UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM PSIQUIATRIA E CIÊNCIAS DO COMPORTAMENTO

Luís Francisco Ramos Lima

AVALIAÇÃO DO ESTADIAMENTO DO TRANSTORNO DE ESTRESSE PÓS-TRAUMÁTICO: UM ESTUDO COM APRENDIZADO DE MÁQUINA

(STAGING EVALUATION OF POSTTRAUMATIC STRESS DISORDER: A MACHINE LEARNING STUDY)

Porto Alegre, 10 de Março de 2021

Luís Francisco Ramos Lima

AVALIAÇÃO DO ESTADIAMENTO DO TRANSTORNO DE ESTRESSE PÓS-TRAUMÁTICO: UM ESTUDO COM APRENDIZADO DE MÁQUINA

(STAGING EVALUATION OF POSTTRAUMATIC STRESS DISORDER: A MACHINE LEARNING STUDY)

Tese apresentada à Universidade Federal do Rio Grande do Sul, Faculdade de Medicina, Programa de Pós-Graduação Psiquiatria e Ciências do Comportamento, como requisito parcial para obtenção do título de **Doutor em Psiquiatria**.

Orientadora: Profa. Dra. Lúcia Helena Machado Freitas (UFRGS/HCPA)
Coorientadora: Profa. Dra. Andrea Feijó de Mello (UNIFESP)
Colaboradora: Dra. Stefania Teche (HCPA)

Porto Alegre, 10 de Março de 2021

CIP - Catalogação na Publicação

Ramos Lima, Luis Francisco
AVALIAÇÃO DO ESTADIAMENTO DO TRANSTORNO DE ESTRESSE PÓS-TRAUMÁTICO: UM ESTUDO COM APRENDIZADO DE MÁQUINA / Luis Francisco Ramos Lima. -- 2021.
167 f.
Orientadora: Lucia Helena Machado Freitas.

Coorientadora: Andrea Feijó Mello.

Tese (Doutorado) -- Universidade Federal do Rio Grande do Sul, Faculdade de Medicina, Programa de Pós-Graduação em Psiquiatria e Ciências do Comportamento, Porto Alegre, BR-RS, 2021.

1. Transtornos Relacionados a Trauma e Estresse. 2. Aprendizado de Máquina. 3. Psiquiatria. I. Freitas, Lucia Helena Machado, orient. II. Mello, Andrea Feijó, coorient. III. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

Aos meus pais, irmãs e sobrinhos, minha profunda gratidão
pelo amor e pelo carinho.
Ao Aldrin, obrigado por estar sempre disposto a nós e por manter
sempre a construção do nosso amor.

AGRADECIMENTOS

Este trabalho jamais poderia ter sido realizado por apenas duas mãos. Gostaria de agradecer à minha querida orientadora, Profa. Dra. Lucia Helena Machado Freitas, por ter me acolhido como orientando mesmo antes do início do doutorado, e por não ter se furtado a ser "apenas" orientadora; agradeço por todo o carinho, paciência, disponibilidade e amizade. Agradeço também à minha querida coorientadora, Profa. Dra. Andrea Feijó de Mello, por ter aceito a proposta do meu trabalho e por ter me acolhido com amizade e afeto em São Paulo; estendo também esse agradecimento ao Prof. Dr. Marcelo Feijó de Mello, por terem mantido sempre as portas abertas e o apoio presente.

Agradeço também por todas as experiências que tive a oportunidade de vivenciar no ambulatório do Núcleo de Estudos do Trauma (NET-Trauma), cuja caminhada começamos a compartilhar em 2015 para não mais sair; foi local de aprendizado, de pesquisa, de dar meus primeiros passos como "preceptor" e de ótimas convivências com residentes e contratados. Personifico esse agradecimento à Stefânia Pigatto Teche: preceptora, coautora de diversos trabalhos, "coorientadora" e, principalmente, querida amiga. Agradeço também ao Programa de Atendimento à Violência e ao Estresse Pós-traumático (PROVE) da UNIFESP e toda a sua equipe, pelas tão enriquecedoras experiências e pelo acolhimento; além disso, de me proporcionarem a possibilidade de retornar um pouco do que recebi na forma de supervisor dos seus residentes. Agradeço também aos pacientes desses dois serviços que voluntariamente compartilharam suas vivências e dores, fundamentais para a execução desse projeto.

Devo também enorme agradecimento à minha "braço direito e braço esquerdo", querida amiga Vitoria Waikamp. Obrigado por acreditar no projeto, participar comigo de maneira sempre tão dedicada e voluntariosa, pela paciência nos momentos difíceis, pela parceria nos diversos trabalhos e artigos, e sobretudo pela amizade e carinho.

Agradeço também aos amigos que deram contribuições tão importantes ao longo desse caminho: ao Prof. Dr. Ives Cavalcante Passos, pela disponibilidade e pelos ensinamentos nessa área tão fascinante que é o *machine learning*; ao Thyago Salgado Antonelli, pelo companheirismo e sorriso no rosto sempre, e pela ótima troca de experiência que tivemos em nossos trabalhos. À querida amiga Andressa Behenck, pela parceria, pela amizade próxima e dedicada, pelas músicas, pelo carinho e por estar sempre próxima, nos momentos de alegria e nos momentos de desafio; certamente a tua presença tornou esse caminho muito menos arenoso.

Agradeço também à Profa. Dra. Cintia de Azevedo Marques Perico por ter me recebido na rede de saúde mental do município de São Bernardo do Campo/SP, me dado a oportunidade de contribuir com a área de ensino junto ao Centro Universitário Faculdade de Medicina do ABC (FMABC) e com o programa de residência da prefeitura de São Bernardo do Campo. À Alessandra Cardeal, rara combinação de chefe e amiga, por ter me dado enorme suporte, paciência e afeto nesse último ano. Agradeço também por me ensinar com suas palavras e atos o que é a promoção da saúde mental no Sistema Único de Saúde.

A idealização e a execução deste trabalho me permitiram combinar duas áreas importantes da minha trajetória pessoal: conciliar conhecimentos adquiridos já longinquamente, na graduação em Ciência da Computação da UFRGS (encerrada em 2001), com os conhecimentos e experiências da Medicina e sobretudo da Psiquiatria. Agradeço aos mestres daquela época, mesmo já afastados pelo tempo e pela minha atual prática profissional. Personifico esse agradecimento à Profa. Dra. Mariana Recamonde Mendoza (do Instituto de Informática da UFRGS, minha já remota casa), por ter gentilmente aceito cooperar com esse projeto e por auxiliar no planejamento e na execução das práticas dessa área tão instigante que é a de *machine learning*.

Agradeço ao Instituto Abuchaim (na pessoa da querida amiga Vanda Barbosa), e ao Centro de Estudos José de Barros Falcão (na pessoa da amiga e colega Profa. Glaydcianne Bezerra) por oportunizarem meus primeiros passos como professor e como supervisor, e aos alunos que me acolheram.

Aos meus pais, por serem fonte de cuidado e amor, e por serem alicerces da construção do meu indivíduo. Às minhas irmãs, sempre fonte de carinho e pela sorte de poder ter como irmãs duas amigas que amo. Aos meus sobrinhos, pelo amor e carinho sempre.

Ao Aldrin, meu marido, que acompanhou os êxitos e as tensões da trajetória deste doutorado desde os seus primórdios. Por estar ao meu lado sempre, sempre disposto a nós e a construir nossa família, por ser fonte de amor e tolerância frente aos desafios inerentes desta etapa.

EPÍGRAFE

"Não, nossa ciência não é uma ilusão. Mas seria uma ilusão acreditar que poderemos encontrar noutro lugar o que ela não nos pôde dar."

Sigmund Freud (1856-1939)

"Toda a nossa ciência, comparada com a realidade,
é primitiva e infantil – e, no entanto
é a coisa mais preciosa que temos."

Albert Einstein (1879-1955)

RESUMO

Os transtornos de estresse relacionados a um evento traumático, como o transtorno de estresse agudo (TEA) e o transtorno de estresse pós-traumático (TEPT), são caracterizados por alta morbidade e prejuízo social significativo. No Brasil, estima-se que 80% da população já foi exposta a pelo menos um evento traumático ao longo da vida em grandes centros urbanos, como São Paulo e Rio de Janeiro; o crescente problema da violência urbana mostra-se fator importante para a gênese dos transtornos relacionados ao trauma. Devido à etiologia do TEPT ser multicausal e complexa, técnicas de *Machine Learning* (Aprendizado de Máquina – ML) tem sido usadas para desenvolver escores de risco, para predição diagnóstica e para definição de tratamento. Contudo, considerando sua heterogeneidade clínica e etiológica, realizar o diagnóstico e definir um tratamento adequado pode ser muitas vezes desafiador. O uso do estadiamento clínico surge como um método mais refinado de diagnóstico, procurando definir a progressão do transtorno em momentos específicos durante o *continuum* da enfermidade. Esta abordagem pode auxiliar em um diagnóstico mais aprimorado, conhecer melhor o prognóstico e escolher o melhor tratamento de acordo com o estágio do transtorno. Assim, o TEPT aparece como um exemplo importante de como um método de estadiamento pode trazer benefícios. O objetivo desta tese é avaliar como os aspectos pessoais, clínicos e relacionados ao trauma dos pacientes atendidos em ambulatórios especializados em trauma psíquico podem estar relacionados à predição do estadiamento clínico de TEPT usando técnicas de ML.

Palavras-chave: Transtorno de Estresse Pós-traumático; Aprendizado de Máquina; Estadiamento; Predição

ABSTRACT

Stress disorders related to a traumatic event, such as acute stress disorder (ASD) and posttraumatic stress disorder (PTSD), are characterized by high morbidity and significant social impairment. In Brazil, it is estimated that 80% of the population has already been exposed to at least one traumatic event throughout life in large urban centers, such as São Paulo and Rio de Janeiro; the growing problem of urban violence proves to be an important factor in the genesis of trauma-related disorders. The etiology of PTSD is multicausal and complex; techniques of Machine Learning (ML) have been used to develop PTSD risk scores, to predict its diagnosis and to choose better treatments. However, considering its clinical and etiological heterogeneity, making the diagnosis and defining an appropriate treatment can often be challenging. The use of clinical staging appears as a refined method of diagnosis, aiming to define the progression of the disorder at specific times during the *continuum* of the illness. This approach may provide improved diagnosis, better understand the prognosis and choose the best treatment according to the stage of the disorder. Thus, PTSD appears as an important example of how a staging method can bring benefits. The objective of this thesis is to evaluate how the personal, clinical and trauma-related aspects of patients who sought care at outpatient clinics specialized in emotional trauma can be related to the prediction of the PTSD staging using ML techniques.

