Journal of Psychiatric Research 121 (2020) 159–172

Contents lists available at ScienceDirect

Journal of Psychiatric Research

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

The use of machine learning techniques in trauma-related disorders: a systematic review

Luis Francisco Ramos-Lima[a,c,*], Vitoria Waikamp[a,c], Thyago Antonelli-Salgado[b], Ives Calvalcante Passos[a,b], Lucia Helena Machado Freitas[a,c]

[a] Post-graduate Program in Psychiatry and Behavioral Sciences, Federal University at Rio Grande do Sul, Porto Alegre, Brazil
[b] Bipolar Disorder Program, Laboratory of Molecular Psychiatry, Clinical Hospital of Porto Alegre, Porto Alegre, Brazil
[c] Psychological Trauma Research and Treatment Program (NET-Trauma), Clinical Hospital of Porto Alegre, Porto Alegre, Brazil

ARTICLE INFO

Keywords:
Machine learning
Forecasting
Psychological trauma
Posttraumatic stress disorders

ABSTRACT

Establishing the diagnosis of trauma-related disorders such as Acute Stress Disorder (ASD) and Posttraumatic Stress Disorder (PTSD) have always been a challenge in clinical practice and in academic research, due to clinical and biological heterogeneity. Machine learning (ML) techniques can be applied to improve classification of disorders, to predict outcomes or to determine person-specific treatment selection. We aim to review the existing literature on the use of machine learning techniques in the assessment of subjects with ASD or PTSD. We systematically searched PubMed, Embase and Web of Science for articles published in any language up to May 2019. We found 806 abstracts and included 49 studies in our review. Most of the included studies used multiple levels of biological data to predict risk factors or to identify early symptoms related to PTSD. Other studies used ML classification techniques to distinguish individuals with ASD or PTSD from other psychiatric disorder or from trauma-exposed and healthy controls. We also found studies that attempted to define outcome profiles using clustering techniques and studies that assessed the relationship among symptoms using network analysis. Finally, we proposed a quality assessment in this review, evaluating methodological and technical features on machine learning studies. We concluded that etiologic and clinical heterogeneity of ASD/PTSD patients is suitable to machine learning techniques and a major challenge for the future is to use it in clinical practice for the benefit of patients in an individual level.

1. Introduction

Trauma-related disorders such as Posttraumatic stress disorder (PTSD) and acute stress disorder (ASD) are considered to be debilitating conditions, developed from exposure to traumatic events including war, mass violence, natural disasters, and accidents. The *DSM-5* (American Psychiatric Association, 2013) lists 20 diagnostic criteria for PTSD divided into four clusters of symptoms: re-experience of the traumatic event; avoidance; persistent negative thoughts or feelings; trauma-related arousal and reactivity. The WHO World Mental Health Survey conducted across 24 countries found a lifetime prevalence of any traumatic event of 70.4% (Benjet et al., 2016), suggesting that constitutional and sociocultural factors are also involved in the development of the disorder, besides the magnitude of trauma (Yehuda, 2004). The prevalence of PTSD in a lifetime is 11% for women and 5.5% for men (Kessler et al., 1995). It is postulated that a dose–response relationship exists between exposure to traumatic events and the subsequent development of PTSD, indicating that prior trauma and/or multiple traumatic event exposures increase the risk of the disorder (Ozer et al., 2008; Kilpatrick et al., 2013).

Establishing the diagnosis of PTSD and ASD has always been a challenge in clinical practice, as well as in academic research. As indicated by its numerous risk factors, the etiologies of trauma disorders are multicausal and complex. In addition, the development of diagnostic criteria for classification systems (such as *DSM-5*) has been elaborated from research with chronic populations and in tertiary care settings; such phenotypic expressions may not reflect the instability and nonspecific nature of the phenomenology of the disorder in its development (McGorry et al., 2006). Evidence-based, trauma-focused therapies with the most support are cognitive- and exposure-based approaches, with prolonged exposure and cognitive processing therapy being the most investigated (Charney et al., 2018). Notwithstanding,

* Corresponding author. Post-graduate Program in Psychiatry and Behavioral Sciences, Federal University at Rio Grande do Sul, Ramiro Barcelos St, 2400, Porto Alegre/RS, 90035-002, Brazil.
E-mail address: lfrlima@hcpa.edu.br (L.F. Ramos-Lima).

https://doi.org/10.1016/j.jpsychires.2019.12.001
Received 29 September 2019; Received in revised form 22 November 2019; Accepted 5 December 2019

establishing first-line psychotherapies may be difficult because of—among other aspects—the burden to patients and patient profiles (Nash and Watson, 2012). Some statistically significant results provided by evidence-based medicine may not represent a real benefit for an individual patient; subjects in clinical trials do not always reflect the multimorbidity profile of "real-life" patients (Greenhalgh et al., 2014). This may be particularly true in the field of PTSD, where clinical heterogeneity can be a very important factor, not always taken into account in research.

