Journal of Psychiatric Research 96 (2018) 15–22

Contents lists available at ScienceDirect

Journal of Psychiatric Research

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

Predicting posttraumatic stress disorder following a natural disaster

Anthony J. Rosellini [a, b, *], Francisca Dussaillant [c], José R. Zubizarreta [a, d], Ronald C. Kessler [a], Sherri Rose [a]

[a] Department of Health Care Policy, Harvard Medical School, Boston, MA, USA
[b] Center for Anxiety and Related Disorders, Boston University, Boston, MA, USA
[c] School of Government, University of Desarrollo, Santiago, Chile
[d] Division of Decision, Risk and Operations, and Department of Statistics, Columbia University, New York, NY, USA

A R T I C L E I N F O

Article history:
Received 1 February 2017
Received in revised form
23 July 2017
Accepted 7 September 2017

Keywords:
Posttraumatic stress disorder
Machine learning
Natural disaster
Earthquake
Risk score

A B S T R A C T

Earthquakes are a common and deadly natural disaster, with roughly one-quarter of survivors subsequently developing posttraumatic stress disorder (PTSD). Despite progress identifying risk factors, limited research has examined how to combine variables into an optimized post-earthquake PTSD prediction tool that could be used to triage survivors to mental health services. The current study developed a post-earthquake PTSD risk score using machine learning methods designed to optimize prediction. The data were from a two-wave survey of Chileans exposed to the 8.8 magnitude earthquake that occurred in February 2010. Respondents (n = 23,907) were interviewed roughly three months prior to and again three months after the earthquake. Probable post-earthquake PTSD was assessed using the Davidson Trauma Scale. We applied super learning, an ensembling machine learning method, to develop the PTSD risk score from 67 risk factors that could be assessed within one week of earthquake occurrence. The super learner algorithm had better cross-validated performance than the 39 individual algorithms from which it was developed, including conventional logistic regression. The super learner also had a better area under the receiver operating characteristic curve (0.79) than existing post-disaster PTSD risk tools. Individuals in the top 5%, 10%, and 20% of the predicted risk distribution accounted for 17.5%, 32.2%, and 51.4% of all probable cases of PTSD, respectively. In addition to developing a risk score that could be implemented in the near future, these results more broadly support the utility of super learning to develop optimized prediction functions for mental health outcomes.

Earthquakes are a common and deadly natural disaster that can result in both ground shaking and tsunami waves. According to the Center for Research on the Epidemiology of Disasters (2016), earthquakes affected roughly one hundred million people and resulted in over 700,000 deaths worldwide between 2000 and 2015. Although earthquake exposure has been associated with several adverse psychosocial consequences (e.g., depression; suicidality), posttraumatic stress disorder (PTSD) is typically found to be the most prevalent negative mental health outcome (North, 2014). A recent meta-analysis of 46 studies of earthquake survivors found an overall post-earthquake PTSD incidence of 23.7% (Dai et al., 2016).

Although a large literature has identified risk factors associated with post-earthquake (including post-tsunami) PTSD (Cairo et al., 2010; Chen et al., 2014; Cheng et al., 2014; Dell'Osso et al., 2013; Kun et al., 2009, 2013; Lai et al., 2004; Priebe et al., 2009; Rosendal et al., 2014; Sattler et al., 2014; Tural et al., 2004; van Griensven et al., 2006; Wang et al., 2009, 2011; Wen et al., 2012; Zhang et al., 2011), few studies have examined how to combine risk factor information into a risk score that can be used to predict who is most likely to develop post-earthquake PTSD. The development and use of clinical tools to identify individuals at high risk of PTSD is consistent with the American Red Cross PsySTART program (Schreiber et al., 2014). In PsySTART, aid workers meet with survivors in the immediate aftermath of a natural disaster to complete a risk factor checklist. Decisions about triaging survivors to mental health interventions are determined based on the total number of 13 risk factors that are present (American Red Cross, 2012). Given that regression coefficients vary widely across risk

* Corresponding author. Center for Anxiety and Related Disorders, Department of Psychological and Brain Sciences, Boston University, 648 Beacon Street, 6th Floor, Boston, MA 02215, USA.
E-mail address: ajrosellini@gmail.com (A.J. Rosellini).

https://doi.org/10.1016/j.jpsychires.2017.09.010

factors for post-earthquake PTSD (Cheng et al., 2014; Tural et al., 2004; Zhang et al., 2011), assuming all predictors equally contribute to PTSD risk may not result in optimal prediction. The one existing study to develop a regression-based risk score for post-earthquake PTSD used rounded main terms coefficients (no interactions) from a logistic regression of 11 risk factors (Liu et al., 2012). However, it is unclear if maximum prediction accuracy was achieved given evidence of interactions among predictors of post-earthquake PTSD (Dell'Osso et al., 2013; Fan et al., 2011).

Another way to develop a post-earthquake PTSD risk score would be through machine learning methods designed to optimize prediction. Machine learning has been used to develop risk scores for PTSD onset related to other types of traumatic events (Galatzer-Levy et al., 2014, 2017; Karstoft et al., 2015a; Karstoft et al., 2015b). A number of popular machine learning algorithms are described in Table 1. There are several reasons why machine learning algorithms might outperform standard parametric method. In comparison to conventional main terms regression examining the direct effect of predictors on an outcome, for example, there are machine learning algorithms available that automatize identification of interactions and non-linearities (e.g., multivariate adaptive regression splines, Friedman, 1991; random forests, Breiman, 2001; Bayesian trees, Chipman et al., 2010). In addition, whereas a conventional regression based on highly correlated independent variables (e.g., injury and amputation, Liu, et al., 2012) might have good prediction accuracy in the sample which it was developed but perform poorly in independent samples (model overfit), machine learning methods can be employed to reduce the likelihood of overestimating prediction performance. For example, penalized regression algorithms (i.e., *regularization*) prevent overfit by shrinking coefficients among collinear variables (Friedman et al., 2010).