Keywords: Posttraumatic Stress Disorder; Machine Learning; Staging; Prediction

APRESENTAÇÃO

Este trabalho consiste na tese de doutorado intitulada "**Avaliação do Estadiamento do Transtorno de Estresse Pós-traumatico: um Estudo em Aprendizado de Máquina**", apresentada ao Programa de Pós-Graduação em Psiquiatria e Ciências do Comportamento, da Universidade Federal do Rio Grande do Sul, em 10 de Março de 2021. Optou-se por escrever o corpo desta tese na língua inglesa para possibilitar maior disseminação do conhecimento no meio acadêmico.

O material é dividido em seis partes, na ordem que se segue: (1) Introdução, (2) Justificativa, (3) Objetivos, (4) Métodos, (5) Resultados e (6) Considerações finais e conclusão. Na seção "Resultados", estão apresentadas as seguintes publicações:

- Artigo 1 (*Short communication*): "**Trauma-Related Disorders in a Low- to Middle-income Country: a Four-year Follow-up of Outpatient Trauma in Brazil**"
- Artigo 2: "**The Use of Machine Learning Techniques in Trauma-Related Disorders: a Systematic Review**"
- Artigo 3: "**Prediction of Posttraumatic Stress Disorder Staging from Clinical and Sociodemographic Features: a Machine Learning Approach**"

Nos anexos, encontram-se as escalas utilizadas e o Termo de Consentimento Livre e Esclarecido. Ainda, encontram-se dois artigos resultantes de colaboração com a Universidade Federal de São Paulo (UNIFESP), ligados diretamente aos objetivos deste estudo:

- Artigo adicional 1 (*Letter to the editor*): "**The Clinician-Administered PTSD Scale (CAPS-5): Adaptation to Brazilian Portuguese**"
- Artigo adicional 2: "**Validation of the Brazilian-Portuguese Version of the Clinician-Administered Posttraumatic Stress Disorder Scale-5 (CAPS-5)**"

LISTA DE ABREVIATURAS E SIGLAS

AUC: Area under the curve

CAPS-5: Clinician-administered posttraumatic scale (version DSM-5)

CNS: Conselho nacional de saúde

CPT: Cognitive Processing therapy

DSM: Diagnostic and Statistics Manual

EBM: Evidence-based medicine

HCPA: Hospital de clínicas de Porto Alegre

HPA: Hypothalamic-pituitary-adrenal

ICD: International classification of diseases

LEC-5: Life events checklist (version DSM-5)

ML: Machine learning

NET-Trauma: Psychological trauma research and treatment program

PCA: Principal component analysis

PROVE: Program for research and care on violence and PTSD

PTSD: Posttraumatic stress disorder

SNRI: Serotonin-norepinephrine reuptake inhibitors

SSRI: Selective serotonin reuptake inhibitors

UFRGS: Universidade federal do Rio Grande do Sul

UNIFESP: Universidade federal de São Paulo

SUMMARY

1 INTRODUCTION

Emotional trauma and its consequences for the individual, especially posttraumatic stress disorder (PTSD), acquired scientific and social importance from the study of "war neurosis", which showed that traumatic situations could have as consequences psychopathological effects of high prevalence and morbidity [1]. In 1970, Shatan and Lifton [2] observed veterans from the Vietnam War, reports on Holocaust survivors and existing research on victims of traffic accidents and burns; they compiled a list of the 27 most common symptoms of the victims of such "neurosis", which was compared with the clinical records of 700 veteran patients in Vietnam, constituting the embryo of the criteria used by DSM-III to include the diagnosis of PTSD [3].

PTSD is currently considered a debilitating condition, developed from exposure to traumatic events including war, mass and interpersonal violence, natural disasters and accidents. Exposure is defined as directly experiencing or witnessing a traumatic event, or learning that a trauma occurred to a close family member or friend. PTSD can also develop from repeated or extreme exposure to aversive details of traumatic events. The fifth edition of the Diagnostic and Statistical Manual (DSM-5) [4] lists 20 diagnostic criteria divided into four clusters of symptoms: re-experience of the traumatic event; avoidance; persistent negative thoughts or feelings; trauma-related arousal and reactivity.

Although most individuals experience a traumatic event during their lifetime, most of the people exposed to a traumatic event do not develop PTSD. A study in the North American population corroborated these findings, showing that 40–70% of the individuals had experienced at least one major trauma during life, suggesting that constitutional and socio-cultural factors are also involved in the development of PTSD, besides the magnitude of trauma [5]. A study in the Brazilian population showed incidence of exposure to at least one traumatic event during lifetime of almost 90% in two large urban centers [6]. Studies on prevalence of PTSD found lifetime rates of 11% for women and 5.5% for men [7]. In the Brazilian context, PTSD has also become an important public health problem: a population-based study conducted in São Paulo and in Rio de Janeiro – two major urban cities in Brazil – found a prevalence of PTSD during lifetime of 11.4–14.7% in women and 4.7–7.8% in men [6], similar to prevalence in other countries such as the USA [7].

PTSD is likely to occur after more severe types of trauma, such as rape, childhood sexual abuse or military combat [8]. A meta-analysis performed by Ozer et al. [9] identified

seven major predictors of PTSD: (a) prior trauma, (b) prior psychological adjustment, (c) family history of psychopathology, (d) perceived life threat during the trauma, (e) posttrauma social support, (f) peritraumatic emotional responses, and (g) peritraumatic dissociation. It is postulated that a dose–response relation exists between exposure to traumatic events and the subsequent development of PTSD, indicating that prior trauma and/or multiple traumatic event exposures increase risk of PTSD [10, 11]. Among the factors that could be involved in the occurrence of trauma, a recent study in our enviroment has indicated that patients who developed ASD/PTSD reported more previous trauma history or childhood trauma [8]. A longitudinal study, conducted in acute care centers from six countries, indicated a PTSD prediction model where female gender, having less than secondary education and prior interpersonal trauma (added to DSM-5 diagnostic criteria) were associated with higher PTSD likelihood at the end of follow-up [11]. Preceding trauma in adulthood and the ocurrence of a childhood trauma are also associated with a worse prognosis of PTSD evolution: previous findings support the hypothesis that exposure to intense stress during early life can have long-term effects on mental health, increasing the risk of developing later psychopathology [12, 13]. These findings are in line with evidence that previous traumatic experiences may have a negative cumulative effect on mental health [6, 14], reinforcing the idea of sensitization of the individual who, exposed to previous trauma, may develop a greater response to subsequent stressors [15].

The etiology of PTSD is multi-causal and complex as indicated by countless risk factors involved in its development. The identification of risk at an individual level for PTSD remains difficult to achieve, leaving a gap between scientific discovery and practical application. For instance, a study conducted by Galatzer-Levy et al. [16] attempted to predict the clinical course of 957 trauma survivors, concluding that the prediction of PTSD from ASD symptoms was not better than chance. Specific risk factors and their relative weight can vary between individuals and traumatic circumstances: as an example, the contribution of female gender to the risk of developing PTSD varies between traumatic events [7] and with specific genetic risk alleles [17]. Several research studies have demonstrated alterations in neural circuitry, molecular biology, endocrinology and immune reactivity and other physiological domains in PTSD. A review to identify potential biomarkers for PTSD has included monoaminergic transmitter systems, the hypothalamic-pituitary-adrenal (HPA) axis, metabolic hormonal pathways, inflammatory mechanisms, indicators of psychophysiological reactivity and neural circuits [18].

Neurobiology in PTSD is becoming an important field of study, influenced by the emergence of modern technologies. Autonomic reactivity (represented mostly by “cluster D” hyperarousal symptoms) shows the involvement of the hypothalamic-pituitary-adrenal (HPA) axis and sympathetic nervous system, which, in the long term, results in exaggerated negative feedback sensitivity [19]. Moreover, specific brain networks and neurochemistry are involved in PTSD; patients can present a myriad of emotional symptoms, which can be undermodulated or overmodulated. Emotional undermodulation, represented by fear, anger, guilt and shame responses, emotional numbing, is consistent with decreased ventromedial prefrontal cortex and activation of rostral anterior cingulate and amygdala [20]. Emotional overmodulation is related to the opposite: increased activation of medial prefrontal cortex and rostral anterior cingulate regions, leading to decreased amygdala activation [21]. PTSD patients also seem to present an enhanced salience network and a decreased default mode network connectivity [22]. It is also postulated the importance of intergenerational transmission of trauma effects, by early environmental exposures to the offspring such as intra-utero, or epigenetic changes associated with a preconception trauma in parents that may affect the germline [23].

Regarding treatment, guidelines for PTSD consider psychotherapy to be the front-line choice, generally recommending this intervention before the prescription of medications [24], in contrast to most other psychiatric disorders associated with significant levels of impairment. Evidence-based, trauma-focused therapies with most support are cognitive- and exposure-based approaches: prolonged exposure (PE) and cognitive processing therapy (CPT) are the two most cited and rigorously investigated [25]. Even though, the use of first-line psychotherapies may sometimes be difficult because of the lack of trained mental health-care practitioners, the burden to patients, and patient preference [26]. It seems that, despite more biological findings in PTSD, there are no specific drug targets, and the only medications that are FDA approved for treatment of PTSD have been developed for other conditions, such as depressive and anxiety disorders [19]. Pharmacotherapies can be used when evidence-based psychotherapies are not available, are ineffective or on the basis of patient preference; pharmacotherapy with the most support for PTSD includes selective serotonin reuptake inhibitors (SSRI) and serotonin-norepinephrine reuptake inhibitors (SNRI) [25].

The current classification format is criticized for not considering those neurobiological aspects and psychosocial factors and for being dependent on clusters of symptoms evaluated in a cross-sectional observation. 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 non-specific nature of the disorder's phenomenology in its development [27]. Psychiatric research traditionally relies on diagnosis made by classification systems. Diagnostic categories, however, suffer from significant underlying heterogeneity, different distributional properties and limited empirical accuracy for predictive models [28]. In PTSD it is no different: the lack of a longitudinal perspective is reflected by the fact that the neurobiology involved in the disorder was recognized as a single entity, without considering the occurrence of several stages in the progression of the disorder [29]. Pathways to PTSD involve distinct genetic, endocrine, demographic and environmental factors that are not shared by all PTSD patients, suggesting that efficient treatment may have to address individual-specific pathogenic pathways [28].

