Author's response to reviews

Title: Combining the Suicide Intent Scale and the Karolinska Interpersonal Violence Scale in Suicide Risk Assessment

Authors:

Jon Stefansson (jon.stefansson@ki.se)
Peter Nordström (peter.nordstrom@ki.se)
Bo Runeson (bo.runeson@ki.se)
Marie Åsberg (Marie.Asberg@ki.se)
Jussi Jokinen (jussi.jokinen@ki.se)

Version: 3
Date: 23 April 2015

Author's response to reviews: see over
The authors would like to thank the Reviewers for a careful review of our manuscript “Combining the Suicide Intent Scale and the Karolinska Interpersonal Violence Scale in Suicide Risk Assessment” and providing us with their valuable comments and suggestions to improve the quality of the manuscript. We have performed a major and minor revisions accordingly.

The following responses have been prepared to address all Reviewers’ comments in a point-by-point fashion, with our responses below (in BOLD type). The page (P) numbers refer to our revised manuscript submitted April 23, 2015.

Reviewer #1. Emilie Olié

The authors propose to measure the predictability of a combination of two scales to identify people at risk of suicide. They consider high-risk patients including only previous suicide attempters. This topic is of great interest. Indeed, many risk factors of suicide are well known, but with very low predictability by themselves. Thus salient measures are still needed to better identify patients that will die by suicide.

1. Authors should add information about previous works using a multidimensional approach of suicidal risk including suicide attempt and suicide (for example, Blasco Fontecilla et al. Journal of Psychiatry Research 2012).

Response 1.1: Thank you for the suggestion. We have added the reference as requested. P8.

2. The authors could better justify their choice of combination of these two scales only.

Response 1.2: We have added in the rationale for choice of these two scales only. The main argument is that we chose scales with documented predictive ability for completed suicide. P4.

3. The assessment “Both childhood trauma, adult violent behavior and suicidal behaviour are interlinked. » needs to be justified.

Response 1.3: Thank you for pointing this out. We have added references concerning the link between childhood trauma, adult violent behaviour and suicidal behaviour. P3.

4. How do the authors explain the small number of included suicide attempters within 5 years? Could the authors give the number of eligible patients during this period?

Response 1.4: We do not unfortunately have the exact number of eligible patients during this period (1993-1998). However in our later study (2001-2005) from the same catchment area, we had information of 258 eligible suicide attempters of which 100 finally participated. With light of this and the fact that the present study also included a CSF protocol (about half of the patients), we consider that the number of included suicide attempters corresponds quite well to the number of consenting participants in the later study with the same inclusion time from the same
catchment area. Since we do not have this information for this cohort, we have added this fact in the limitations. P9.

5. Suicide intent was assessed for the last suicide attempt. But what about previous behaviors? Would it be interesting to consider also SIS scores of first suicide attempt and/or the most lethal attempt?

Response 1.5. Thank you for pointing this out. It would indeed be interesting to consider SIS scores of the first attempt/the most lethal attempt. Unfortunately, we do not have that information systematically assessed in the study protocol. Please see also response to your question 7.

6. Is KIVS a validated scale? If yes, in which population(s)?

Response 1.6. KIVS is a validated scale (Jokinen et al., 2010). The Buss-Durkee Hostility Inventory (BDHI), Urge to act out hostility subscale from the Hostility and Direction of Hostility Questionnaire (HDHQ) and the Early Experience Questionnaire (EEQ) were used for validation in suicide attempters. We have added short information of psychometric properties in the manuscript. P5.

7. First suicide attempters and major repeaters tend to be two distinct populations. How have the authors taken into account this issue?

Response 1.7. Thank you for this important question from a suicidological point of view. We have now assessed this by using item 18 of SIS which classifies previous attempt(s) into three categories: 1) none, 2) one or two and c) three or more. There was no association between repeaters (or major repeaters) vs first suicide attempters concerning suicide risk. However, repeaters scored higher in KIVS total score and 3 of 4 KIVS subscales, but not in SIS. We have added results of these analyses in the results section. When using repeater status as a covariate (confounder) together with SIS and KIVS scores in logistic regression for suicide completion, the results remained practically the same. P7, 8.

