of the SIBAT may be feasible in settings such as when they are awaiting assessment in emergency departments or clinic waiting rooms. The comprehensive nature of the SIBAT may bring areas of attention to the clinician's awareness, without requiring the clinician to ask extensive screening questions in a protracted assessment.

As the SIBAT has many items, analysis of such data in a large dataset would be facilitated by approaches such as artificial intelligence/machine learning. The use of machine learning in the study of psychiatric illness has been discussed as a tool that could lead to the development of new hypotheses around the nature of psychiatric illnesses themselves and the treatment therof (Oquendo et al., 2012). This approach is in keeping with the US National Institute of Mental Health Research Domain Criteria initiative, which promotes a dimensional approach to psychiatric research given the current limitations of research centered on syndrome-based clinical diagnoses (Sanislow, 2016). Given the complex, multifactorial and fluctuating nature of suicidality, such strategies may be helpful in improving our understanding of the way in which contributing risk factors interact (Giddens and Sheehan, 2014).

The categorization of patients at increased risk of suicide could enhance our ability to assess risk and may improve our ability to provide the optimal treatment for specific presentations. Numerous research groups have since used machine learning to classify and characterize groups of patients with increased suicide risk (Kessler et al., 2015; Kim et al., 2018; Morales et al., 2017; Oh et al., 2017; Sinyor et al., 2014; Walsh et al., 2017). Cluster analysis in a group of Korean patients presenting after a suicide attempt extracted two groups, one with more impulsive, low lethality attempts and another with more well-planned attempts using more lethal methods (Kim et al., 2018). Another study conducted in Toronto identified five clusters following analysis of data from a coroner's chart review of deaths ruled as suicide (Sinyor et al., 2014). Their findings suggest that individuals who die by suicide are more likely to have certain combinations of factors that predispose them to risk. While most traditional approaches to suicide risk assessment have focused on identifying the presence of risk factors universally associated with increased risk, these data suggest that the presence of specific constellations of factors may potentially increase risk in a synergistic fashion.

The development of algorithms to predict future suicidal behaviours is another potential application of machine learning. In a study of US Army soldiers, a suicide risk algorithm was generated using administrative data from a population of soldiers recently admitted for a psychiatric disorder.[33] Of suicides occurring in the following 12 months, 52.9% occurred within the 5% of participants predicted as having the highest suicide risk. Another study applied machine learning to data from a repository of electronic health records to develop a machine learning algorithm that predicted future suicide attempts (AUC = 0.84, precision = 0.79, recall = 0.95, Brier score = 0.14) (Walsh et al., 2017). Of note, both of these studies used data from large data repositories, and so factors related to a patient's current state such as hopelessness, affective symptoms, psychotic symptoms, sleep disturbance, recent stressful events or social isolation were not included.

One important factor in the performance of machine learning is the quality of data obtained. Improving the quality of data collection has been noted to be critical in realizing the full potential of machine learning methods in suicide research (Torous et al., 2018). The studies identified above used varying methods to collect data for analysis, with some studies using multiple rating scales (Kim et al., 2018; Morales et al., 2017; Oh et al., 2017). The variability in these approaches to data collection makes comparing results between groups more challenging.

The SIBAT offers a consistent and comprehensive approach to data collection that may have applications in machine learning approaches to suicide research. As mentioned above, many factors associated with suicide risk are typically not captured in administrative data sets as they often rely on subjective reporting and change over time. If they are assessed and documented at all, these factors are usually recorded in assessment notes. Due to the lack of a consistent structure in these data, analyzing these factors using machine learning becomes more challenging. Once validated for use in research settings, the SIBAT could allow us to collect data on these factors in a structured fashion more amenable to analysis using machine learning techniques. Furthermore, as machine learning approaches require very large data sets for optimal performance, the SIBAT could be useful as a standard measure of factors associated with suicidality, as it may allow data from different studies and different study groups to be combined for analysis using machine learning approaches. In addition, the use of a consistent measure across research groups may improve our ability to compare models developed in different settings and to assess the performance of models in different populations.

The SIBAT could be combined with machine learning strategies in various ways. It could be used to collect data from patients assessed in settings such as the emergency department, psychiatric inpatient units, or psychiatric outpatient clinics. These data could then be combined with data on future suicide attempts and death from suicide and the combined data set could be analyzed using supervised learning to predict which combinations of risk factors are most predictive of future adverse outcomes. Another approach could be to analyze data using cluster analysis. This may improve our understanding of the different combinations of factors that may increase an individual's risk of suicide and, through identification of specific subpopulations, may improve our ability to identify which treatments are most likely to benefit a patient with particular characteristics. Given the length of the SIBAT, machine learning techniques, such as dimensionality reduction, allow for item reduction and these findings could have implications for clinical assessments as well.