1.1 MACHINE LEARNING

The practice of evidence-based medicine (EBM) means integrating individual clinical expertise with the best available external clinical evidence from systematic research. EBM has helped us understand risk factors, optimal treatments and prognosis of disorders by using traditional statistical methods that primarily provide average group-level results [30]. Some statistically significant results may not represent a real benefit for an individual patient and subjects in clinical trials may not always reflect the multimorbidity profile of real-life patients [31]. This may be particularly true in the PTSD field, where clinical heterogeneity can be a very important factor, not always taken into account in research. In fact, several studies conducted in specific populations, such as American veterans [32, 33] or groups of patients sharing a particular traumatic event, such as large-scale accidents or mass disasters, are observed in the current literature of PTSD [34, 35], possibly reproducing a partial vision of the disorder. Techniques that aim at developing tailor-made psychiatric care to the individual from more heterogeneous information have been gaining ground in psychiatric research. Computational psychiatry approaches combine multiple levels and types of computation with multiple types of data in an effort to improve understanding, prediction and treatment of mental illness [36].

Machine Learning (ML), 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 data and acquire information from that data [37]. Unlike conventional statistics,

ML can help develop much more sophisticated data models using advanced mathematical techniques and handling complex data sets with heterogeneous distribution. Therefore, ML methods using heterogeneous data can be applied to *prediction* (risk factors, treatment outcomes), to improve *classification* of disorders or improve *treatment selection*. Since PTSD is a disorder that presents clinical and biological heterogeneity, which may constitute a barrier to understanding the causative mechanisms and to developing optimal treatments and diagnostic tools, it is suitable to use a ML approach to improve this understanding.

The success of a ML algorithm depends upon which features to extract and the criteria for developing (training) a data model. These features and criteria are likely to vary across ML applications, as well as the final performances of algorithms. Typically, these algorithms are implemented in two key stages [38]: first, data are separated into algorithm 'training' and 'testing' sets with the former being used to 'train' a ML model. Second, the accuracy of the 'learning' model is evaluated using the 'testing' set that consists of observations previously 'unseen' by the model. ML models can be evaluated in terms of *sensitivity*, *specificity* and *area under the curve* (AUC) of a receiver operating characteristic test. In many cases, cross-validation is employed, in which a subset of the data is used to build/train the model, and a different subset is used to test/validate the model as described above. This cycle can be repeated to improve model performance [16, 37]. If a test set that has never been used previously could be evaluated separately and the model is successful, it could indicate that a more stable association is found, which may allow generalization of this ML model [39].

The 'learning' method of a ML algorithm is usually made by either a *supervised* or an *unsupervised* approach [40]: 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 continuous variable, such as "weight" or a scale number) methods. Supervised learning is often used to estimate prediction and risk: the Framingham Risk Score for coronary heart disease may be the most famous use of supervised learning in medicine [41]. Unsupervised learning does not depend on previous associations and no corresponding output variables. The goal for unsupervised learning is to model the underlying structure or distribution in the data in order 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*).

While ML techniques were developed mainly from the theory and application of artificial neural networks, it now encompasses a much more diverse set of algorithms and

techniques [37]. There are several reasons why ML algorithms may outperform standard parametric method: for example, in comparison to conventional main terms regression examining the direct effect of predictors on an outcome, some ML algorithms can automatize identification of interactions and non-linearities (e.g., random forests and bayesian trees). In addition, where a conventional regression based on highly correlated independent variables may have good prediction accuracy in the sample in which it was developed but perform poorly in independent samples (model overfitting), ML methods can be used to reduce the likelihood of overestimating prediction performance [42]. A number of popular ML algorithms are described in Table 1.

Table 1. Brief descriptions of common ML algorithms.

Algorithm	Description
Conventional Logistic Regression	• Traditional parametric logistic regression • Prone to overfit if independent variables are highly collinear • Optimal functional form of independent variables unknown (e.g., linear versus non-linear)
Regularization Ridge Elastic net Lasso	• Penalized regression reduces overfit due to collinear independent variables • Ridge regression shrinks coefficients for collinear independent variables toward zero, but does not fully eliminate any independent variable • Elastic net regression allows various penalties where coefficients for collinear independent variables are shrunk toward zero (but not eliminating contributions to the predicted probability) and/or to zero (eliminating their contributions to the predicted probability) • Lasso regression shrinks coefficients for collinear covariates to zero, eliminating their contributions to the predicted probability
Decision trees Random forest Bayesian regression trees	• Decision tree methods capture interactions and non-linear associations • Independent variables are partitioned (based on values), stacked to build decision trees, and assembled into a Bayes tree aggregate "forest" • Random forests build numerous trees in bootstrapped samples and generates an aggregate tree by averaging across trees (reducing overfit) • Bayesian trees are based on an underlying probability model (priors) for the structure and likelihood for data in terminal nodes; the aggregate tree is generated by averaging across tree posteriors (reducing overfit)
Support vector machines Linear kernel Polynomial kernel Radial kernel	• Support vector machines treat each independent variable as a dimension in high dimensional space and attempt to identify the best hyperplane to separate the sample into classes (e.g., cases and non-cases) • Goal is to find the hyperplane with the maximum margin between the two closest points in space • Captures linear associations, but alternate kernels can be used to capture non-linearities (polynomial and radial basis kernels were used here)

Source: adapted from Rosellini et al., 2018 [42]

ML has been used to develop risk scores for PTSD onset related to several types of traumatic events [16, 28, 43, 44]. Different predictors will almost certainly be found in

different weight in diverse populations (for example, military veterans, disaster survivors, civilians in less developed countries) and in different screening settings (outpatient setting, emergency rooms, medical clinics in war zones, trauma units) [45], demanding different ML strategies. An investigation by Kessler et al. [46] indicated random forest as the best ML approach to predict PTSD after trauma exposure. Another study indicates support vector machines as suited for data with heterogeneous non-normal distributions consistent with the data that PTSD researchers are interested in integrating for prediction (self-reported and biological, for instance) [16]. Also as a prediction tool, Kobach et al. [47] used a random forest approach from clinical data on Burundian ex-combatants to discover that the number of lifetime experienced traumatic events would be the main predictor for posttraumatic stress in that scenario.

Regarding classification of disorders, a single diagnostic classifier of PTSD may never perform with 100% accuracy, which is why it will be essential to pursue the characteristics of subjects for whom such a classifier does not work: larger samples using ML techniques could be a solution for this problem [48]. For instance, a growing number of studies have applied ML techniques to neuroimaging data to characterize psychiatric diseases including PTSD [49]. A study on magnetoencephalography using Bayesian classification methods indicated that veterans with PTSD had significantly stronger neural activity in prefrontal, sensorimotor and temporal areas, as well as stronger activity in the bilateral amygdalae, parahippocampal and hippocampal regions compared with those without PTSD [50]. Several other studies applied ML techniques over functional and structural magnetic resonance imaging to identify patients with PTSD [51, 52, 53]. Tylee et al. [48] used a support vector machine to develop a pilot study using blood-based gene expression biomarkers in an attempt to identify marines with and without PTSD, achieving 90% prediction accuracy based on a panel with 20 exons.

Despite its many uses, a ML model is not without its critics. The complexity ‘hidden’ in ML techniques and the necessary knowledge on computer science in order to generate them may be part of the reason why ML have not pervaded trauma care and clinical practice in psychiatry. Another reason may be that many clinicians or researchers do not understand the principles of ML algorithms and their potential to change practice. A model performance can be altered by the nature of the data (source, amount and type) [54]: for instance, a small sample size could inflate the performance of a ML model. Issues such as computational power, heterogeneity of data, generalization and lack of interpretability of ML results have also been found and need to be addressed.

1.2 STAGING

The concept of staging has been used for many years as a useful tool in other complex diseases, such as diabetes mellitus, cardiovascular and neoplastic diseases [55]. Staging emerges as a more refined method of diagnosis; differing from a conventional diagnostic practice, staging tries to define the extent of progression of disease at particular points in time through the continuum of the illness. This approach could help to refine diagnosis, adjust prognosis and choose the best treatment according to illness stage [56]. Staging links the clinical aspects to treatment selection and prediction: it may be that its role in treatment selection is more crucial to change outcome than its role in prediction, particularly since early successful treatment may change the prognosis and thus prevent progression to subsequent stages [57].

Indexed approaches (for example, DSM-5 and ICD-11) identify the existence of a particular disease state. Despite the real utility of these approaches, there is a growing need in psychiatry to rate severity, extent and characteristics of mental illness. Fava and Kellner, in 1993 [58], first proposed the application of staging to psychiatric disorders, developing a cross-sectional staging approach to manic episodes based on symptom severity. According to them, the neglect of staging in psychiatry parallels its reliance on cross-sectional descriptions instead of longitudinal study of prodromes, the fully developed disorder and residual states. In 2006, McGorry and colleagues [57] introduced a staging model that highlighted the longitudinal course of psychiatric diseases in the psychotic spectrum, also integrating mood disorders. Aspects such as biological and endophenotypic markers, indicators of illness extension and progression were included in their model.

Currently, some areas of psychiatry are advanced in developing specific models of staging. With regard to severe psychiatric diseases, a recent Spanish systematic review [59] has found staging approaches to bipolar disease (Berk et al., 2007 [60]; Kapczinski et al., 2009 [61]; Duffy, 2014 [62]), unipolar depression (Fava and Kellner, 1993 [58]; Fava and Tossani, 2007 [63]; Hetrick et al., 2008 [64]), and schizophrenia (Lieberman et al., 2001 [65]; Singh et al., 2005 [66]; McGorry et al., 2010 [67]; Agius et al., 2010 [68]). For instance, Kapczinski et al. [61] proposed a staging approach to classify clinical stages in bipolar disease, considering clinical features, genotype modulation, effects of environmental pathogens (such as life stress) and biomarkers. A systematic review from the International Society for Bipolar Disorders Task Force reinforced the importance of this approach, indicating that late-stage bipolar patients have a worse overall prognosis and poorer response

to standard treatment, consistent with patterns for end-stage medical disorders [56] and suggesting the maintenance of staging to identify these patients earlier. Table 2 below presents some examples of most known staging models.

