Esta es la versión de autor del artículo publicado en:
This is an author produced version of a paper published in:


DOI: http://dx.doi.org/10.1016/j.jpsychires.2012.06.013

Copyright: © 2012 Elsevier Ltd. All rights reserved

El acceso a la versión del editor puede requerir la suscripción del recurso
Access to the published version may require subscription
Dear Author,

Please check your proof carefully and mark all corrections at the appropriate place in the proof (e.g., by using on-screen annotation in the PDF file) or compile them in a separate list. Note: if you opt to annotate the file with software other than Adobe Reader then please also highlight the appropriate place in the PDF file. To ensure fast publication of your paper please return your corrections within 48 hours.

For correction or revision of any artwork, please consult http://www.elsevier.com/artworkinstructions.

Any queries or remarks that have arisen during the processing of your manuscript are listed below and highlighted by flags in the proof.

<table>
<thead>
<tr>
<th>Location in article</th>
<th>Query / Remark: Click on the Q link to find the query’s location in text Please insert your reply or correction at the corresponding line in the proof</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>The citations ‘Patton JH, 1995; Efron B., 2004’ have been changed to match the author name in the reference list. Please check and correct if necessary.</td>
</tr>
<tr>
<td>Q2</td>
<td>Please note that the funding details provided under the ‘Role of funding source’ and ‘Acknowledgments’ were similar. Hence the section ‘Role of funding source’ is not retained. Kindly check and amend if necessary.</td>
</tr>
<tr>
<td>Q3</td>
<td>Please check the expanded journal titles in references.</td>
</tr>
<tr>
<td>Q4</td>
<td>Please provide an update for reference “Delgado-Gomez et al., in press”.</td>
</tr>
<tr>
<td>Q5</td>
<td>Please check the layout of Table 3, and correct if necessary.</td>
</tr>
<tr>
<td>Q6</td>
<td>Please confirm that given names and surnames have been identified correctly.</td>
</tr>
</tbody>
</table>

Please check this box if you have no corrections to make to the PDF file

Thank you for your assistance.
Combining scales to assess suicide risk

Hilario Blasco-Fontecilla a,*, David Delgado-Gomez b, Diego Ruiz-Hernandez c, David Aguado d, Enrique Baca-Garcia a, e, Jorge Lopez-Castroman f

a Department of Psychiatry, Puerta de Hierro Hospital, CIBERSAM, Calle de Los Madroños 5, 28400 Collado Villalba, Spain
b Department of Statistics, Carlos III University, Calle Madrid 126, Getafe 28903, Spain
c Department of Psychiatry, Puerta de Hierro Hospital, CIBERSAM, Calle de Los Madroños 5, 28400 Collado Villalba, Spain
d Department of Psychiatry, Puerta de Hierro Hospital, CIBERSAM, Calle de Los Madroños 5, 28400 Collado Villalba, Spain
e Department of Psychiatry, Puerta de Hierro Hospital, CIBERSAM, Calle de Los Madroños 5, 28400 Collado Villalba, Spain
f Institute of Knowledge Engineering, Autonoma University, 28049 Madrid, Spain

A R T I C L E   I N F O

Article history:
Received 20 November 2011
Received in revised form 22 May 2012
Accepted 19 June 2012

Keywords:
Lars-en algorithm
Suicidal behavior
Personality traits
Life events

A B S T R A C T

Objectives: A major interest in the assessment of suicide risk is to develop an accurate instrument, which could be easily adopted by clinicians. This article aims at identifying the most discriminative items from a collection of scales usually employed in the assessment of suicidal behavior.

Methods: The answers to the Barrat Impulsiveness Scale, International Personality Disorder Evaluation Screening Questionnaire, Brown–Goodwin Lifetime History of Aggression, and Holmes & Rahe Social Readjustment Rating Scale provided by a group of 687 subjects (249 suicide attempters, 81 non-suicidal psychiatric inpatients, and 357 healthy controls) were used by the Lars-en algorithm to select the most discriminative items.

Results: We achieved an average accuracy of 86.4%, a specificity of 89.6%, and a sensitivity of 80.8% in classifying suicide attempters using 27 out of the 154 items from the original scales.