Machine learning, a field of computer science and a part of artificial intelligence, refers to the science and engineering by which machines (i.e., computer systems) can analyze and acquire information from data (Liu and Salinas, 2017). Machine learning can help develop sophisticated data models using advanced mathematical techniques and handling complex data sets with heterogeneous distribution. The 'learning' method is usually made by a supervised or an unsupervised approach (Bishop, 2006). In supervised learning, the user feeds the machine with input data and expected outcome: the machine learns a mapping from the input to the outcome target, through classification (where the output variable is a category, such as 'disease' or 'no disease') or regression (where the output variable is a numeric variable) methods. Common examples of supervised learning algorithms are Logistic Regression, Support Vector Machines and Neural Networks. Supervised learning is often used to estimate prediction and risk: the Framingham Risk Score for coronary heart disease may be one of the most famous uses of supervised learning in medicine (Deo, 2015; Kannel et al., 1975). Unsupervised learning does not depend on previous associations and output variables: the goal is to model the underlying data structure to learn more about the data. It can be performed by discovering groups of similar cases (clustering) or determining the distribution of available data (density estimation). Network analysis allows visualization of the connectivity among symptoms and clusters of symptoms providing knowledge about the strength and quantity of relationships (Sullivan et al., 2018), taking into account regression and clustering techniques. A revision of the relevant principles of machine learning and its limitations can be found elsewhere (Schultebraucks and Galatzer-Levy, 2019; Librenza-Garcia et al., 2017; Deo, 2015).

Machine learning techniques can be applied to improve classification of disorders, to predict risk factors and treatment outcomes, and to improve person-specific treatment selection (Hahn et al., 2017). Since PTSD and ASD are disorders that present clinical and biological heterogeneity, which may constitute a barrier to understanding the causative mechanisms and to developing optimal treatments and diagnostic tools, machine learning is a suitable approach to better achieve this understanding. The present study aims to systematically review data in which PTSD and ASD were assessed through machine learning techniques regarding classification, prognostic, and treatment selection studies. Furthermore, we proposed a method of quality measurement of these studies.

2. Methods

This study followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) reporting guidelines (Liberati et al., 2009) and is registered on the International Prospective Register of Systematic Reviews (PROSPERO identifier CRD42019115850). We searched PubMed, Embase, and Web of Science for articles published between January 1960 and May 2019 using terms associating machine learning techniques with PTSD and ASD. The complete filter is available in the supplementary material. Additionally, we searched the reference lists to find potential articles to include. There were no language restrictions.

Articles met the inclusion criteria if they assessed PTSD and/or ASD patients in childhood or adulthood using machine learning techniques. Technical and theoretical studies that used machine learning techniques but did not assess ASD/PTSD patients and studies evaluating traumatic brain injuries (TBI) not related to PTSD were excluded. Also, we excluded preclinical and review studies addressing ASD/PTSD.

2.1. Data collection, extraction, and statistical analysis

Two researchers (LFRL and VW) independently screened titles and abstracts of the identified articles. They then obtained and read the full texts of potential articles; TAS made the final decision in cases of disagreement. All processes during primary and secondary screening were supervised by ICP. Data extracted from the articles included year of study publication, type of data used in the machine learning model (i.e., neuroimaging, blood biomarkers, clinical and demographical characteristics, among others), sample size, scales and diagnoses assessed in the study, machine learning algorithm, and statistical measure of performance (i.e., accuracy, sensitivity, specificity, and area under the curve [AUC]). Information such as use of controls on the sample, outcome assessment, machine learning characteristics (description, metrics), use of testing data set, feature selection, use of hyperparameters, and handling of missing data were also retrieved through quality evaluation of the studies. We contacted the authors of three studies for additional information—the authors of two studies provided the relevant data on request. ICP and LHMF aided in interpreting the results. All authors discussed the results and contributed to the final version of the manuscript.

We also developed a quality assessment to use in this review, as there is currently no instrument for this purpose in machine learning studies. We considered the methodological features comprising sample representativeness, confounding variables, and outcome assessments as the most clinically relevant aspects among machine learning-based healthcare research. The remaining dimensions assess the quality of specific components of the machine learning approach that were used in a given study. In summary, we reviewed the algorithm used, the description of accuracy of a given model or other performance metrics, how missing data and class imbalance problems had been handled, evidence that the model had been tested on unseen data, and evidence that results were optimized using hyperparameter optimization and feature selection procedures. Supplementary Table 1 describes the dimensions used in this specific analysis. The results of the quality assessment are described in Section 3.4 and presented in Supplementary Table 2.

3. Results

We found a total of 806 potential abstracts and included 49 articles in the present review. Fig. 1 shows the study selection process. A list of the included articles as well as the most relevant characteristics and findings are presented in Table 1 (Prognostic studies), Table 2 (Classification studies), and Table 3 (Network analysis and unsupervised studies).

Thirty-three articles assessed prognosis, most in order to predict risk factors related to the development of PTSD or to identify its early symptoms (Table 1). Of these, eight used neuroimaging studies (Zandvakili et al., 2019; Nicholson et al., 2018; Yuan et al., 2018; Im et al., 2017; Jin et al., 2017; Li et al., 2016; Wang et al., 2016; Cisler et al., 2015); 13 used questionnaires in open or semi-structured format (Leightley et al., 2019; Rosellini et al., 2018a; Augsburger and Elbert, 2017; Conrad et al., 2017; Gradus et al., 2017; He et al., 2017; Reece et al., 2017; Schalinski et al., 2016; Karstoft et al., 2015a, 2015b; Köbach et al., 2015; Kessler et al., 2014; Marinić et al., 2007); three used biological samples (Galatzer-Levy et al., 2017; Hemmings et al., 2017; Tylee et al., 2015); and nine used audio and/or medical records (Harrington et al., 2019; Marmar et al., 2019; Papini et al., 2018; Saxe et al., 2017; Wortwein and Scherer, 2017; Dabek and Caban, 2015a, 2015b; Vergyri et al., 2015; Galatzer-Levy et al., 2014). Eight articles used machine learning techniques to build diagnostic classification