Table 2. Examples of staging models in psychiatry

	Clinical aspects
Kapczinski et al., 2009 [61]	Latent phase: At risk for developing BD, positive family history, mood or anxiety symptoms without criteria for threshold BD
Bipolar disease	Stage I: Well-defined periods of euthymia without overt psychiatric symptoms Stage II: Symptoms in inter-episodic periods related to comorbidities Stage III: Marked impairment in cognition and functioning Stage IV: Unable to live autonomously owing to cognitive and functional impairment
Hetrick et al., 2008 [64]	Stage 0: Increased risk of anxiety or depressive disorder; no symptoms currently
Unipolar depressive disorder	Stage 1 (1a, 1b): Mild or nonspecific symptoms of anxiety or depression, including neurocognitive deficits of severe mood disorder; mild functional change or decline; moderate but subthreshold symptoms of anxiety or depression, with moderate neurocognitive changes and functional decline to caseness (GAF < 70) Stage 2: First episode of major depressive disorder; full-threshold disorder with moderate to severe symptoms, neurocognitive deficits and functional decline (GAF = 30–50) Stage 3 (3a, 3b, 3c): Incomplete remission from first episode of care; primary and specialist care services; recurrence or relapse of depressive disorder, which stabilizes with treatment at a level of GAF, residual symptoms, or multiple relapses Stage 4: Severe, persistent or unremitting illness as judged on symptoms, neurocognition and disability criteria
Lieberman et al, 2001 [65]	Stage 1: Premorbid, mild physical anomalies, poor motor coordination, mild cognitive impairments, social deficits
Schizophrenia	Stage 2: Prodromal, nonspecific mood symptoms: anxiety, sadness, lability, irritability; sleep disturbances; cognitive impairment in attention, concentration; mild psychotic symptoms Stage 3: Onset/deteriorative, psychosis, cognitive impairment, negative symptoms, and social deficits Stage 4: Chronic/residual, negative symptoms, cognitive impairment, social deficits, and psychosis

GAF: Global functioning Source: author

Considering psychiatric disorders with genetic and environmental influences, heterogeneous population, with different proposed treatments and different prognoses, PTSD

appears to be an important example of where staging could provide benefit. The impact of traumatic stress exposure could be modeled as a continuous variable, with PTSD at the more severe end of the spectrum. The course of PTSD rests on the fact that post-trauma neurobiological alterations reach a point that results in persistent or progressive illness [69]. Previous trauma in adulthood and the occurrence of adverse situations in childhood are also associated with a worse prognosis of PTSD evolution. Previous findings support the hypothesis that exposure to intense stress during early life can have long-term effects on mental health via neurobiological effects and the development of allostatic load, increasing the risk of developing later psychopathology [70, 71].

McFarlane et al. [29] developed the first attempt at a staging model in PTSD, considering clinical presentation and neurobiological alterations at each stage. Details of the proposed staging are presented in Table 3.

Table 3. Staging model of PTSD.

Clinical aspects	**Possible neurobiological changes**
Stage 0: Trauma exposed asymptomatic but at risk	Down-regulation of glucocorticoid receptor sensitivity, increased amygdala reactivity, 5FKH genotype
Stage 1a: Undifferentiated symptoms of mild anxiety and distress	Inflammatory cytokine activation, decreasing response inhibition in the frontal cognitive systems
Stage 1b: Subsyndromal distress with some behavioral and functional decline	Increased physiological reactivity to trauma-related stimuli and startle response, prolonged autonomic arousal on provocation
Stage 2: First episode of full-threshold symptoms that has different trajectories	Early and potentially reversible neurobiological disinhibition of frontolimbic circuitry
Stage 3: Persistent symptoms that may fluctuate with ongoing impairment: • 3a Incomplete remission of first episode • 3b Recurrence or relapse of PTSD and persistent impairments • 3c Multiple relapses or worsening following incomplete treatment response	Decreased anterior cingulate and hippocampal volume, hypertension and metabolic syndrome
Stage 4: Severe unremitting illness of increasing chronicity with substantial disability	High allostatic load, high levels of inflammation, medical comorbidities, entrenched sensitization of a range of neurobiological systems

Source: extracted from McFarlane et al., 2017 [29]

According to this staging structure, Stage 0 includes those individuals who have not yet developed symptoms following trauma exposure but have accumulated and amplified risk of symptom development following further trauma exposure, representing a presymptomatic vulnerability phase. Stage 1 could be related to so-called "partial" or "sub-syndromic" PTSD, which consists of the development of a certain set of symptoms, but which does not reach the totality required by DSM-5 (Stein et al. [72] postulated that the effects of “partial PTSD” are prevalent and lead to personal injury and interference in life). The authors highlight the significance of separating a full acute disorder (Stage 2) from one in which there are persistent symptoms that are relatively long-standing (Stage 3). Stage 4 refers to a clinic presentation where symptoms last a prolonged period of time, increasing the probability of a severe, chronic and unremitting illness.

Although this staging model may allow a framework for examining different clinical and biological models for PTSD and how they overlap, developing a strategy for interventions at different stages of PTSD becomes a critical need. Regarding PTSD, there is much more to be done. Even if the staging model proposes stage-targeted treatments that may provide a better clinical outcome with fewer side effects, there are still differences among the patients of a particular stage [55]. Within the concept of ‘precision psychiatry’, personalized treatments could be offered according to these differences. In addition to integrating biographical, clinical and biological information regarding each individual, precision psychiatry involves the use of technological, data and computer science to improve diagnosis and treatment selection [73]. Figure 1 illustrates how the concepts of staging and precision psychiatry are integrated.

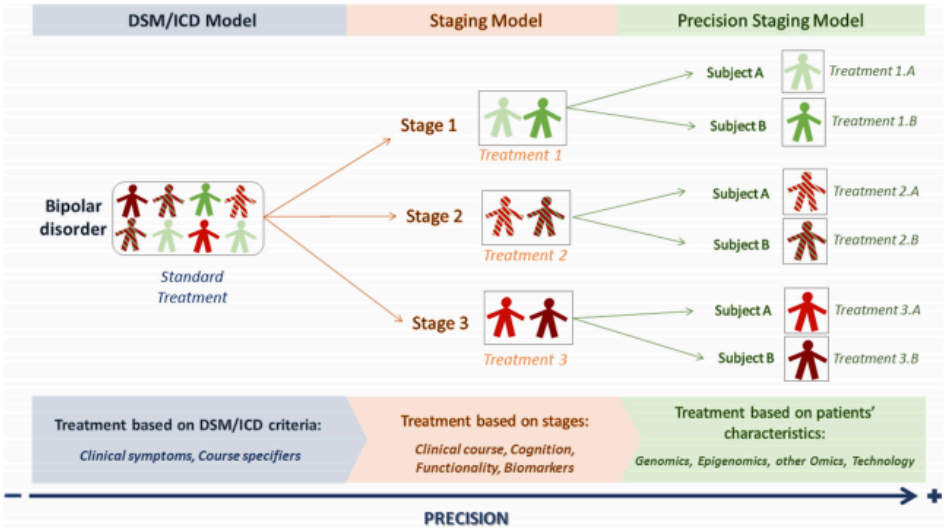


Figure 1. From cross-sectional approaches to precision staging model.

Source: Extracted from Salagre et al., 2018 [55]

In the last decades, several social phenomena have brought to light the problem of urban violence becoming an important factor for the genesis of PTSD. Many studies in the field of traumatic stress are being conducted to observe psychiatric comorbidities to PTSD (such as Major Depression and Substance Use Disorders), and the consequences of mass violence, natural disasters or war veterans [14, 74, 75]. In a psychiatric outpatient setting, a more heterogeneous population and a greater diversity of trauma may be found. In this way, a staging approach considering a heterogeneous environment is desirable in order to improve diagnosis, treatment selection and prognosis information. In order to fill this gap, the main goal of this research is to discover how clinical and individual aspects of patients could be related to a staging approach on PTSD using ML techniques.

2 JUSTIFICATION

The importance of research on the proposed theme starts with the high prevalence and morbidity of PTSD. It is a disorder characterized by significant social impairment and, when it evolves with chronic symptoms, presents an incipient response to treatment [6]. There is still a need to understand the different stages of the disorder, as well as to offer suitable treatments at each stage according to clinical and biological presentations.

ML has demonstrated significant potential to provide new insights into complex data. The field of ML for stress research is still in its infancy, but is already showing promising results in predicting PTSD-related health outcomes [39]. As far as we know, the initiative to investigate patient-related information associated with a widely used PTSD severity and diagnosis scale like CAPS-5, and relate them to a PTSD staging method is unprecedented.

We hope that a better understanding of the factors associated with PTSD staging can improve the diagnosis, provide more targeted treatments and better understand the progression of the disorder.

3 OBJECTIVES

3.1 OVERALL OBJECTIVE

To investigate personal, clinical and trauma-related characteristics of traumatic event patients related to PTSD staging through a machine learning approach.

3.2 SPECIFIC OBJECTIVES

To help in achieving the overall objective, the following specific objectives are presented:

a) To know the profile of patients who are victims of traumatic events and seek care in an outpatient clinic for emotional trauma;
b) To identify the ML models best suited to PTSD research;
c) To develop a ML model that could best predict PTSD staging from personal, clinical and trauma-related characteristics of patients.

4 METHODOLOGY

4.1 STUDY DESIGN AND POPULATION

The design chosen is a cross-sectional study. The patients included in this study were a subset of prospectively collected patients at two centers: the Psychological Trauma Research and Treatment Program (NET-Trauma), from the Psychiatric Service of the Hospital de Clínicas de Porto Alegre (HCPA), which belongs to the Universidade Federal do Rio Grande do Sul (UFRGS), and the Program for Research and Care on Violence and PTSD (PROVE), an outpatient service that belongs to the Department of Psychiatry at the Universidade Federal de São Paulo (UNIFESP). This study is approved by the Ethical Committees of UFRGS and UNIFESP under CAAE numbers 14644819.0.0000.5327 (UFRGS/HCPA) and CAAE 98548718.8.1001.5505 (PROVE/UNIFESP).

The exposure factor was the occurrence of a traumatic event up to six months before seeking outpatient care. The population of the study comprised all patients with a history of a previous traumatic event who sought care in the NET-Trauma and PROVE outpatient clinics.

4.2 STUDY FACTORS

Information associated with the traumatic event and psychiatric history, diagnostic criteria for PTSD (according to DSM-5), severity of symptoms (CAPS-5 score) and clinical evaluation. The study protocol is detailed in item 4.6 below.

4.3 OUTCOME

Clinical staging assigned in the patient evaluation and determined from the model developed by McFarlane et al. [29]. The staging structure is presented in Table 3.

4.4 LOGISTICS

The NET-Trauma unit is a community-based mental health service from HCPA, offering evidence-based treatments since 2003 for victims of traumatic events from the metropolitan region of Porto Alegre, a state capital in Southern Brazil. The PROVE center is

an outpatient service from the Psychiatry Department of the Escola Paulista de Medicina, UNIFESP, created in 2004. Patients can be referred to NET-Trauma or PROVE services by a public Basic Health Unit, emergency services or by their physicians/psychiatrists.

