Journal of Psychiatric Research 59 (2014) 68–76

Contents lists available at ScienceDirect

Journal of Psychiatric Research

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

Quantitative forecasting of PTSD from early trauma responses: A Machine Learning application

Isaac R. Galatzer-Levy [b, *], Karen-Inge Karstoft [b, c], Alexander Statnikov [d, e], Arieh Y. Shalev [a, b]

[a] Center for Traumatic Stress Studies, Hadassah University Hospital, Jerusalem, Israel
[b] Department of Psychiatry, NYU School of Medicine, New York, NY, USA
[c] Department of Psychology, University of Southern Denmark, Odense, Denmark
[d] Center for Health Informatics and Bioinformatics, NYU School of Medicine, New York, NY, USA
[e] Department of Medicine, NYU School of Medicine, New York, NY, USA

ARTICLE INFO

Article history:
Received 25 March 2014
Received in revised form
28 July 2014
Accepted 28 August 2014

Keywords:
Posttraumatic stress disorder (PTSD)
Course and prognosis
Early prediction
Forecasting
Machine Learning
Markov boundary feature selection
Support Vector Machines

ABSTRACT

There is broad interest in predicting the clinical course of mental disorders from early, multimodal clinical and biological information. Current computational models, however, constitute a significant barrier to realizing this goal. The early identification of trauma survivors at risk of post-traumatic stress disorder (PTSD) is plausible given the disorder's salient onset and the abundance of putative biological and clinical risk indicators. This work evaluates the ability of Machine Learning (ML) forecasting approaches to identify and integrate a panel of unique predictive characteristics and determine their accuracy in forecasting non-remitting PTSD from information collected within10 days of a traumatic event. Data on event characteristics, emergency department observations, and early symptoms were collected in 957 trauma survivors, followed for fifteen months. An ML feature selection algorithm identified a set of predictors that rendered all others redundant. Support Vector Machines (SVMs) as well as other ML classification algorithms were used to evaluate the forecasting accuracy of i) ML selected features, ii) all available features without selection, and iii) Acute Stress Disorder (ASD) symptoms alone. SVM also compared the prediction of a) PTSD diagnostic status at 15 months to b) posterior probability of membership in an empirically derived non-remitting PTSD symptom trajectory. Results are expressed as mean Area Under Receiver Operating Characteristics Curve (AUC). The feature selection algorithm identified 16 predictors, present in ≥95% cross-validation trials. The accuracy of predicting non-remitting PTSD from that set (AUC = .77) did not differ from predicting from all available information (AUC = .78). Predicting from ASD symptoms was not better then chance (AUC = .60). The prediction of PTSD status was less accurate than that of membership in a non-remitting trajectory (AUC = .71). ML methods may fill a critical gap in forecasting PTSD. The ability to identify and integrate unique risk indicators makes this a promising approach for developing algorithms that infer probabilistic risk of chronic posttraumatic stress psychopathology based on complex sources of biological, psychological, and social information.

1. Introduction

Chronic PTSD is prevalent, distressful, and debilitating (Kessler, 2000) and often follows an unremitting course (Galatzer-Levy et al., 2013; Peleg and Shalev, 2006). The early manifestations may provide sufficient information to identify individuals at risk for chronic PTSD. Studies to date have identified numerous risk indicators of chronic PTSD, many of which are accessible shortly after trauma exposure. These include, but are not limited to early symptoms of PTSD, depression or dissociation, physiological arousal (e.g., heart rate), early neuroendocrine responses, gender, lower socioeconomic status, the early use of opiate analgesics, the occurrence of traumatic brain injury, and a progressively growing number of genetic and transcriptional factors (Boscarino et al., 2012; Brewin et al., 2000; Etkin and Wager, 2007; Karl et al., 2006; Ozer et al., 2003). Despite these discoveries, the individual identification of risk for PTSD remains elusive, thereby leaving a major gap between scientific discovery and practical application.

One reason for such a gap is the current use of computational models that do not match the disorder's inherent complexity in

* Corresponding author. New York University School of Medicine, 1 Park Ave, New York, NY 10016, USA. Tel.: +1 847 420 2527.
E-mail address: Isaac.Galatzer-Levy@nyumc.org (I.R. Galatzer-Levy).

http://dx.doi.org/10.1016/j.jpsychires.2014.08.017
0022-3956/

etiology. As attested by its numerous risk factors, the etiology of PTSD is multi-causal, multi-modal, and complex. As such, the longitudinal course of PTSD reflects a converging interaction of numerous, multimodal risk factors. Moreover,specific risk markers and their relative weight can vary between individuals and traumatic circumstances. For example, head injury increases the likelihood of developing PTSD (Bryan and Clemans, 2013) but has a low occurrence overall among survivors. Similarly, the contribution of female gender to the risk of developing PTSD varies between traumatic events (Kessler et al., 1995) and with specific genetic risk alleles (Ressler et al., 2011). Thus, to accurately predict PTSD in individuals, one must account for complex and variable interactions between putative markers.