Ensembling methods refer to a type of machine learning in which multiple algorithms are consolidated into a single algorithm with improved prediction performance. Super learning (van der Laan et al., 2007; van der Laan and Rose, 2011) is an ensembling method that is particularly well suited to develop risk scores (see Rose, 2013 for a mortality risk score example) because of its flexibility in generating a consolidated algorithm from a number of different prediction approaches. In other words, a super learner algorithm is able to simultaneously (i) capture the relationships of predictors with an outcome (e.g., using conventional regression), even if predictors are highly correlated (e.g., using penalized regression), and (ii) detect interactions and nonlinear associations (e.g., using decision-tree or spline algorithms).

Super learning has been used in one study to develop a risk score for PTSD related to any type of traumatic event (Kessler et al., 2014). In this study, the super learner algorithm outperformed a select number of individual algorithms (including logistic regression) in predicting PTSD based on several hundred risk factors. However, that study was limited as the data were cross-sectional, relied on retrospective reports of PTSD symptoms and risk factors, and a small proportion of the sample had disaster-related PTSD. Accordingly, the goal of the current study was twofold: (i) to demonstrate how machine learning methods can be used to develop a more accurate post-earthquake PTSD risk score than conventional regression methods, and (ii) to develop a preliminary model-based risk score for post-earthquake PTSD that could be expanded or adapted in future epidemiological disaster research.

1. Materials and methods

1.1. Sample

The data came from a two-wave household survey of individuals living in Chile at the time of the 8.8 magnitude earthquake occurring on February 27, 2010. The survey was conducted by Chile's Ministry of Planning and Cooperation (Division Observatorio Social, 2010) using fully-structured face-to-face interviews. The pre-earthquake survey was a biennial nationally representative socioeconomic-health survey conducted between November and December 2009. In order to understand the public health impact of the earthquake, including the tsunami that occurred in some coastal areas, a subsample of baseline respondents were re-interviewed between May and June 2010. Of the 27,000 households asked to participate in the post-earthquake survey, 22,456 agreed (response rate = 83.2%). Additional details of the survey design are available elsewhere (Division Observatorio Social and AGCI Ministerio de Relaciones Exteriores, 2015). The sample used here consisted of all 23,907 adults who participated in both surveys and completed the post-earthquake PTSD assessment.

1.2. Outcome measure

Probable DSM-IV PTSD was assessed in the post-earthquake survey using the Davidson Trauma Scale (DTS; Davidson et al., 1997). The DTS was administered in Spanish to assess past-week PTSD symptoms specifically in relation to the earthquake and tsunami. The DTS assesses the frequency and severity of all 17 DSM-IV PTSD symptoms using a 0–4 scale (total score 0–136). The reliability and factor structure of the DTS has been supported among Chileans exposed to the 2010 earthquake (Leiva-Bianchi and Araneda, 2013). Comparison of the English DTS to independent structured interview-based PTSD diagnosis suggests that a DTS total score ≥40 indicates a probable PTSD diagnosis (Davidson et al., 1997). PTSD research conducted in Spanish speaking countries suggests that the ≥40 cut score demonstrates good concurrent validity with a PTSD diagnosis using the Spanish Clinician Administered PTSD Scale ($\kappa = 0.78$; Coronas et al., 2008). Several studies of Spanish-speaking samples have applied this cut score (e.g., Leiva-Bianchi and Araneda, 2013; Ruiz-Parraga and Lopez-Martinez, 2014). In the current sample, 9.9% of respondents lived in areas where the earthquake could not be felt (or no tsunami waves). We identified cases of probable PTSD based on (i) living in an area affected by the disaster (conservative confirmation of PTSD Criterion A1), and (ii) having a DTS total score ≥40 (13.3% had probable PTSD, $n = 3182$).

1.3. Independent variables

We reviewed two areas of literature to identify risk factors for PTSD among adults: (i) studies of risk factors specifically relevant to post-earthquake PTSD (cited in the introduction), and (ii) systematic reviews of risk factors for PTSD related to *any* type of natural disaster (Goldmann and Galea, 2014; Norris et al., 2002) and *any* type of traumatic event (Ozer et al., 2003; Sayed et al., 2015). Consistent with the literature, we organized risk factors into pre-earthquake factors (present *before* the trauma), peri-earthquake factors (objective and subjective experiences-severity immediately surrounding the trauma), and post-earthquake factors (present *after* the trauma).

We identified all survey questions that could be used to operationalize risk factors identified in the literature review. As the purpose of the surveys was not to study the complete range of all previously identified PTSD risk factors, especially peri-earthquake factors, we supplemented the survey data with other publically available data about the severity-impact of the earthquake (described below). The goal of the analysis was to optimize prediction of probable PTSD, not to test a conceptual model. We consequently operationalized as many risk factors as possible, regardless of if they had been directly (e.g., sex, age, property