4.5 SAMPLE SELECTION AND INCLUSION CRITERIA

Two samples were collected: the first sample (Sample 1) consisted of patients enrolled in the screening process of the NET-Trauma outpatient service between August 2018 and January 2020. The second sample (Sample 2) consisted of patients enrolled in the screening process of the PROVE outpatient service between January 2016 and March 2019. All patients who sought care for a previous traumatic event, who were willing to accept the research terms and who met the inclusion criteria were included. In both samples, the participants spontaneously sought psychiatric assistance after experiencing a traumatic event and all agreed to participate in the study.

The exclusion criterion of the study was non-compliance with the inclusion criteria, which were:

- Age >16 years;
- Traumatic event up to 6 months;
- Ability to understand and sign the free informed consent form (Termo de Consentimento Livre e Esclarecido - TCLE);
- Ability to understand and respond to questionnaires.

4.6 STUDY PROTOCOL

The study protocol was composed by a questionnaire to collect information pertinent to the traumatic event, and the patients' socio-demographic and clinical information. It includes socio-demographic data (age, gender, ethnicity, educational level), information about trauma and psychiatric history, collected during first patient evaluation. Furthermore, this protocol contains two structured questionnaires: the Clinician-Administered PTSD Scale (CAPS-5) and the Life Events Checklist (LEC-5).

4.6.1 Clinician-Administered PTSD Scale – Brazilian version: this instrument is considered the gold standard in the diagnosis of PTSD. This scale was developed in 1989 by the National

Center of PTSD in the United States of America and first validated in 1995 [76]. The current scale [77] contains 30 items with symptoms grouped into diagnostic clusters according to DSM-5 structure and it is sensitive to diagnosis within the past month. This scale is administered by the clinician who applies a severity graduation between 0 (absent symptom) and 4 (extreme / incapacitating symptom) to each item. To be considered a symptom in the diagnosis (related to DSM-5 criteria), a given item should be evaluated with at least 2 points in severity (the so-called "SEV2 rule"). The authors suggest a minimum score between 23 and 26 points to consider the diagnosis in the English version [77]. The adaptation and validation of the CAPS-5 to Brazilian Portuguese was one of the efforts of this research, cooperating with a project from PROVE/UNIFESP (see Annex 1).

4.6.2. Life Events Checklist – Brazilian version: three formats of this instrument are available: standard, extended self-report, and interview (conducted by a clinician). The LEC-5 interview, used in this protocol, comprises 17 items and is designed to investigate exposure to potentially traumatic events. This instrument is often used in combination with other measures such as CAPS-5 for the purpose of establishing exposure to a traumatic event according to the PTSD Criterion A. The questions include life events such as natural disasters, physical or sexual aggression, severe injuries, and witnessing violent death. A first version was developed in DSM-IV [78] and adapted in DSM-5 *a posteriori* [79]. A cross-cultural adaptation for Brazilian Portuguese related to this last version was released in 2016 [80] (see Annex 2).

4.7 EXECUTION OF THE STUDY

The diagnosis of PTSD were assessed by the presence or absence of symptoms listed in DSM-5 according to clinical evaluation at the time of the first consultation. The diagnosis of comorbid psychiatric disorders will be performed by clinical evaluation according to criteria established in DSM-5. A list of traumatic events was accessed by the LEC-5 application; the severity of PTSD-related symptoms were evaluated by the CAPS-5 application.

During the first consultation, the following procedures will be performed:

- Application of the inclusion criteria;
- Anamnesis and mental status examination for evaluation and diagnostic confirmation;

- Signature of the free informed consent form;
- Application of the study protocol (including LEC-5 and CAPS-5).

To help achieve the specific objectives of this project, the following strategies will be applied (Table 4):

Table 4. Specific objectives and strategies of execution.

Specific objectives	**Strategies**
1) To know the profile of patients who are victims of traumatic events and seek care in an outpatient clinic of emotional trauma	Since a major part of PTSD research is made through specific populations (war veterans, survivors of great catastrophes or depressive patients), establishing a profile of PTSD patients is important. A paper (**article 1**) observing a 4-year follow-up of NET-Trauma patients (from 2014 to 2017) were developed to reach this objective, using available data from a previous NET-Trauma research project (CAAE 58511716.5.0000.5327). Regarding this specific objective, cooperation with PROVE/UNIFESP was established and generated two articles annexed to this thesis: an adaptation to Brazilian Portuguese of the CAPS-5 instrument (**additional article 1**). After that, a subset of data collected in NET-Trauma was shared with UNIFESP to perform validation of the instrument (**additional article 2**).
2) To identify the ML models best suited to PTSD research	This objective is related to knowing which ML techniques are best suited to perform the analysis of this study. To reach this objective, a systematic review of ML techniques applied to PTSD research was carried out (**article 2**).
3) To develop a ML model that could best predict PTSD staging from personal, clinical and trauma-related characteristics of patients	After identifying suitable ML techniques, they will be executed to find the best model that could predict the PTSD staging using both samples (**article 3**).

Source: author

4.8 STATISTICAL ANALYSIS

The study protocol was implemented and applied using the Redcap® environment supplied by HCPA. The statistical program for data analysis and construction of ML models was *R*, software version 3.6.3 (https://www.Rproject.org). For demographic information, the distribution of variables was described as mean and standard deviation or frequency and

proportion, where applicable. For the comparison between continuous variables, the Mann-Whitney U tests and Student's *t*-test were used. For categorical variables, Pearson's chi-square test or the Fisher's exact test were used. A significance level of 5% ($p \leq 0.05$) with a confidence level of 95% were considered for all statistical tests.

4.9 DATA PROCESSING

In this section, the data processing for the execution of ML is described according to the steps presented in Table 5 below.

Table 5. Strategies on Data Processing.

Data processing	**Strategies**
Treatment of quantitative and categorical variables	When necessary, quantitative variables were converted to *z-scores* based on the mean and standard deviation of the sample. The conversion of original values into *z-scores* was performed when it is important that the quantitative variables should be comparable to each other. Categorization works by selecting one or more cut-off points such that the values between them will represent a certain category. Categorical variables were transformed into a set of binary variables using one-hot encoding.
Missing data	Most ML learning techniques need complete data to perform the analysis, requiring the researcher to decide what to do to deal with missing data. The commonest treatment for this problem is not to consider cases that contain missing values. Another possible strategy is to eliminate variables with missing data (reducing dimensionality). In this study, no missing data were found.
Bias and asymmetric variables	A model with high bias is an "underfit" and a model with high variance is an "overfit". Data bias is addressed using methods to avoid overfitting training data, according to the ML technique used. The distribution of quantitative variables could be analyzed through basic statistics (mean, median, fashion, quartiles, standard deviation).
Standardization of variables	Variables with different scales of values can disrupt the learning process. Some algorithms require input values in the range of -1 to 1 or 0 to 1. Before running a learning process, it is important to evaluate where standardization is necessary. Since most of the quantitative data comes from a standardized instrument, this step was not needed.

Source: adapted from Baştanlar & Ozuysal, 2014 [81] and Olivera, 2014 [82].

4.10 MACHINE LEARNING APPROACH

As a result of the review of ML techniques described in specific objective 3 (see Table 4), the models suitable to predict the clinical staging of PTSD from the personal, sociodemographic and trauma-related variables were chosen. We developed four predictive models for comparison: (1) the 20-item CAPS-5; (2) the 15-item CAPS-5; (3) the 20-item CAPS-5 plus clinical and demographic data; and (4) the 15-item CAPS-5 plus clinical and demographic data. For each group, all ML techniques were applied.

Performance metrics as area under the receiver operating characteristic curve (AUC), sensitivity, specificity, accuracy, F1 score and Cohen's Kappa score. Cohen's Kappa statistics are widely used in multiclass classification to inform how well the classifier is performing compared with a random classifier: values less than 0 indicate no agreement; 0–0.20 as slight, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1 as almost perfect agreement [83].

According to the choice of ML models, the following strategies was carried out to ensure the quality of the study (Table 6):

Table 6. Quality issues of ML techniques.

Issue	**Strategies**
AUC score average for multiclass approaches	The AUC score normally applies to binary problems. For multiclass problems, there are two possible ways to obtain 'average' scores: (1) using macro-averaging, which reduces the multiclass predictions down to multiple sets of binary predictions. It calculates the corresponding metric for each of the binary cases and then averages the results together (macro-averaging reduces the problem to multiple one-vs-all comparisons); and (2) using micro-averaging, which treats the entire set of data as an aggregate result and calculates a single metric rather than k metrics that get averaged together (by calculating all of the true positive results for each class and using that as the numerator, and then calculating all of the true positive and false positive results for each class, and using that as the denominator). In this case, rather than each class having equal weight, each observation gets equal weight. This gives the classes with the most observations more power. Both methods must be compared to choose the most reliable for the imbalanced multiclass problem.
Dimensionality reduction: feature extraction and feature	Dimensionality reduction is one of the major tasks in the analysis of multidimensional data, aiming to decrease the number of dimensions (variables) /features. In feature selection, *k* of the *v* variables (features) that

selection	give the most information is selected and the other dimensions are eliminated, aiming to select the subset of variables capable of producing the best possible prediction model. In feature extraction, the original *v* variables are transformed to a new set of *k* variables. A popular technique of feature selection is the principal component analysis (PCA), where the data is analyzed and grouped with the most informative components (dimensions).
Randomization and cross-validation	Randomization is required to ensure that the result of the learning process is independent of the selection of training data. If the dataset is large enough, it is possible to randomly divide it into K parts, and then randomly divide each part into two as the training and validation sets. A popular strategy when a data set is not large enough is the *cross-validation* approach, which consists in splitting the data set into K−1 parts to be used as a training set and the remaining part is used as a validation set. This procedure is repeated for all K possible choices. In this study, a nested cross-validation strategy was employed to obtain an unbiased estimation of the true generalization performance, to avoid data leakage, and to provide robust parameter estimates particularly for smaller samples.

Source: extracted from Olivera, 2014 [82]

Data pre-processing and ML techniques were implemented using the *caret* package version 6.0-85 from software R. Details of each strategy and how they were implemented after data collection is described in article 3, supplementary material, in the Appendix section.

4.11 ETHICAL ASPECTS

This study protocol and logistics followed the conditions established in Resolution 466/12 of the Brazilian National Health Council (CNS). The confidentiality of the information collected from the research individuals was ensured through adequate training of the team involved in the data collection, as well as the ethical commitment of the researchers participating in the research in the handling and treatment of the information. All information collected is archived in an appropriate place and not used for any purpose other than the proposed objectives.