8. The assessment needs to be reconsidered. « In this study, neither high scores in SIS or KIVS predicted suicide on the short term. However, their combined use detected the suicide completers in the high risk category from a long term perspective. »

Response 1.8. We have changed the wording. “In this study, the mean time to suicide after attempted suicide was six years. High scores in SIS and KIVS independently predicted suicide indicating that their combined use may detect suicide attempters in the high risk category from a long term perspective.” P7.

9. The authors should emphasize the interest of using two subscales of the KIVS rather than the global scale.

Response 1.9. Thank you pointing this out. We have emphasized the two subscales of the KIVS. P7.

10 The paragraph “Both scales used in this clinical study reflect underlying neurobiological vulnerability to suicidal behavior. Exposure to interpersonal violence as a child as well as aggression
Response 1.10. We have rewritten this part to be more to the point. “Both scales used in this clinical study have been studied in relation to underlying neurobiological vulnerability to suicidal behavior. Both high exposure to interpersonal violence as a child as well as accentuated aggression dyscontrol measured with the Karolinska Interpersonal Violence Scale were associated with low levels of serotonin metabolite 5-hydroxyindolacetic acid in the cerebrospinal fluid (CSF) in suicide attempters [23] whereas high suicide intent measured with SIS was associated with low CSF oxytocin [24]. Furthermore, KIVS measures both distal risk in form of early life adversity and developmental factor impulsive aggression, while SIS may also capture more precipitating aspects of suicidal crisis.” P9

Response 1.11. Changes made as requested, we have moved the figure legend before figure 1 for clarity.

Response 1.12. Thank you for your comment. We believe that it is important to show that SIS and KIVS measure different forms of risk and displaying the correlations (both total scores and subscales) in a table makes that clear for the reader.

Reviewer #2 Martino Belvederi Murri:

The research question is relevant to the clinical field and well defined. Indeed, several studies attempted at identifying predictive factors for suicide, with relatively low success. The analyses are appropriately conducted and the writing style is concise and clear.

1. The author mention that the SIS alone reached a PPV of 16.7% and adding the KIVS it raises to 18.8%. I wonder whether a 2.1% raise in the positive predictive value really corresponds to a meaningful improvement in clinical practice?

Response 2.1. Thank you for your comment on clinical relevance. We do agree that the change in PPV is quite small and the change accounted for only higher sensitivity reached with combined use of the two scales, we have now discussed PPV change in relation to meaningful improvement in
Response to Reviewers’ Comments

2015-04-22

"Combined use of both scales led to a higher specificity. Using both scales, the number of false positives was reduced by 9 patients compared to using SIS alone (26 vs 35 false positives) and by 17 patients compared to using KIVS alone (26 vs 43 false positives) leading to a specificity of 63%. In other words, 11-21% of suicide attempters in this cohort, classified as false positive high suicide risk patients using one scale only, could be reclassified as patients with lower suicide risk combining the both scales in the suicide risk assessment." P7.

2. Also in table 3, we see that only 77 patients had both SIS and KIVS ratings. With only 7 suicides in the whole sample, I fear that the change of predictivity might also be influenced by the exclusion of some subjects.

Response 2.2. As pointed out, we had unfortunately some missing data concerning the SIS and KIVS ratings. This is quite common in clinical studies on severely ill patients. We handled this by not including these individuals in the analyses. The possibility of the change of predictivity being influenced by missing data is mentioned as a limitation. P9.

3. Second, in the previous study the authors found that “Four items were used to test a shorter version of the SIS in the suicide prediction. The positive predictive value was 19% and the AUC was 0.82” then the authors should give better argumentations of how this is an improvement. By clinical experience, one can deem that exposure to violence increases the likelihood of self-harm, but, put this way, these data seem to show overall little change compared to the use of only the SIS.

Response 2.3. The shorter version of SIS in our previous article gave indeed somewhat better prediction. However, since this finding has not been replicated in an independent sample, or the shorter version been validated, we decided not to use the short version in the analyses in this study. When using the shorter version of SIS together with KIVS total score would give AUC 0.87 and ppv 21.4%. Since the improvement of ppv is approximately 2% in both analyses (SIS total/SIS short version with KIVS total), we report the results using the SIS total score in the manuscript.