Conclusions: The 27 items reported here should be considered a preliminary step in the development of a new scale evaluating suicidal risk in settings where time is scarce.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Suicide is a major health issue. One suicide is completed every 40 s, leading to approximately one million deaths every year worldwide (WHO, 2002). Moreover, suicide is the third most important cause of death worldwide among people aged 15–44 (Holmes et al., 2007). Notwithstanding human costs, the economic burden of suicidal behavior has been estimated annually in $33 billion in the United States (Coreil et al., 2001). Fortunately, suicidal behavior might be prevented to a great extent (Jamison, 2000). Treating subjects at risk with the appropriate preventive measures, such as cognitive behavior therapies (Brown et al., 2005) can reduce suicide rates up to 25% (Isaacson, 2000). More recently, a 75% reduction of suicide rates has been reported in a large depression care program (Hampton, 2010). In order to detect subjects at risk, researchers have investigated the factors underlying suicidal behavior. The most relevant risk factors are major depression (Mann et al., 1999b), high impulsiveness (Patton et al., 1995), aggressiveness (Mann et al., 1999b), personality disorders (Mann et al., 1999a), life events (Kolves et al., 2006), and social-demographic factors (Smith et al., 1988).

Unfortunately, most of these studies did not measure the effectiveness of the risk factors to identify subjects at risk. They just tested if there was a statistically significant relationship between the studied variable (e.g. high impulsiveness) and suicidal behavior. Therefore, the clinical usefulness of these studies is limited. One notable exception is the seminal Pokorny’s article (Pokorny, 1983). Pokorny applied discriminant analysis to several features including, among others, socio-demographic variables, the 24 items of the brief psychiatric rating scale, and the items included in the nurses’ observation scale for inpatient evaluation. Although it was an innovative approach, his results showed a weak performance, with accuracy, sensitivity and specificity levels below 70%. More recently, Hendin (Hendin et al., 2010) slightly improved these results achieving an accuracy of 71.67% with a specificity of 74% and a sensitivity of 60%. The improvement was basically due to the use of a different set of predictive variables, as they used a simple classifier consisting on the sum of the individual scores associated...
to each variable. This study suggested that the use of a more suitable set of predictive variables together with the use of more sophisticated classifiers might improve the classification accuracy of people at risk of suicidal behavior. This intuition was confirmed by us (Delgado-Gomez et al., 2011) in a study aimed at discriminating between suicide attempters (SA) and non-SA. In this study, we used two personality scales as predictors, and a collection of modern classification techniques such as linear discriminant analysis, Fisher linear discriminant analysis, boosting, and support vector machines (SVM). The best results were obtained with SVM, which achieved a classification accuracy of 80.3%, with a specificity of 86.8% and a sensitivity of 76.1%. Recently, Stefansson et al. (Stefansson et al., 2010) have shown that the prediction of suicide can be improved by means of an appropriate selection of the items. However, their results were obtained ad hoc. Therefore, we have applied the Lars-en algorithm in order to automatically select the most discriminative items (Delgado-Gomez et al., in press). Using the selected items of scales measuring life events and personality disorders, it was possible to achieve a classification accuracy of 83%. A question that remains open is the accuracy that could be obtained if the Lars-en algorithm were applied to a set of items assessing the most relevant risk factors for suicidal behavior.

The present study extends our previous findings and is conceived as a further step toward the development of more precise and reliable measures of suicide risk (Garcia-Nieto et al., 2012). The major goal is to maximize the classification accuracy applying simultaneously Lars-en to sociodemographic factors and items from four scales measuring some of the most relevant risk factors for suicidal behavior (impulsive aggression, life events and personality disorders). As a by-product, we developed a small set of items to classify subjects as SA or non-SA. This set of items might help to develop a tool to support clinical decisions with regard to suicide risk.

2. Methods

2.1. Participants

To accomplish our objectives, data from 687 subjects were used. Participants were 18 years or older and provided written informed consent before participating in the study. Subjects that showed incapability to provide informed consent were excluded (e.g. presence of a life-threatening medical condition, significant organic brain disease). The cases included 249 first-time SA (157 women and 92 men) admitted to two university hospitals in Madrid, Spain, between 1999 and 2003. Non-SA (n = 438) included 81 psychiatric inpatients (54 women and 27 men) without current or past history of suicidal behavior and 357 healthy controls (blood donors; 131 women and 226 men) recruited in the same hospitals. Healthy controls had neither Axis I diagnoses nor a history of suicidal behavior. The appropriate ethics committee approved the study. The study was carried out in accordance with the latest version of the Declaration of Helsinki.

Mean age (±SD) of the SA, psychiatric controls, and healthy controls were 37.2 (±14.5), 42.1 (±13.0), and 34.6 (±10.7), respectively (F = 13.18; df = 2; p < 0.001). We used the Mini International Neuropsychiatric Interview (MINI), a short, easy to administer, and efficient structured diagnostic interview to assess Axis I disorders in psychiatric inpatients and SA (Sheehan et al., 1998). Tables 1 and 2 show information with regard to sociodemographic and clinical factors of the study groups.