Participation in this study was voluntary and participants could withdraw from the research process at any time. All participants were informed of the research procedures as well as the guarantee of confidentiality of the data in any report relating to this research or any document that may result from it. All information will be provided at the end of the free

informed consent term that should be signed by all research subjects before any procedures related to this protocol are performed. The NET-Trauma outpatient center provides a consent term that is related to a follow-up project; as this study is subordinated to this major project, the consent term provided to the patients was the same from the NET-Trauma project (see Annex 3). The author of this project and his advisors signed a data confidentiality term.

5 RESULTS

5.1 ARTICLE 1

SHORT COMMUNICATION: TRAUMA-RELATED DISORDERS IN A LOW- TO MIDDLE-INCOME COUNTRY: A FOUR-YEAR FOLLOW-UP OF OUTPATIENT TRAUMA IN BRAZIL.

Luis Francisco Ramos-Lima [1],
Paulo Ricardo Assis de Souza [2],
Stefania Pigatto Teche [3],
Lucia Helena Machado Freitas [3]

[1] Post-graduate Program in Psychiatry and Behavioral Sciences, Federal University at Rio Grande do Sul, Porto Alegre, Brazil.
[2] Faculty of Medicine (FAMED), Federal University at Rio Grande do Sul, Porto Alegre, Brazil.
[3] Post-graduate Program in Psychiatry and Behavioral Sciences, Federal University at Rio Grande do Sul, Porto Alegre, Brazil; Psychological Trauma Research and Treatment Program (NET-Trauma), Clinical Hospital of Porto Alegre, Porto Alegre, Brazil.

Psychiatry Res. 2019;280:112525. doi:10.1016/j.psychres.2019.112525

MAIN PURPOSES:

- To investigate clinical, trauma-related characteristics of NET-Trauma population;
- To provide knowledge about heterogeneous clinic outpatient setting;
- To evaluate possible features to use in the machine learning models (article 3).

1. Introduction

The World Health Organisation (WHO) World Mental Health Survey, which included seven low- to middle-income countries (LMICs), found that the prevalence of a traumatic event in a lifetime was 70.4% (Benjet et al., 2015). Social phenomena have highlighted urban violence as an important factor for the genesis of acute stress disorder (ASD) and post-traumatic stress disorder (PTSD). PTSD is considered a debilitating condition that is a result of exposure to traumatic events including war, mass violence and natural disasters.

In the Brazilian context, PTSD has become an important public health problem. A population-based study in two major urban cities in Brazil found that the prevalence of PTSD during a lifetime was 11.4%–14.7% in women and 4.7%–7.8% in men (Ribeiro et al., 2013). While several studies about PTSD are related to specific types of trauma (Schnyder et al., 2017; Silva et al., 2013; Tay et al., 2016), only a few studies were developed from heterogeneous mental health services.

We remark that more information is needed about PTSD in LMICs. A small number of PTSD studies are performed in LMICs (Schnyder et al., 2017), and a more heterogeneous population may be found in a psychiatric outpatient setting. The present study aims to estimate the prevalence and follow-up of trauma-related disorders in patients who seek assistance from an open access psychiatric ward at a public university hospital in Southern Brazil.

2. Methods

This study is an evaluation of data collected through routine clinical care of patients who attended the Psychological Trauma Research and Treatment Programme (*Núcleo de Estudos e Tratamento do Trauma–NET-Trauma*) from the Clinical Hospital of Porto Alegre (HCPA), during the period of 2014 to 2017.

The study protocol included sociodemographic data, information about trauma and psychiatric history, and follow-up information. The diagnosis of ASD and PTSD was assessed by the presence of symptoms listed in the DSM-5 according to clinical evaluation at first consultation. The Ethical Committee of HCPA approved all phases of the study.

Remission was attributed according to clinical observation of the symptoms (based on DSM-5) at discharge and considered as a perceived improvement in greater than 50% of the symptoms compared with the initial evaluation. The proposed treatment was based on current

guidelines (US Department of Veterans Affairs, 2010; National Institute for Health and Care Excellence, 2013) using pharmacological and/or psychotherapeutic approaches according to clinical evaluation.

3. Results

During the period from 1 January 2014, to December 31, 2017, 251 patients were referred to the outpatient clinic; 179 attended the first consultation and agreed to participate in the study. Fifty-six patients (31.3%) were diagnosed with PTSD and 17 (9.5%) with ASD. Female patients had a higher prevalence of ASD/PTSD (61.6% vs 38.4%, $p = 0.048$). The main profile was composed of patients of caucasian ethnicity, married and who had 10 to 12 years of schooling.

Observed patients experienced the traumatic events firsthand (89.4%) mostly by an unidentified aggressor (72.6% of the sample). The presence of any previous trauma in adult life appears to be related to the development of ASD/PTSD (14.2% non-PTSD vs 35.6% ASD/PTSD, $p = 0.001$). The presence of childhood trauma/stress in individuals who developed ASD/PTSD in adult life compared with those who did not suffer these conditions (11.3% non-PTSD vs 23.3% ASD/PTSD, $p = 0.033$) also presented a statistical difference. The presence of current or previous psychiatric disorders and a family history of a psychiatric disorder did not appear to be statistically related to ASD/PTSD in our observation.

The most prevalent traumatic event in our results was assault (27.4% vs 42.5% ASD/PTSD; $p = 0.035$). In contrast, accidents appeared to be more prevalent in patients who did not develop PTSD (13.2% vs 4.1% ASD/PTSD; $p = 0.045$). A higher prevalence of events related to sexual violence in the group without the development of ASD/PTSD (21.7% vs 15.5% ASD/PTSD; $p = 0.266$) was found; however, this difference was not statistically significant. Other life-threatening events, physical violence and natural disasters did not show significant differences between groups.

Table 1 presents indicators of follow-up from all 179 patients. Patients who were diagnosed with ASD/PTSD had twice the number of consultations (median 9 vs 4.5 visits ASD/PTSD; $p < 0.001$). The mean follow-up time was approximately 68 days for patients without the diagnosis of ASD/PTSD, whereas that of those who fulfilled the criteria was approximately 143 days ($p < 0.001$). Patients in the ASD/PTSD group who completed treatment with remission stayed longer at the outpatient clinic: the mean follow-up time rose to approximately 205 days and 80% were discharged after 3 months of follow-up.

Table 1: Indicators of patient follow-up.

Variables	**Non-PTSD (n=106)**	**ASD/PTSD (n=73)**	**All**	***P* value**
Number of consultations (median, IQR)	4.5 (9)	9 (11.5)	6 (10)	[a] <0.001
Follow-up time of total sample (mean in days, SD)	68.5 (7.8)	143.5 (16.1)	99.1 (8.5)	[a]<0.001
Follow-up time of discharged patients (mean in days, SD)	107.4 (11.3) (n=56)	205.6 (23.5) (n=35)	145.1 (12.4)	[a]0.004
End of follow-up				
Abandonment	25 (23.6%)	33 (45.2%)	58 (32.4%)	[a]0.003
Discharge with remission	56 (52.9%)	35 (48.0%)	91 (50.8%)	
Referral	25 (23.6%)	5 (6.8%)	30 (16.8%)	

[a] *p<0.05*; SD: standard deviation; IQR: interquartile range

4. Discussion

The present study shows a prevalence of 40.8% of ASD/PTSD. In order to compare a group of individuals who developed more symptoms with those who were exposed to trauma but did not develop these disorders fully, this research considered the occurrence of ASD and PTSD as instances of a trauma-related disorders spectrum. A Brazilian study found a prevalence of PTSD of around 8%–10% in a lifetime (Ribeiro et al., 2013); the higher prevalence found in our study may be a result of the fact that it was measured in a specialised psychiatric trauma clinic, which receives patients with more comorbidities and previous vulnerability. Mental disorders have different aetiologies, are more common in women and particularly affect individuals with accumulating social and family disadvantages, such as LMICs population (World Health Organisation, 2014). A greater prevalence of 61.6% of ASD/PTSD was found on observing female patients. We hypothesised that in a country such as Brazil, considered a LMIC, individuals with lower levels of education share some difficulties regarding access to health services. These include indirect costs to household (transport cost), information on health care providers and cultural beliefs (Jacobs et al., 2012).

A greater occurrence of previous trauma in adulthood or childhood was reported in patients who developed ASD/PTSD. Previous findings support the hypothesis that exposure to intense stress during early life may have a negative cumulative effect on mental health

(Ribeiro et al., 2013; Kostaras et al., 2016) and may increase the risk of developing a subsequent psychopathology (Pupo et al., 2015; Cordero et al., 2017).

In a specific study for victims of urban violence, assaults were responsible for 25% of the cases of PTSD (Pupo et al., 2015); in our results, assaults accounted for 42.5% of the cases of ASD/PTSD and were related to the diagnosis ($p = 0.035$). Regarding the higher prevalence of events related to sexual violence in the group without the development of ASD/PTSD, we postulate whether the 'natural' course of symptoms is in part responsible for remission by adaptation (Rothbaum et al., 1992) or if the influence of immediate care of these patients prevents ASD progression – more research is necessary to confirm these hypothesis.

The number of consultations and the mean follow-up time of patients who developed ASD/PTSD were double compared with that of patients who did not receive the diagnosis, possibly reflecting the complexity of treatment of these patients. A previous meta-analysis found a mean time of spontaneous remission of approximately 40 months for PTSD patients (Morina et al., 2014); our data indicate a shorter discharge when remission is achieved, in an average of 205.6 days (approximately 7 months) of follow-up. It suggests that the intervention may shorten the disorder and promote faster recovery of these patients.

Providing information from a naturalistic setting was considered a strength of this study. Limitations included that the intensity of the symptoms during follow-up was not evaluated. This study was conducted in a large urban centre of the country in the only outpatient public clinic that specialises in psychiatric trauma; the results may not be applied to other areas where social and cultural conditions differ.

This study addressed an important gap in trauma research in LMIC by examining factors associated with characteristics of trauma. As a recommendation, areas involved in receiving these patients should be familiar with psychological first aid techniques in order to guarantee initial patient care, to ensure safety and comfort, to offer an empathic approach and to help activate social support (Brymer et al., 2006). Additional effort is needed to educate the population on trauma-related disorders, to ensure access to health care and to offer adequate treatment.

Acknowledgements: This study was supported in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Brazil, Finance Code 001) and by the Conselho Nacional de Pesquisa (CNPq, Brazil).