## 5. Limitations

The population examined in this study is not a group traditionally associated with being at elevated risk of suicide. As the most likely implementation of this scale would be in populations at higher risk, such as psychiatric inpatients or patients seen for suicidality in the emergency department, it is unclear whether these findings could be generalized to the populations in which the SIBAT would most likely be used. In addition, less than half of participants invited to the study had completed the SIBAT. It is possible that this relates to the fact that completion of the SIBAT was the only component of the AddGenes study for which no reimbursement was provided, which may have introduced selection bias. As such, we note that, while the SIBAT has demonstrated high internal consistency and concurrent validity in the population studied, further investigation in other settings will be necessary before this tool can be considered a valid measure.

In addition to administering the SIBAT in populations associated with elevated risk of suicide, further analytical approaches, such as item response theory and factor analysis are necessary before the SIBAT can be considered a valid and reliable measure. Due to the population assessed in this study, as well as the relatively low sample size studied when considering the SIBAT's length and complexity, it was felt that further data collection in future studies will be necessary in order to perform these steps. As such, while our findings are promising for the potential of the SIBAT, our findings should be considered preliminary at this time. Additional investigation in more relevant populations and using the aforementioned techniques need to occur in order to determine if the SIBAT is a valid and reliable scale.

Furthermore, as the SIBAT was only administered once in this study and no follow-up data were collected, measures such as test-retest reliability, scale adherence and most notably predictive validity could not be assessed. While our initial data suggest that the SIBAT may evaluate state-related components of suicidality, repeated, longitudinal administration of the SIBAT would be necessary to establish the sensitivity of the SIBAT to changes over time.

As noted in the introduction, no current tool has been shown to perform adequately in the prediction of suicide as an event. As this study compared data obtained using the SIBAT to the suicidality component of the MINI, it is important to note that the MINI itself has limitations in its use. Most items from the suicidality component of the MINI use dichotomous responses, which may lead to lower correlations with outcomes of interest. A previous study examining the ability of the MINI to predict future threats and acts of suicidal behavior or nonsuicidal self-injury found that the MINI had a positive predictive value of 43% using a threshold of $\geq$ 10 and a positive predictive value of 39% using a threshold of $\geq$ 6 (Ionescu et al., 2021). Given these limitations, comparison of data collected using the SIBAT to clinically relevant outcomes, such as future suicide attempts and death by suicide, may allow for greater insight into the predictive validity of this new instrument.

## 6. Future directions

Further studies assessing the dimensionality and construct validity of the SIBAT, especially in populations associated with higher risk of suicide, are necessary to establish the SIBAT as a valid and reliable scale. Factor analysis and item reduction will also allow for further development of the scale for future use. Additionally, studies in which the SIBAT is administered repeatedly over time and studies examining long-term clinically relevant outcomes, such as future suicide attempts or death by suicide, would help identify potential uses for the SIBAT in research and in practice.

As indicated in the discussion, future administration of the SIBAT could include data collection for use in machine learning. As the SIBAT covers a large number of factors associated with suicide risk, data collected may improve our understanding of suicidality and assist us in developing treatments for different presentations.

## 7. Conclusions

The findings of this study suggest that the self-report component of the SIBAT has good internal consistency overall and items in module 5 have good concurrent validity with the suicidality component of the MINI. The total score of modules 2 and 3 combined had a moderate association with the suicidality component of the MINI and a stronger association with SIBAT module 5. The median time to complete modules 2, 3 and 5 was 14.3 min. While these data are preliminary, they support further assessment of the validity of the SIBAT in populations at higher risk of suicide in order to establish if the SIBAT is a valid and reliable scale. Assessment of the SIBAT involving repeat assessment, association with long-term outcomes and refinements of the measure could provide insight into its potential use in research and clinical settings.

## Author statement

Thank you for taking the time to review our submission and providing us with feedback. We have made changes to our manuscript and highlights in accordance with your suggestions. We would like to note that we were unable to make revisions to the analytical methods used in this submission as we did not feel that some of the methods suggested, such as factor analysis, would perform adequately given the population studied and the heterogeneity of suicidality as a construct. We have made revisions to the text however, to make it more clear that our findings are intended as a preliminary results that support further development and investigation of the SIBAT, rather than establish the SIBAT as a valid and reliable measure on their own. We believe that future studies involving large populations from samples deemed at-risk

of suicide are necessary for further validation of the SIBAT and feel our findings suggest that these studies would be worthwhile. Some other comments related to the development of the SIBAT are not mentioned in our submission as we were not involved in that stage of the process, but we agree that further description of these processes will also be necessary and will discuss with the team responsible for the selection of items in the future. We hope that you will find our submission to be more appropriate for publication with the changes we have made.