4. Third: instead, the use of SIS+KIVS “adult subscale” seemed to reach higher values (26.3%), while maintaining acceptable levels of specificity and sensitivity. Possibly, this should be evident in the abstract and in the discussion, rather than the other finding?

Response 2.4. Thank you for pointing this out. This finding is now added in the abstract and discussed more in detail. P8.

5. However I would consider that, As in every study attempting at predicting suicide over the long-term, prediction is made difficult by “clinical interference” (i.e. the case management after the assessment). One cannot ignore that patients assessed as displaying more severe suicidal intent, will likely to be treated more intensively. Possibly (and luckily) this limits the maximum PPV that any study can reach.

Response 2.5. Thank you for pointing this out. This important aspect is added in the discussion as a limitation since we do not have information of treatments received. P9.

6. - I think the authors should clarify the relationship of this study with the other recent study on 181 suicide attempters recruited between 1993 and 2005, seemingly from the same setting [M. Rajalin et al./Journal of Affective Disorders 148 (2013) 92–97]. Thus, a hundred patients (cohort 2) are
excluded from the present study, possibly because “In the first part of the inclusion period the information was collected from patient records”, instead of the KIVS. Or maybe because of a longer follow up available for cohort 1. In any case, it would have been interesting to compare if the same information, drawn through another instrument has similar predictive value. I think these explanations should be given and the study should be cited in the paper.

Response 2.6. Thank you for pointing this out. Unfortunately, the Beck Suicide Intent Scale (SIS) was not a part of the study protocol in the latter study, while KIVS was part of the study protocol in both studies. Thus, we were not able to include patients from the cohort 2 in the current manuscript due to lack of SIS ratings. (In the study of Rajalin et al, information of heredity for suicide was collected from the patient records for this cohort). This is now clarified in the methods and the study Rajalin et al., 2013 is cited. P4.

7. - the sample size is relatively small: only 7 completed suicides. The authors point this out, but without sufficient emphasis in my opinion. Conclusions should be taken with caution before replication in larger samples.

Response 2.7. We have modified the conclusions according this comment. P9.

Minor comments

8. How many patients received a borderline PD diagnosis? Were patients displaying Non-Suicidal Self-Injury included? NSSI might represent an important factor to take in account (as it does in clinical practice)

Response 2.8. We have added information concerning BPD diagnosis in the sample, 8 patients fulfilled criteria for BPD. P5. NSSI is prevalent in BPD, however NSSI frequency, methods etc were not systematically assessed in the study protocol. We do agree that information of NSSI might be important, however studies are lacking concerning this behaviour as a risk factor for completed suicide. Please see also response 7 reviewer 1.

9. Were all eligible patients approached to take part in the study? Is data available on how many patients were not included in the study because of refusal? Representativity of the sample is an important issue that might be further highlighted.

Response 2.9. We do not unfortunately have the exact number of eligible patients during this period (1993-1998). However in our later study (2001-2005) from the same catchment area, we had information of 258 suicide attempters of which 100 finally participated. With light of this and the fact that the present study also included a CSF protocol (about half of the patients), we consider that the number of included suicide attempters corresponds quite well to the number of consenting participants in the later study with the same inclusion time from the same catchment area. Since we do not have this information for this cohort, we have added this fact in the limitations. P9.
10. What were the causes of death for the seven non-suicides? Was there any margin for ambiguity in the classification?

Response: 2.10. This is an important aspect when doing mortality studies. All causes of death were scrutinized from the death certificates. This is clarified in the methods. Suicides were classified as suicide deaths if it was given as a cause of death in the death certificate. The natural causes of death were sarcoma, cerebral bleeding, stroke, pneumonia and atherosclerosis. Two patients died of unnatural causes (classified as accidents in the death certificates). P5.

11. There might be some misuse of English terms (e.g. line 101 “evinced”)

Response: 2.11. We have checked the language.

12. Please report AUC p values and confidence intervals from ROC curve analyses


13. The authors might want to consider reporting a very brief description of the types of suicide attempts made in the cohort at the time of the evaluation.

