Psychiatry Research 269 (2018) 474–480

Contents lists available at ScienceDirect

Psychiatry Research

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

Identifying Chinese adolescents with a high suicide attempt risk

Xu Yanjing*, Wang Chihhuan, Shi Mengmeng

Department of Teacher Education, Zhejiang Normal University, Zhejiang, China

ARTICLE INFO

Keywords:
Suicide attempt risk
Decision tree
CRUISE
Adolescents

ABSTRACT

Adolescent suicide has become a serious public health problem in China. Since suicide attempts are considered the strongest predictor of suicide completion, it is valuable to identify adolescents with a high suicide attempt risk. This study aimed to develop a decision tree model for the interactive prediction of high suicide attempt risk in Chinese adolescents. A classification tree analysis was conducted in a sample of senior high school students (N = 6,686) based on the CRUISE program. The results indicated that depression, anxiety, social support, gender, self-esteem, family cohesion and adaptability were significant predictors of high suicide attempt risk, and interactions among these predictors constructed a hierarchical decision tree model. The tree model offered a series of reliable rules to identify Chinese adolescents with high suicide attempt risk, for example, adolescents with high depression scores had the highest probability (69.22%) of having a suicide attempt, female adolescents with low social support and low depression scores had the second highest probability (57.58%), and adolescents with low anxiety, low family adaptability, and medium depression scores had the third highest probability (55.77%). These exploratory findings suggested that different screening criteria are needed to detect at-risk Chinese adolescents with different severities of depression.

1. Introduction

Adolescent suicide has been considered a major public health problem. Although the suicide rate varies among different countries, suicide has become one of the leading causes of death for adolescents worldwide (Silverman et al., 2007). In China, it has been reported that suicide is the first leading cause of death for adolescents (Xian yun, 2001). A meta-analysis found that 17.7% of Chinese adolescents reported having suicide ideation, 7.3% adolescents reported having planned to commit suicide, and 2.7% adolescents tried to commit suicide but survived (Dong et al., 2014).

Given that adolescent suicide has long-lasting effects on the family members and friends left behind, a considerable number of studies have been conducted to develop reliable decision rules that can be used to classify adolescents into the high-risk category for attempting suicide. Moreover, since suicide has been considered to be an interaction of various risk factor (Mościcki, 2001), the investigation of interactions associated with adolescent suicide attempt risk is a fundamental step toward identifying at-risk populations, which will allow for earlier prevention and identification.

However, the main difficulty of these suicide related studies was that the data are often complex, non-normal, and involve high-order interactions (Triponez et al., 2013). The commonly used statistical methods (e.g., logistic regression) often fail to find meaningful patterns from such data and have been recognized as cumbersome in addressing these types of classification problems (Long et al., 1993). Classification tree analyses (CTAs) are modern statistical techniques ideally suited for both exploring and modeling such data (De'ath and Fabricius, 2000). CTAs are data-driven techniques for analysis of complex ecological data, including many algorithms such as CART (Classification and Regression Tree), QUEST (Quick Unbiased Efficient Statistical Tree), CHAID (Chi-squared Automatic Interaction Detection) and CRUISE (Classification Rule with Unbiased Interaction Selection and Estimation).

In general, CTA builds the decision tree model based on recursive partitioning of data to maximize the predictive accuracy of a given categorical outcome (Hill et al., 2017). The optimal cutoff point is identified using binary splits to maximize the accuracy of that predictor for classifying cases on the dependent variable, and this cutoff point then becomes a “node”. The first node is based on the best correlation with the dependent variable; the subsamples are then divided into further non-overlapping subsamples. This partitioning process will not stop until the trees are pruned to the best level using a cost complexity criterion. Moreover, Kim indicated that CRUISE was shown to perform the best among these algorithms according to several characteristics (Kim and Loh, 2001). CRUISE uses two techniques to improve the

* Corresponding author.
E-mail address: 493780546@qq.com (Y. Xu).

https://doi.org/10.1016/j.psychres.2018.08.085
Received 20 September 2017; Received in revised form 22 August 2018; Accepted 24 August 2018
Available online 28 August 2018

interpretability of its trees. First, it splits each node into multiple subnodes, with one for each class. Second, it selects variables based on one-factor and two-factor effects, indicating that CRUISE can immediately identify a variable with a significant two-factor interaction even when it does not have a significant one factor effect. Furthermore, CRUISE uses a bootstrap bias correction to reduce the bias due to differences between numerical and categorical variables. Therefore, CRUISE was considered a more sophisticated method among the CTAs.

It has been reported that there are three main advantages of a CTA compared to traditional methods (De'ath and Fabricius, 2000), including (1) the ability to use various types of outcome variables, including numeric, categorical, ratings, and survival data; (2) easy graphical interpretation of complex results involving interactions; and (3) the ability to handle missing values in both response and predictor variables. Thus, CTA has been increasingly considered an alternative method to many traditional methods, such as multiple regression, logistic regression, and survival models.