## Declaration of Competing Interest

KJA was awarded a fellowship grant from Janssen Inc., Canada, to support the research activities of a fellow (trainee, ECC). The funds were used to support a small proportion of research assistant (GA, KW, EY) time involved in participant recruitment, data collection, and database maintenance. However, none of the funds were used for development of the app (which was done by ECC in his own time). The SIBAT was developed by The SIBAT Consortium, a group of clinical trial and academic experts in scale development, suicidality, and clinical management of suicidal patients, led by Dr. Larry Alphs, and is copyright (2016) to Janssen Scientific Affairs, LLC. The distributor of the SIBAT is Mapi Research Trust. ECC is the author of the mobile application version of the SIBAT used in this study. RJL was on an advisory board for Pfizer, though no renumeration was provided, cost of travel and stay was covered by Pfizer. KJA declares consultancy services for Otsuka Canada Pharmaceutical Inc., Lundbeck Canada, and HLS Therapeutics and research grants from Janssen Inc., Canada (the above fellowship grant and a further fellowship grant).

### *Data Access*

Data is available from the corresponding author upon request.

### *Acknowledgment*

We thank David Streiner for statistical advice and other editorial input.

### *Copyright and Access to the SIBAT*

SIBAT is protected by international copyright, with all rights reserved to Janssen Scientific Affairs, LLC. Do not use without permission. For information on, or permission to use SIBAT, please contact Mapi Research Trust, official distributor of the SIBAT on behalf of Janssen Scientific Affairs, LLC on ePROVIDE™ at https://eprovide.mapi-trust.org/instruments/suicide-ideation-and-behavior-assessment-tool.

### *Funding*

KJA was awarded a fellowship grant from Janssen Inc., Canada to support activities of a fellow (trainee). ECC was the named fellow on a grant from Janssen Inc., Canada, with the funds being used to contribute to support for research assistants involved in participant recruitment and database maintenance. ECC also received two Rudelle Hall Graduate Scholarships (2018 and 2019/20) and awards from the Department of Psychiatry Residency Project Fund at the University of Alberta. A grant from the American Foundation for Addiction Research to KJA and the Alberta Centennial Addiction and Mental Health Research Chair and transitional funding to KJA provided support for participant recruitment, data collection, and database maintenance. PJC received a Fulbright Canada – Palix Foundation Distinguished Visiting Research Chair in Brain Science, and Child and Family Health and Wellness (host institution: the University of Alberta). The funding sources had no involvement in any aspect of the study design, data collection, data analysis, or interpretation.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.psychres.2021.114128.

## Appendix 1. Sample Question Stems from SIBAT

- The number of members in my family who have died by suicide is…
- Over the past 7 days I have felt hopeful.
- Over the past 7 days my sleep has been good.
- I have drunk alcohol on __ of the past 7 days.
- I am glad to be alive.
- Nothing in life gives me pleasure.
- Nobody will care if I am dead.
- My emotional (mental) pain is so severe that I want to end my life.
- I have been shamed and should die.
- Which of the following ratings best describes your intent to end your life in the past 7 days?

## References

Alphs, L., Canuso, C., Williamson, D.E., et al., 2015. Suicide ideation and behavior assessment tool (SIBAT): a novel measure of suicidal ideation and behavior and perceived suicide risk. In: Poster presented at: International Society for CNS Clinical Trials and Methodology 2015 Autumn Conference. Amsterdam, The Netherlands.

Alphs, L., Fu, D.J., Williamson, D.E., et al., 2018a. Empirical evaluation of suicidal ideation and behavior categories established by the Columbia Classification Algorithm of Suicide Assessment. C-CASA. In: Poster presented at: 14th Annual Scientific Meeting of the International Society for CNS Clinical Trials and Methodology (ISCTM). Washington DC.

Alphs, L., Fu, D.J., Williamson, D., et al., 2020. Suicide ideation and behavior assessment tool (sibat): evaluation of intra- and inter-rater reliability, validity, and mapping to columbia classification algorithm of suicide assessment. Psychiatry Res. 294, 113495 https://doi.org/10.1016/j.psychres.2020.113495.