We also used the Risk-Rescue Rating Scale (RRRS), a 10-item interviewer-administered questionnaire that provides an estimate of the seriousness of a SA (Weisman and Worden, 1972). The first five items of the scale describe risk factors of a SA. The Risk Rating ranges from 5 (5–6 points indicate “low risk”) to 15 (13–15 points indicate “high risk”). 81.3% of the SA were classified as low risk SA. Furthermore, item 15 of the Beck Suicidal Intent Scale (SIS) was implemented to elucidate the degree of premeditation (Beck et al., 1974). In 63% of SA there was no premeditation at all (impulsive SA) and only 23.7% thought on suicide more than three hours before the attempt.
2.2. Measures

We selected age and sex as socio-demographic features and four clinical scales. Both age and sex are well-known risk factors of suicidal behavior. Suicide attempts are 2–3-fold more frequent among women than among men (Canetto and Sakinofsky, 1998). In addition, age is reported as a risk factor for future attempts in most studies (Christiansen and Jensen, 2007), and an age pattern has been described for the first suicide attempt (Slama et al., 2009). The selected scales assess central features of suicidal behavior: personality, life events and the impulsive aggression construct. Personality and life events are nuclear factors to explain suicidal behavior in the context of the stress-diathesis model (Mann et al., 2005, 2006; Perroud et al., 2011).

- The 11th version of the Barrat Impulsiveness Scale (BIS-11) is a widely used measure of impulsiveness. In the present study we used the Spanish version (Oquendo et al., 2001). The BIS-11 is a 30-item self-reported Likert scale that comprises three subscales to assess cognitive, motor, and non-planned impulsiveness. Items are scored from 1 (rarely/never) to 4 (almost always/always).

- The International Personality Disorder Evaluation Screening Questionnaire (IPDE-SQ). The IPDE-SQ (Loranger et al., 1994) is a screening psychological instrument designed to identify relevant traits and behaviors in the assessment of personality disorders according to the main international classifications of mental disorders. This questionnaire examines the presence in adults (if apparent for at least five years) of diagnostic criteria for any personality disorder and comprises 77 True/False self-report items.

- The Holmes and Rahe Social Readjustment Rating Scale (SRRS). Life events within two years preceding a suicide attempt were coded according to the standardized and adapted Spanish version (Gonzalez de Rivera and Morera, 1983) of the SRRS (Holmes and Rahe, 1967). The SRRS includes 43 life events ranked according to the degree of severity. Death of spouse is considered the most severe item while minor violations of the law is the mildest.

- The Brown-Goodwin Lifetime History of Aggression (BGHA). The BGHA scale is a 11-item questionnaire measuring how many times different types of aggressive behavior occurred across childhood, adolescence, and adulthood (Brown et al., 1979). It includes 11 questions over a large range of aggressive behaviors. Different scores are registered for childhood, adolescence and adulthood. Subjects were requested to consider each of the 11 aggressive behaviors into a 4-point Likert scale (0 = never; 1 = rarely; 2 = occasionally; 3 = frequently).

2.3. Data analysis

An experiment was conducted with two objectives, namely, maximizing the classification accuracy of SAs, and selecting the most discriminative set of items. Our analyses followed three steps. Initially, we explored the classification accuracy that could be obtained by automatically selecting the best items of the previously described scales and socio-demographic factors. Then, we determined the items that best discriminated between SA and non-SA. Finally, the classification accuracy achieved using the selected items was compared with the obtained by each individual scale.

In order to reach our objectives, we used elastic net (Lars-en) (Zou and Hastie, 2005), which is a variant of stepwise linear regression that usually improves its performance. Also based on forward selection, this method modifies the optimization function by adding constraints to the L1 and L2 norms of the vector of coefficients. This technique minimizes

\[ |y - X\beta|^2 + \lambda_1 |\beta_1| + \lambda_2 |\beta_2| \]

where

\[|\beta_1| = \sum_{j=1}^{p} |\beta_j|\]

and

\[|\beta_2| = \sum_{j=1}^{p} \beta_j^2\]

The response is centered and the predictors are standardized before applying the technique. This rather simple modification has proven very powerful, and can be understood as a generalization of other two popular methods: least angle regression (Lars) (Efron et al., 2004) and least absolute shrinkage and selection operator (Lasso) (Tibshirani, 1996).