Response: 2.12. We have added this information as requested.P5.

Reviewer #3. Nuria Machin

1. Is the question posed by the authors well defined? To a limited extent. There is a clearly described hypothesis (the potential benefit of combining two scales) but the reasons for choosing these two scales is only briefly indicated.

Response 3:1. Thank you for pointing this out. We have added rationale for choosing these two scales. The main argument is that we chose scales with documented predictive ability for completed suicide. P4.

2. They do not comment on the fact that the study was entirely in Sweden. It may be that there are no reasons why exactly the same results would be found in any country but there could be differences e.g. the reality of which scales are preferred in individual countries. Some comments on this would be appropriate. Another matter which is not addressed is: to what extent do the results from this study or its follow-up feed into policy and / or clinical practice? The authors comment that a larger clinical trial is now required but what would they anticipate at the end of that?

Response 3:2. Thank you for pointing this out. This is an interesting aspect. Suicide Intent Scale is widely used and widespread in different countries and translations. The Karolinska Interpersonal Violence Scale was published in 2010 and is quite a new instrument, translated to English. New data on KIVS will be published very soon in different patient cohorts. We have added in the discussion the fact that the study was performed in Sweden and discussed eventual differences...
which scales are preferred in individual countries. We have also added a comment to what extent the results from this study or its follow-up feed into policy and/or clinical practice. P9,10.

Reviewer #4. Bogdan Nemeş

1. The authors do not clearly state how they combined the results of the two instruments. I presume that they defined a positive result of the combined test (equivalent to high suicide risk) if both instruments had scores over the established thresholds, but one may think about combining the two scores in a mathematical function and study the relevance of this new combined score for predicting suicide attempts. Therefore, I consider that the authors need to explicitly state this. Major compulsory revision 1.

Response 4:1. Thank you pointing this out. We defined a positive result of the combined test (equivalent to high suicide risk) if both instruments had scores over the established thresholds from the ROC curves. Since there was no significant correlation between the scales, we used them in a regression model as independent predictors of completed suicide. This is now clarified in the manuscript. The idea of combining the two scores in a mathematical function and study the relevance of this new combined score for predicting suicide is interesting and could possibly lead to a new scale. However, our aim was to test the predictive validity of the two existing scales. P6.

2. Conclusions should be formulated based on the aims declared. This paper proves that using both SIS and KIVS, combined, is better for predicting suicide attempts than using them separately, and that the two instruments assess different aspects of suicide risk. It is this that should be stated in the conclusions. Furthermore, the need for replication should be stated, in my opinion, in the discussion section, since this paper is not making a review of the available data on suicide assessment. Major compulsory revision 2

Response 4:2. Changes made as requested. P9,10.

Positive predictive values and areas under the curve are given in the results section of the abstract, and specificity in the conclusion section. I suggest reconciling the two sections. Minor essential revision 1. ! Higher specificity of the combined use does not necessarily suggest that the two instruments measure different components of suicide risk, as conveyed by how the conclusions are written in the abstract. The main argument supporting the hypothesis that they are indeed measuring different components of suicide risk is the fact that they are poorly correlated. Major compulsory revision 3.

Response 4:3. We have reconciled the two sections as proposed, please see the abstract. We have changed the conclusions according to the reviewer’s comment on the fact that the scales were poorly correlated and thus measure different components of suicide risk.

Minor revisions:
On row 97-98, the authors state that: “The assessment of suicide risk is the most important and difficult part in the management of patients after a suicide attempt.” This is debatable; consider rewriting. Discretionary revision 1.

**Response:** Wording changed P3.

On row 109-110 and 111-112 the authors cite NICE Guidelines using another format for citations – minor essential revision 2.

**Response:** Changes made as requested.

On row 179-180 the authors state: “…one patient had substance related disorder…” and on row 182 they state “...21% of the suicide attempters had a substance related disorder.” Please clarify – minor essential revision 3.

**Response:** One patient had substance related disorder as the primary Axis 1 diagnosis, while 21% had it as a comorbid diagnosis (Axis 1 secondary, third or fourth diagnosis).