Alphs, L., Williamson, D., Turkoz, I., et al., 2018b. Validation of suicidal ideation and behavior assessment tool (SIBAT): intra- and inter-rater reliability. In: Poster presented at: 17th European Symposium on Suicide and Suicidal Behaviour (ESSSB). Ghent, Belgium.

Alphs, L., Williamson, D., Turkoz, I., et al., 2018c. Validation of Suicidal Ideation and Behavior Assessment Tool (SIBAT): Intra- and inter-Rater reliability. Poster presented at: American College of Neuropharmacology. Hollywood, Florida, USA.

American Psychiatric Association, 2016. Practice Guidelines For the Psychiatric Evaluation of Adults third editionArlington (VA)American Psychiatric AssociationISBN:978-0-89042-465-0.

Boudreaux, E.D., Horowitz, L.M., 2014. Suicide risk screening and assessment: designing instruments with dissemination in mind. Am. J. Prev. Med. 47 (3), S163–S169. Supplement 2https://doi.org/10.1016/j.amepre.2014.06.005.

Carter, G., Milner, A., McGill, K., et al., 2017. Predicting suicidal behaviours using clinical instruments: systematic review and meta-analysis of positive predictive values for risk scales. Br. J. Psychiatry 210 (6), 387–395 https://doi.org/10.1192/bjp.bp.116.182717.

Chan, M.K., Bhatti, H., Meader, N., et al., 2016. Predicting suicide following self-harm: systematic review of risk factors and risk scales. Br. J. Psychiatry 209 (4), 277–283 https://doi.org/10.1192/bjp.bp.115.170050.

Fu, D.J., Ionescu, D.F., Li, X., et al., 2020. Esketamine nasal spray for rapid reduction of major depression symptoms in patients who have active suicidal ideation with intent: double-blind, randomized study (ASPIRE I). J. Clin. Psychiatry 81 (3), 19m13191https://doi.org/10.4088/JCP.19m13191.

Giddens, J.M., Sheehan, D.V., 2014. Is there value in asking the question "do you think you would be better off dead?" In assessing suicidality? A case study. Innov. Clin. Neurosci. 11 (9–10), 182–190.

Hawgood, J., De, Leo D., 2016. Suicide prediction – a shift in paradigm is needed. Crisis 37 (4), 251–255 https://doi.org/10.1027/0227-5910/a000440.

Ionescu, D.F., Fu, D.J., Qiu, X., et al., 2021. Esketamine nasal spray for rapid reduction of depressive symptoms in patients with major depressive disorder who have active suicide ideation with intent: results of a phase 3, double-blind, randomized study (ASPIRE II). Int. J. Neuropsychopharmacol. 24 (1), 22–31 https://doi.org/10.1093/ijnp/pyaa068.

Kessler, R.C., Warner, C.H., Ivany, C., et al., 2015. Predicting suicides after psychiatric hospitalization in US army soldiers: the army study to assess risk and resilience in servicemembers (Army STARRS). JAMA Psychiatry 72 (1), 49–57 https://doi.org/10.1001/jamapsychiatry.2014.1754.

Kim, H., Kim, B., Kim, S.H., et al., 2018. Classification of attempted suicide by cluster analysis: a study of 888 suicide attempters presenting to the emergency department. J. Affect. Disord. 235, 184–190 https://doi.org/10.1016/j.jad.2018.04.001.

Kleiman, E.M., Turner, B.J., Fedor, S., et al., 2017. Examination of real-time fluctuations in suicidal ideation and its risk factors: results from two ecological momentary