Conflicts of interest: The authors declare that they have no conflicts of interest.

References

American Psychiatric Association, 2013. Diagnostic and statistic manual of mental disorders. 5th. ed. Washington (DC): American Psychiatric Publishing.

Benjet, C., Bromet, E., Karam, E.G., Kessler, R.C., McLaughlin, K.A., Ruscio, A.M., Shahly, V., Stein, D.J., Petukhova, M., Hill, E., Alonso, J., Atwoli, L., Bunting, B., Bruffaerts, R., Caldas-de-Almeida, J.M., de Girolamo, G., Florescu, S., Gureje, O., Huang, Y., Lepine, J.P., Kawakami, N., Kovess-Masfety, V., Medina-Mora, M.E., Navarro-Mateu, F., Piazza, M., Posada-Villa, J., Scott, K.M., Shalev, A., Slade, T., ten Have, M., Torres, Y., Viana, M.C., Zarkov, Z., Koenen, K.C., 2016. The epidemiology of traumatic event exposure worldwide: results from the World Mental Health Survey Consortium. Psychol Med. 46(2):327-43. doi: 10.1017/S0033291715001981.

Brymer, M., Layne, C., Jacobs, A., Pynoos, R., Ruzek, J., Steinberg, A., Vernberg, E., Watson, P., 2006. Psychological first aid: Field operations guide (2nd ed.). Los Angeles: National Child Traumatic Stress Network & National Center for PTSD.

Cordero, M.I., Moser, D.A., Manini, A., Suardi, F., Sancho-Rossignol, A., Torrisi, R., Rossier, M.F., Ansermet, F., Dayer, A.G., Rusconi-Serpa, S., Schechter, D.S., 2017. Effects of interpersonal violence-related post-traumatic stress disorder (PTSD) on mother and child diurnal cortisol rhythm and cortisol reactivity to a laboratory stressor involving separation. Horm Behav, 90:15–24. doi: 10.1016/j.yhbeh.2017.02.007.

Jacobs, B., Ir, P., Bigdeli, M., Annear, P.L., Van Damme, W., 2012. Addressing access barriers to health services: an analytical framework for selecting appropriate interventions in low-income Asian countries. Health Policy Plan. 27(4):288-300. doi: 10.1093/heapol/czr038.

Kostaras, P., Bergiannaki, J.D., Psarros, C., Ploumbidis, D., Papageorgiou, C., 2017. Posttraumatic stress disorder in outpatients with depression: Still a missed diagnosis. J Trauma Dissociation. 18(2):233-247. doi: 10.1080/15299732.2016.1237402.

Morina, N., Wicherts, J.M., Lobbrecht, J., Priebe, S., 2014. Remission from post-traumatic stress disorder in adults: a systematic review and meta-analysis of long term outcome studies. Clin Psychol Rev. 34(3):249-55. doi: 10.1016/j.cpr.2014.03.002.

National Institute for Health and Care Excellence, 2013. Post-traumatic stress disorder. Evidence Update December 2013. https://arms.evidence.nhs.uk/resources/hub/1031525/attachment. (Accessed 01 May 2019)

Olff, M., Vermetten, E., 2013. Psychotrauma research in the Netherlands. Eur J Psychotraumatol. 4. doi: 10.3402/ejpt.v4i0.20873.

Pupo, M.C., Serafim, P.M., de Mello, M.F., 2015. Health-related quality of life in posttraumatic stress disorder: 4 years follow-up study of individuals exposed to urban violence. Psychiatry Res. 228(3):741-5. doi: 10.1016/j.psychres.2015.05.030.

Rothbaum, B.O., Foa, E.B., Riggs, D.S., Murdock, T., Walsh, W., 1992. A prospective examination of post-traumatic stress disorder in rape victims. J Trauma Stress. 5(3), 455-475. doi: 10.1002/jts.2490050309.

Schnyder, U., Schäfer, I., Aakvaag, H.F., Ajdukovic, D., Bakker, A., Bisson, J.I., Brewer, D., Cloitre, M., Dyb, G.A., Frewen, P., Lanza, J., Le Brocque, R., Lueger-Schuster, B., Mwiti, G.K., Oe, M., Rosner, R., Schellong, J., Shigemura, J., Wu, K., Olff, M., 2017. The global collaboration on traumatic stress. Eur J Psychotraumatol. 8(sup7):1403257. doi: 10.1080/20008198.2017.1403257.

Silva, T.L.G., Mello, P.G., Silveira, K.A.L., Wolffenbüttel, L., Lobo, B.O.M., Bicca, C.H.M., Grassi-Oliveira, R., Kristensen, C.H., 2013. Primeiros Socorros Psicológicos: relato de intervenção em crise em Santa Maria. Rev Bras Psicoter. 15(1):93-104 (Portuguese)

Tay, A.K., Rees, S., Steel, Z., Tam, N., Soares, Z., Soares, C., Silove, D.M., 2016. Six-year trajectories of post-traumatic stress and severe psychological distress symptoms and associations with timing of trauma exposure, ongoing adversity and sense of injustice: a latent transition analysis of a community cohort in conflict-affected Timor-Leste. BMJ Open. 6(2):e010205. doi: 10.1136/bmjopen-2015-010205.

US Department of Veterans Affairs, 2010. VA/DoD Clinical Practice Guideline for the Management of posttraumatic stress disorder and acute stress disorder. https://www.healthquality.va.gov/PTSD-Full-2010c.pdf (Accessed 01 May 2019)

World Health Organisation, Calouste Gulbenkian Foundation, 2014. Social determinants of mental health. Geneva: World Health Organisation.

5.2 ARTICLE 2

THE USE OF MACHINE LEARNING TECHNIQUES IN TRAUMA-RELATED DISORDERS: A SYSTEMATIC REVIEW

Luis Francisco Ramos-Lima [1],
Vitoria Waikamp [2],
Thyago Antonelli-Salgado [3],
Ives Cavalcante Passos [4],
Lucia Helena Machado Freitas [2]

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

J Psychiatr Res. 2020;121:159-172. doi:10.1016/j.jpsychires.2019.12.001

MAIN PURPOSES:

- To investigate how machine learning techniques are being applied to PTSD research
- To determine the most used machine learning algorithms in PTSD research to be used to construct staging model (article 3)

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, 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 at., 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, Super 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. Figure 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; Karstoft et al, 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; Dabek and Caban, 2015b; Vergyri et al., 2015; Galatzer-Levy et al., 2014). Eight articles used machine learning techniques to build diagnostic classification tools (Table 2): five of these used neuroimaging and DTI studies (Salminen et al., 2019; Rangaprakash et al., 2017; Zhang et al., 2016; Liu et al., 2015; Gong et al., 2014); two studies used neuropsychological tests (Breen et al., 2019; Omurca and Ekinci, 2015); and two utilized audio records (Banerjee et al., 2017; van den Broek et al., 2013). Another seven studies were found (Table 3): three of these used clustering analysis to identify subtypes of symptomatology (Grisanzio et al., 2018; Ma et al., 2016;

Galovski et al., 2016) and four used network analysis (Bartels et al., 2019; Fried et al., 2018; Sullivan et al., 2018; Mitchell et al., 2017) to find relationships between different symptoms.