In order to conduct the experiment, the data set was randomly divided into three sets: training set, evaluation set and test set. Each set was composed of 146 non-SA and 82 SA. The training set was used to train the Lars-en algorithm. Because the classifier accuracy depends on the selected variables and the threshold, the evaluation set was used to determine the optimal configuration parameters. These parameters were first adjusted according to the values that maximized the accuracy in the evaluation set and then used for the test set. Average results after 100 repetitions of this process are reported below.

The scales included in this study were presented to all the subjects. However, some items were removed before starting the experiment. Regarding the IPDE-SQ, item 25 (“I have never threatened suicide or injured myself on purpose”) reflects suicidal behavior and therefore was excluded from the analysis. Item 49 (“I often seek advice or reassurance about everyday decisions”) was also removed due to the existence of non-responding subjects. There were no missing values in the remaining items. According to the SRRS scale, we found that certain life events were extremely rare in our sample. For instance, only 0.5% people presented with “jail term or probation”, and none of them showed “change in religious activities”, “minor financial loan”, or “change in schools”. Uncommon life events have limited clinical interest and furthermore, if included in the analysis, the covariance matrix would be singular and the estimation of parameters would have been problematic. Therefore life events with a frequency lower than 5% were removed from the study. The remaining set of variables for the analysis consisted of 154 variables (30 BIS-11 items, 75 IPDE-SQ items, 33 BGHA items, 14 SRRS items, age and sex).

3. Results

Classification results are displayed in Table 3. The items selected by the Lars-en algorithm attained an average accuracy of 85.3% in classifying SA.

For a better understanding of the predictive capacity of these scales, Fig. 1 shows the respective average receiver operating characteristic (ROC) curves together with the average area under the curves.
Once we observed that a suitable accuracy could be obtained, it was necessary to decide which items should be part of an accurate scale for measuring risk of suicidal behavior. In order to do this, each item was assigned to an index that indicates its relevance. The index is defined as the absolute average weight associated to the item in the projections obtained by the Lars-en algorithm in the 100 repetitions of the experiment. The scale was constructed by selecting the n most relevant items that maximized the average classification accuracy in the evaluation set. Fig. 2 displays the average accuracy obtained in the evaluation set when the n more relevant items were included. The maximum was attained using the 27 most relevant items.

Table 4 displays the selected items and their weights. The total score of a particular subject is obtained summing (or subtracting, if negative) each item value (0 or 1 in the SRRS and IPDE-SQ; 1 to 4 in the BIS-11; and 0 to 3 in the BGHA scale) multiplied by each item weight. For instance, a 30 years-old subject who is in the process of marital separation (SRRS), but considers that he/she usually get fun and enjoyment out of life (IPDE-SQ) and always plan for job security (BIS-11) would score 30 (years) × 0.2 (weight) + 1 × 15 (weight) + 1 × (−8.1) (weight) + 3 × (−2.6) (weight) = 51 (see Table 2 for weight's information). Using this set of items, we obtained an average accuracy of 86.4%, a specificity of 89.6%, and a sensitivity of 80.8% in classifying suicide attempters. These results are similar to those reported in Table 1, thus giving further support to the item selection by the Lars-en algorithm. Of note, the proposed scale did not include sex. Two facts might explain this result. First, sex was the 30th more relevant feature. As it can be appreciated in Fig. 2, the difference in the fitness function when selecting 27 or 30 items is minimal. Secondly, the effect of sex might be captured in the scale through other items. For instance, the selected item “revision of personal habits” which is closely associated with suicidal behavior in females.

Finally, we tested if the scores obtained by the SA and non-SA groups were significantly different. Then, we calculated the scores of the whole sample using the proposed scale. The Lilliefors goodness-of-fit test of composite normality showed that the projections of the SA and the non-SA (psychiatric inpatients and blood donors) groups followed Gaussian distributions. Mean (±SD) total scores of blood donors, psychiatric inpatients, and suicide attempters were 8.4 (±14.6), 37.1 (±21.9), 75.9 (±28.7), respectively. In order to provide a visualization of these groups, Fig. 3 adjusts a mixture of Gaussians to the histogram of the scores. Intuitively, it is appreciated that there are differences between the three groups. An ANOVA test with Bonferroni correction verified this assumption (F = 717.23; df = 2; p < 0.001).

4. Discussion

In this article, the Lars-en algorithm analyzes the accuracy that can be obtained in the classification of individuals as SA or non-SA, by selecting the most discriminating items of well-known psychiatric scales. Our results indicate that SA can be accurately classified using a set of 27 items. This set of items showed an average accuracy of 86.4%, a specificity of 89.6%, and a sensitivity of 80.8% in classifying SA.