To date, CTAs (especially the CART) have been already applied in suicide-related studies. A CART was conducted in 220 suicide attempters, and the results implied that clinic doctors should pay particular attention to depressed individual with affective conflicts or non-depressed individual with affective conflicts and low monthly income (Jia et al., 2005). Another CART model comprised of 75 factors was built to describe early childhood predictors of suicidal thoughts to identify subgroups of adolescence at high risk of suicidality, and their findings suggests female adolescents reported stressful life events between 0 and 5 years with family member died at the same time had the highest probability to have suicidal thoughts (Dykxhoorn et al., 2017). Bae et al. (2015) built a classification tree model in Korean adolescents by applying the CHAID and found that depressed females with higher delinquency and those with potential depression with lower family intimacy had the highest possibility to attempt suicide. Most recently, Hill et al. (2017) utilized a large, longitudinal data set and CRUISE in American adolescents to provide a means for prospectively identifying adolescents at risk for suicide ideation, and their results offer community organizations options for instituting large-scale screenings for suicide ideation risk.

Despite the increasing number of studies examining the interactive effects of various predictors of suicidality, only one study in the literature concentrated on adolescent suicide attempt risk. Furthermore, to our knowledge, no existing literature has investigated the interactions between suicide related factors in Chinese adolescents by applying CTAs. Considering that a suicide attempt is a strong predictor of completed suicide (Bridge et al., 2006), the present study aimed to address this issue and provide an empirical basis for decision-making and referral via a classification tree approach to predict high suicide attempt risk using data from a large sample of Chinese adolescents.

To establish the classification tree model for the prediction of high suicide attempt risk in Chinese adolescents, ten risk factors were selected based on previous findings and theories, including gender (Joe and Marcus, 2003), school grade (Björkenstam et al., 2011), depression (Esposito and Clum, 2002), anxiety (Yen et al., 2014), hopelessness (Abramson et al., 1998), self-esteem (Kuhlberg et al., 2010), social support (Kleiman and Liu, 2013), life satisfaction (Valois et al., 2004), family adaptability and cohesion (Wang et al., 2012).

2. Method

2.1. Participants

Participants in this study were derived from 8 high schools in 7 cities in Zhejiang province, and they were asked to complete a packet of questionnaires during a class period while one teacher supervised them with instructions. Assent was obtained from each participant prior to participation. The participants were informed regarding the aims of this study and their identities were ensured to be confidential. A total of 7023 participants completed the test battery, while 337 (4.8%) of these participants had more than 10% omissions in data. Of the remaining sample, 47.5% were males, 52.5% were females, 55% were in grade one, 24.1% were in grade two, and 20.9% were in grade three. In addition, the percentage of missing data in the remaining sample ranged from 1.5% to 4.3%. Data analysis was conducted using the remaining sample (N = 6686).

2.2. Measurements

2.2.1. Dependent variable: suicide attempt risk

The Suicide Behaviors Questionnaire-Revised (SBQ-R, 4 items) is a self-report questionnaire designed to assess the risk of a suicide attempt for children and adolescents, and a total score of 7 or higher indicates a high risk of suicide attempt for non-suicidal samples (Osman et al., 2001). Considering that the SBQ-R and the dichotomization ($\geq$7) were developed in Western samples and have rarely been used in Chinese adolescents, two preliminary analyses were conducted. First, we translated the original SBQ-R into Chinese and conducted a pilot study (N = 166) to test its reliability and validity in Chinese adolescents. The results of the pilot study found a moderate level of internal consistency (α = 0.795), while good validity was demonstrated by a significant correlation between the SBQ-R and the Beck Depression Inventory (r = 0.765, p < 0.001). The internal consistency of the current data was 0.784. Second, we applied a receiver operating characteristic (ROC) analysis to examine the risk cutoff score for the SBQ-R among Chinese adolescents using the first item of the SBQ-R. A score of 4 (i.e. I have attempted to kill myself and did not want to die/really hoped to die) on the first item of the SBQ-R was considered the characterization of suicide attempt risk. Results (Table 1) indicated that a total score of 7.5 is the cutoff score for the SBQ-R in Chinese adolescents. However, we reasoned that the accuracy of this cut-off score ($\geq$7.5) might be confounded considering the state variable (the first item of the SBQ-R) is a part of the test variable (raw score of SBQ-R) in the ROC analysis. Therefore, we suggested a score of 7 or higher indicates a high risk of suicide attempt in Chinese adolescents in line with previous studies.

2.2.2. Independent variables: demographic, psychological and family variables

Demographics included gender and school grade. The psychological variables assessed included the following: (1) Depression was assessed by the Beck Depression Inventory. This inventory is a 21-question multiple choice self-report inventory to assess depression severity. The items are answered on 4-point scale ranging from 0 (e.g., I do not feel

Table 1
Receiver operating characteristic analysis to distinguish high suicide attempt risk.

Score	Sensitivity	Specificity	Youden's index*
2	1	0	0
3.5	1	0.366	0.366
4.5	1	0.467	0.467
5.5	1	0.508	0.508
6.5	0.981	0.664	0.645
7.5	0.957	0.708	0.665
8.5	0.884	0.758	0.642
9.5	0.845	0.793	0.638
10.5	0.777	0.8	0.577
11.5	0.687	0.937	0.624
12.5	0.554	0.964	0.518
13.5	0.424	0.982	0.406
14.5	0.347	0.993	0.34
15.5	0.272	0.998	0.27
16.5	0.17	1	0.17
17.5	0.121	1	0.121
19	0	1	0

* Youden's index = Sensitivity + Specificity-1