assessment studies. J. Abnorm. Psychol. 126 (6), 726–738 https://doi.org/10.1037/abn0000273.
Mental Health Commission of Canada, 2018. Research On Suicide and Its prevention: What the Current Evidence Reveals and Topics For Future Research. Ottawa, ON. ISBN:978-1-77318-089-2.
Morales S., Barros J., Echávarri O. et al. 2017. Acute mental discomfort associated with suicide behavior in a clinical sample of patients with affective disorders: ascertaining critical variables using artificial intelligence tools. Front Psychiatry. 8:7. https://doi.org/10.3389/fpsyt.2017.00007.
Morrison, K.B., Laing, L., 2011. Adults' use of health services in the year before death by suicide in Alberta. Health Rep. 22 (3), 15–22.
National Action Alliance for Suicide Prevention: Transforming Health Systems Initiative Work Group, 2018. Recommended Standard Care For People With Suicide risk: Making health Care Suicide Safe. Education Development Center, Inc. Washington, DC.
Oh, J., Yun, K., Hwang, J.H., et al., 2017. Classification of suicide attempts through a machine learning algorithm based on multiple systemic psychiatric scales. Front Psychiatry 8, 192 https://doi.org/10.3389/fpsyt.2017.00192.
Oquendo, M.A., Baca-Garcia, E., Artés-Rodríguez, A., et al., 2012. Machine learning and data mining: strategies for hypothesis generation. Mol. Psychiatry 17 (10), 956–959 https://doi.org/10.1038/mp.2011.173.
Pokorny, A.D., 1983. Prediction of suicide in psychiatric patients. Report of a prospective study Arch. Gen. Psychiatry 40 (3), 249–257. https://doi.org/10.1001/archpsyc.1983.01790030019002.
Public Health Agency of Canada, 2019. Suicide in Canada: Key statistics (infographic) [accessedSeptember 19]. https://www.canada.ca/en/public-health/services/publications/healthy-living/suicide-canada-key-statistics-infographic.html.
Roaldset, J.O., Linaker, O.M., Bjørkly, S., 2012. 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 https://doi.org/10.1080/13811118.2013.722052.
Rosen, A., 1954. Detection of suicidal patients: an example of some limitations in the prediction of infrequent events. J. Consul. Psychol. 18 (6), 397–403 https://doi.org/10.1037/h0058579.
Runeson, B., Odeberg, J., Pettersson, A., et al., 2017. Instruments for the assessment of suicide risk: a systematic review evaluating the certainty of the evidence. PLoS ONE 12 (7), e0180292 https://doi.org/10.1371/journal.pone.0180292.
Sanislow, C.A., 2016. Updating the research domain criteria. World Psychiatry 15 (3), 222–223 https://doi.org/10.1002/wps.20374.
Sinyor, M., Schaffer, A., Streiner, D.L., 2014. Characterizing suicide in Toronto: an observational study and cluster analysis. Can. J. Psychiatry 59 (1), 26–33 https://doi.org/10.1177/070674371405900106.
Statistics Canada, 2021a. Table 13-10-0392-01 Deaths and Age-Specific Mortality rates, By Selected Grouped Causes https://doi.org/10.25318/1310039201-eng (accessed 12 March 2021).
Statistics Canada, 2021b. Table 13-10-0394-01 Leading causes of death, Total population, By Age Group https://doi.org/10.25318/1310039401-eng (accessed 12 March 2021).
Steeg, S., Quinlivan, L., Nowland, R., et al., 2018. Accuracy of risk scales for predicting repeat self-harm and suicide: a multicentre, population-level cohort study using routine clinical data. BMC Psychiatry 18 (1), 113 https://doi.org/10.1186/s12888-018-1693-z.
Streiner, D.L., 2002. Breaking up is hard to do: the heartbreak of dichotomizing continuous data. Can. J. Psychiatry 47 (3), 262–266 https://doi.org/10.1177/070674370204700307.
Streiner, D.L., 2003. Starting at the beginning: an introduction to coefficient alpha and internal consistency. J. Pers. Assess. 80 (1), 99–103 https://doi.org/10.1207/S15327752JPA8001_18.
Torous, J., Larsen, M.E., Depp, C., et al., 2018. Smartphones, sensors, and machine learning to advance real-time prediction and interventions for suicide prevention: a review of current progress and next steps. Curr. Psychiatry Rep. 20 (7), 51 https://doi.org/10.1007/s11920-018-0914-y.
Turkoz, I., Fu, D.J., Canuso, C., et al., 2018. Suicidal ideation and behavior assessment tool (SIBAT): psychometric characterization of patient-reported modules. Poster presented at: the International Society for. Clinical Trials Methodol. 2018 Autumn Conference. California, USA.
Von Hippel, P.T., 2009. How to impute interactions, squares, and other transformed variables. Sociol. Methodol. 39 (1), 265–291 https://doi.org/10.1111/j.1467-9531.2009.01215.x.
Walsh, C.G., Ribeiro, J.D., Franklin, J.C., 2017. Predicting risk of suicide attempts over time through machine learning. Clin. Psychol. Sci. 5 (3), 457–469 https://doi.org/10.1177/2167702617691560.
Williamson, D., Canuso, C., Fu, D.J., et al., 2017. Patient report with the suicide ideation and behavior assessment tool (SIBAT): acceptability and sensitivity to rapid change. Biol. Psychiatry 81 (10), S349.