The commonly utilized General Linear Modeling (GLM) was designed to test focused hypotheses without generalization beyond the data under study (Hald, 2007). This modeling approach identifies the probability of rejecting a null hypothesis of no effect along with an estimate of shared variance between the dependent and independent variable (s) (Cohen, 1994). Such models are important for rigorously testing novel hypotheses because they assume no relationship between variables. In this context, a significant *p*-value indicates disconfirmation of the assumption of no relationship between variables. Further, estimates of unique and shared variance provide information about how much of the variability in the dependent variable (DV) is accounted for by the independent variable (s) (IVs). This approach has intrinsic limitations. First, it is built on the assumption that variables would follow a normal distribution given an infinite sample size (Stigler, 1986). Second, information provided is limited to null hypothesis testing along with effect size estimates of shared variance. Third, relatively large sample-to-variable ratios are needed for such analyses because relationships between variance components are being analyzed. What is needed to forecast later outcomes is the probability of the DV given the IV(s) along with an estimate of accuracy. Further, as many identified predictors may be redundant, methods to identify variables that provide unique information are also important. Finally, as many different variables may provide predictive information, analysis and integration of many variables simultaneously is required.

Machine Learning (ML) can handle large complex data with heterogeneous distributions (Hastie et al., 2003), determine probabilistic relationships from complex conditional dependencies between variables, and test the reliability of the results through repeated cross validation. The use of ML to determine a later outcome is known as forecasting. Forecasting is increasingly used in developing personalized medicine (e.g., using tissue biomarkers to predict the course of malignancies (Cruz and Wishart, 2006)). Recent ML neuroimaging studies have shown promising results in predicting the course of neuropsychiatric disorders (Orrù et al., 2012).

Psychiatry currently relies on descriptive information. Establishing the usefulness and limits of non-invasive, low cost information such as this is an important baseline to build upon with new, potentially more invasive and more expensive methods for identifying risk. However, the implementation of ML methods to clinical observations is limited. Investigators have used clinical information to forecast violent behavior among outpatients with schizophrenia (Tzeng et al., 2004) achieving moderate success of 76.2% positive prediction of later violent behavior. In contrast, ML has failed in forecasting the course of bipolar disorder from clinical data (Moore et al., 2012). A review of studies forecasting the risk for psychotic disorders (Strobl et al., 2012) showed an advantage for ML-based Support Vector Machines (SVMs) classification. In this example, predictive accuracy ranged from 100% to 78% relative to GLM-based variance accounted for ranging from 81% to 67%. This study further demonstrated that forecasting from multimodal information (e.g., quality of life and neuroimaging data) increases the accuracy of prediction. This conclusion was also reached by Marinić et al. (2007) who used ML random forests to compare the classification of PTSD based on the clinical assessment of PTSD alone versus the clinical assessment of PTSD along with other clinical assessment tools. This study demonstrated large improvements (from predictive accuracy of 70.59% based only on PTSD symptom assessment to 78.43% integrating other symptoms).

The current work work evaluates the use of ML to forecast chronic PTSD from data available shortly after a traumatic event. This study utilizes data from a previously published, fifteen months long, longitudinal study of 975 trauma-exposed individuals admitted to a general hospital emergency department within hours of their traumatic events (Shalev et al., 2012). In the current work, we compare a) forecasting accuracy based on all available information utilized indiscriminately, b) forecasting accuracy based on a subset of variables that are selected using a feature selection algorithm, and c) forecasting accuracy from ASD symptoms alone. The current study also compares the prediction of PTSD diagnosis at fifteen months with that of an empirically derived non-remitting PTSD symptoms trajectory (Galatzer-Levy et al., 2013). To meet the specific challenge of early prediction, this work uses data obtained within ten days of a traumatic event.

2. Materials and methods

2.1. Participants and procedures

Data for the current study come for the Jerusalem Trauma Outreach and Prevention Study (J-TOPS; (Shalev et al., 2012), ClinicalTrial.Gov identifier: NCT00146900). The J-TOPS combined a systematic outreach and comprehensive follow-up design with an embedded, randomized, controlled trial of early interventions. Sampling procedures and population parameters of study subjects are fully described in Shalev et al. (2012).

Subjects in the current study were adults who were admitted to Hadassah University Hospital emergency department (ED) immediately following potentially traumatic events (age 18–70). Following identification in the ED, potential subjects were screened using a short telephone interview 9.21 ± 3.20 days following ED admission. Those with acute PTSD symptoms were invited for clinical interviews, which took place 29.51 ± 4.93 days after ED admission. Participants were re-evaluated five, nine, and fifteen months after ED admission regardless and blind of their participation in the nested clinical trial. Participants provided informed consent for all aspects of the study with procedures approved and monitored by the Hadassah University Hospital's institutional review board (IRB).

2.1.1. Current study sample

Included in this study are participants with valid data at 10 days post-trauma and at least two additional time points. The resulting sample consisted of 957 participants. The initial traumatic event exposure included motor vehicle accidents (84.1%), terrorist attacks (9.4%), work accidents (4.4%), and other incidents (2.0%). Participants in this study did not differ from the J-TOPS larger sample in gender distribution, age, ten days symptom severity, and the number of new traumatic events occurring during the study (for full description, see (Galatzer-Levy et al., 2013); Table 1).

2.2. Instruments

To forecast PTSD, we used *all information items* collected during participants' ED admissions and phone interviews during the first ten days following trauma. The resulting 68 items (alias, “features”) include demographic data, ED observations, and instruments