According to their weights, the items most closely associated to SA status came from the IPDE-SQ (personality traits) and SRRS (life events). This is in accordance with the stress-diathesis model of...
suicide (Mann et al., 1999b) and literature suggesting that suicide is related to personality and triggered by dramatic life events (Blasco-Fontecilla et al., 2010). Additionally, 26% (7/27) of the items included in the scale were related to the BIS-11 and BGHA scales in agreement with the expected association of impulsive aggression traits with suicidal behavior (Giegling et al., 2009). If we analyze individually the items that received the highest weights, we can observe that “change in the number of arguments” and “marital separation” were two of the most relevant life events, which probably reflects the relevance of distressful life events in social and marital domains for SAs. All these findings, along with the fact that the most relevant personality items according to their relative weights indicated emptiness (“I often feel empty inside”), unhappiness (“I usually get fun and enjoyment out of life”), and dependency needs (“I worry about being alone and having to care for myself”), suggest that individuals unable to cope with problematic relationships with other people are at risk of attempting suicide.

Another interesting finding is that 18% (5/27) of the selected items had negative weights. Subjects endorsing items with negative weights such as “I usually get fun and enjoyment out of life” (IPDE-SQ), “I plan trips well ahead of time” (BIS-11), or “Armed aggression to others” (BGHA) were less likely to be a SA. This is in accordance with the scarce available literature. For instance, an inverse association between suicide rates and happiness or life satisfaction has been reported (Bray and Gunnell, 2006). A more recent study also reported an inverse association between well-being and suicide intent in a sample of 469 SA (Sisask et al., 2008). The negative weight of the items “I plan trips well ahead of time” and “I plan for job security” is coherent with the fact that most suicide attempts in our sample were impulsive (only 23.7% thought on suicide more than three hours before the attempt). “Armed aggression to others” was the factor with the largest negative weight, which might be compatible with the classical Freudian assumption that the externalization of aggression could protect from suicidal tendencies (Freud, 1947), although this is a controversial, poorly studied topic (Ferreira de Castro et al., 1986).

Therefore, a sense of happiness, planning ahead, and externalization of aggression should be considered protective factors by clinicians evaluating suicidal risk.

Major strengths of the present study are the relatively large sample and the use of an efficient algorithm (Lars-en) that selects the most suitable items to assess suicidal risk. Our study presents some limitations and should just be considered a preliminary step in the development of a new scale for suicide risk assessment. The findings reported here require replication in other samples and different settings. The resulting set of items should be eventually validated and compared to specific scales that assess suicidal risk. Although our results suggest that suicide attempters can be accurately detected without information about Axis I disorders (which are consistently associated with suicidal behavior), including this information might increase the accuracy of our results.

5. Conclusion

The reduced number of items selected by the Lars-en algorithm in this study suggests that this set of items could be used as a quick, feasible but accurate instrument to assist clinicians in the evaluation of suicide risk. For instance, it might help primary care physicians in deciding which patients are at risk of suicide and should be referred to a psychiatrist or even hospitalized. It might also assist psychiatrists in evaluating short-term suicide risk in the emergency departments.

Contributors

Dr. Baca-Garcia designed the study and wrote the protocol. Dr. Blasco-Fontecilla and Dr. Lopez-Castroman managed the literature searches. Dr. Delgado-Gomez, Dr. Aguado and Dr. Ruiz-Hernandez undertook the statistical analysis. Dr. Blasco-Fontecilla and Dr.
Delgado-Gomez wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.

Conflict of interest

None author report any conflict of interest.

Acknowledgments

This article was supported by the National Alliance for Research on Schizophrenia and Affective Disorders (NARSAD), Fondo de Investigacion Sanitaria (FIS) PI060092, Fondo de Investigacion Sanitaria FIS RD06/0011/0016, ETES (PI07/90207), the Conchita Rabago Foundation, and the Spanish Ministry of Health, Instituto de Salud Carlos III, CIBERSAM [Intramural 521 Project, P91B; SCO/3410/2004]. Dr. Blasco-Fontecilla acknowledges the Spanish Ministry of Health (Río Hortega CM08/00170), Alicia Koplowitz Foundation, and Conchita Rabago Foundation for funding his postdoctoral stage at CHRU, Montpellier, France. The work by Jorge Lopez-Castroman was supported by a FondaMental Foundation grant.

References


Bray I, Gunnell D. Suicide rates, life satisfaction and happiness as markers for population mental health. Social Psychiatry and Psychiatric Epidemiology 2006;41:333–7


Christiansen E, Jensen BF. Risk of repetitive suicide attempt, suicide or all deaths after an episode of attempted suicide: a register-based survival analysis. The Australian and New Zealand Journal of Psychiatry 2007;41:257–65