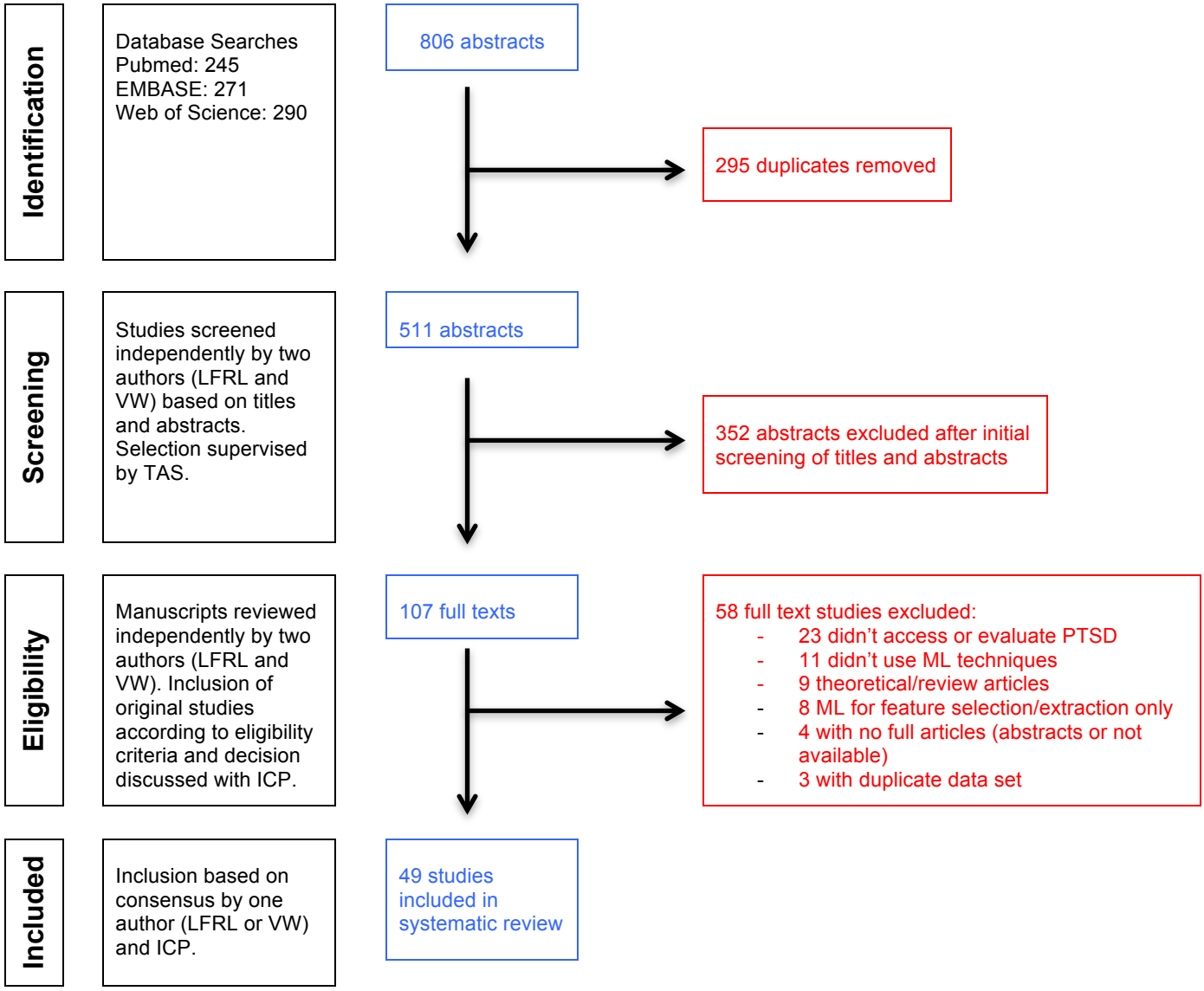


Figure 1. Flowchart of review process and study selection

Table 1. Prognostic studies using machine learning techniques in PTSD

First author, year	Data utilized	Diagnostic PTSD tools used	Sample size and diagnosis	ML model	Accuracy	Other measures	Event	Commentary
PROGNOSTIC STUDIES USING NEUROIMAGING OR ELECTROENCEPHALOGRAPHIC STUDIES								
Zandvakili et al., 2019	Resting-state EEG using Alpha, Beta, Theta and Delta frequency bands pre- and post-TMS	PCL-5	35 PTSD+MDD subjects (post-TMS EEG in 30 subjects)	SVM, LASSO	75.4% (alpha) 77.4% (beta) 73.8% (theta) 78.6% (delta)	AUC 0.71 [95% CI=0.54-0.87] (PTSD) AUC 0.83 [95% CI=0.69-0.94] (MDD)	War veterans	Prediction of non-response to TMS with high specificity, and identified pre- and post-TMS status using EEG coherence
Nicholson et al., 2018	Resting-state fMRI using mALFF and Amygdala maps	CAPS-IV, CAPS 5, SCID	181 subject - 81 PTSD - 49 PTSD+DS - 51 HC (age-matched)	Multiclass GPC	mALFF: - 89.02% (PTSD) - 89.80% (PTSD+DS) - 96.08% (HC) Amygdala maps: - 85.37% (PTSD) - 83.67% (PTSD+DS) - 93.48% (HC)	PPV (mALFF): - 94.81% (PTSD) - 89.90% (PTSD+DS) - 87.50% (HC) PPV (amygdala maps): - 85.37% (PTSD) - 75.93% (PTSD+DS) - 93.48% (HC)	Heterogeneous	PTSD+DS group was characterized by higher MDD comorbidity and increased PTSD symptoms
Yuan et al., 2018	Resting-state fMRI using mALFF and degree centrality (DC)	CAPS, SCID	22 subjects (PTSD using paroxetine, treatment lasted three months)	SVM	72.5% pre-treatment	AUC 0.72	Earthquake survivors	mALFF+DC data could predict the long-term clinical outcome of PTSD
Im et al., 2017	sMRI, dMRI with local and network features	CAPS-IV, SCID	59 subjects - 30 TC - 29 HC	LR	AUC 0.73 [95% CI=0.54-0.91]		Subway fire disaster	Follow-up of 5 years

					(six brain regions in 1.3 years)			
Jin et al., 2017	Resting-state fMRI	CAPS-IV, SCID [1]	149 subjects - 73 PTSD - 76 TC	RCE-SVM	94.2% (Static FC) 90.9% (Dynamic FC)		Earthquake survivors	Individuals with PTSD are characterized by reduced temporal variability of brain connectivity
Li et al., 2016	DTI	CAPS	65 subjects - 43 mTBI - 22 HC	Bayesian discrimination analysis	75.56%	Sensitivity 73%, Specificity 78%	mTBI victims on ER	DTI provide a potential indicator at subacute stage for PTSD
Wang et al., 2016	DTI	CAPS	33 subjects - 17 PTSD - 16 TC (matched)	SVM (3 kernels)	79.86% (RBF) 68.32% (Polynomial) 63.45% (Linear)	AUC 0.816 (RBF) 0.704 (Polynomial) 0.611 (Linear)	Traffic collision survivors	Early prediction of PTSD patients survived from traffic collision
Cisler et al., 2015	fMRI	SCID, PCL-C	16 subjects (PTSD, women)	SVM	76%	Sensitivity: 74% Specificity: 76%	Physical and sexual violence victims	Differ among trauma related and neutral related memories
PROGNOSTIC STUDIES USING QUESTIONNAIRES								
Leightley et al., 2019	Questionnaires	PCL-C	13,690 subjects (data from 2004 to 2009)	SVM, RF, ANN, Bagging	91% (SVM) 97% (RF) 89% (ANN) 95% (Bagging)	Sensivity 70% (SVM), 60% (RF), 61% (ANN), 69% (Bagging) Specificity 92% (SVM), 98% (RF), 92% (ANN), 96% (Bagging)	Military personnel	Variables contributed to the performance: alcohol misuse, gender and deployment status
Rosellini et al., 2018a	Questionnaires with 67 risk factors	DTS	23,907 subjects	LR, Elastic Net, LASSO,	AUC 0.7904		Earthquake survivors	Use of 36 algorithms and a

	of PTSD			RF, Bayesian addictive trees, SVM (3 types), Ridge, Adaptive splines, SuperLearner	(SuperLearner x LR) [95% CI 0.7827-0.7981]			super learner algorithm to predict post-earthquake PTSD
Augsburger and Elbert, 2017	Questionnaires and computer-based tasks	PSS-I	56 subjects (70% separate training set)	Stochastic GBM		R2 0.53	Worldwide refugees	Risk behaviours: exposure to torture and war
Conrad et al., 2017	Diagnostic interviews by trained local counsellors	PDS	652 subjects - 441 TE (training sample) (n=211 new testing dataset)	RF (using conditional inference), LASSO, LR	77.25% RF 74.88% LASSO 75.36% LR (traumatic load)	Sensitivity: - 98.01% RF - 94.04% LASSO - 94.04% LR Specificity: - 25.00% RF - 26.27% LASSO - 28.33% LR	Survivors of uganda rebel war	Prediction accuracy of PTSD risk was slightly improved by ML and is accompanied with expenses in time and calculation effort
Gradus et al., 2017	Mail survey	PCL-M	2061 subjects (1062 men, 1099 women)	CTA, RF	AUC 0.91 (male) 0.92 (female)		Veterans of Iraq and Afghanistan wars	Prediction of suicide ideation
He et al., 2017	Online survey with open questions	SCID, CAPS-IV	300 subjects - 150 PTSD - 150 TC	Decision trees, NB, SVM, PSM	82% (using unigrams on PSM)	Sensitivity 95% (using trigrams on SVM) Specificity 81% (using unigrams+bigrams on PSM)	Heterogeneous	Textmining method, PSM with unigrams attained the highest prediction accuracy
Reece et al., 2017	Questionnaires and posts from Twitter	TSQ	174 subjects - 63 PTSD - 111 HC (243.775 posts, 70% training,	RF	AUC 0.89 (PTSD)		Twitter users	Two groups to detect PTSD and depression

			validated on another sample)					
Schalinski et al., 2016	Questionnaires	Clinical (ICD-10), PSS-I	129 subjects (75% training set)	Conditioned RF		R2 0.58 [95%CI=0.44-0.68] (adverse childhood experiences to PTSD)	Recruited at the local center for Psychiatry	Physical neglect at the age of 5 was most pronounced for PTSD symptoms in adults
Karstoft et al., 2015a	Telephone-based interviews	PSS-I	957 subjects (TE)	SVM	AUC 0.75 [95%CI=0.67-0.80]		ER	Limited predictive power of data features available, all collected within ten days of a traumatic event
Karstoft et al., 2015b	Questionnaires	PCL-C	561 subjects (95% male)	SVM	AUC 0.84 [95%CI=0.81-0.87] (pre-deployment) 0.88 [95%CI=0.85-0.91] (post-deployment)		Soldiers from Afghanistan	Potential for pre- and early post-deployment prediction of resilience or PTSD
Kobach et al., 2015	Interviews conducted in demobilization camp	PSS-I	367 subjects (male)	RF (using conditional inference)	R2 0.30		Ex-combatents	Number of lifetime experienced traumatic events is the main predictor for PTSD
Kessler et al., 2014	Face-to-face interviews	SCID	47,466 subjects (TE)	RF, SuperLearner, Ridge, LR, LASSO, Elastic Net	AUC - 0.96 (SuperLearner) - 0.90 (RF)		Epidemiological data from WHO	High conditional PTSD risks associated with rape and sexual assault
Marinić et al.,	Questionnaires	CAPS,	102 subjects	RF, SVM	80.39%	Sensitivity and	War veterans	Group of comorbid

2007		SCID	(male) - 51 PTSD, - 51 controls (other psychiatric conditions)		(CAPS+PANSS)	specificity 80%		diagnoses surfaced as important
PROGNOSTIC STUDIES USING BIOLOGICAL SAMPLES								
Galatzer-Levy et al., 2017	Questionnaires, blood samples	CAPS-IV	152 subjects	LGMM, SVM	AUC 0.82 [95% CI=0.80–0.85]		Heterogeneous (from ER)	Combined clinical, neuroendocrine, psychophysiological and demographic information to predict pathways of non-remitting PTSD
Hemmings et al., 2017	Microbial DNA from stool samples and questionnaires	CAPS-5	30 subjects - 18 PTSD - 12 TC	RF		r = −0.387 P = 0.035	Heterogeneous	Decreased total abundance of bacterial taxa was associated with higher CAPS scores
Tylee et al., 2015	Sample peripheral blood and questionnaires	CAPS, PCL	50 subjects - 25 PTSD - 25 TC (n=40 training set)	SVM	90% (exon-based) 80% (gene based)	Sensitivity: - Exon-based 100%, - Gene-based 80% Specificity: - Exon-based 80% - Gene-based 80%	US Marines from Iraq or Afghanistan	Diverse group of genes and exons differentially expressed
PROGNOSTIC STUDIES USING AUDIO AND MEDICAL RECORDS								
Harrington et al., 2019	Medical records	ICD-9 and 10, PCL (DSM-IV), PC-PTSD	500 subjects - 198 "likely" PTSD - 84 possible	LASSO	AUC 0.95	Sensivity 99% Specificity 99%	War veterans	Diagnosis were made by chart review

			PTSD - 203 "likely not" PTSD (20% testing dataset)					
Marmar et al., 2019	Audio recordings of clinical interviews	CAPS	129 subjects - 52 PTSD - 77 TC	RF	89.1%	AUC 0.954	War veterans	Probability of PTSD was higher for markers that indicated slower, more monotonous speech, less change in tonality, and less activation
Papini et al., 2018	Medical records, questionnaires	PC-PTSD	271 subjects - 110 PTSD (at three months) - 231 TC	Decision trees	78%	AUC 0.85 [0.83-0.86] Sensitivity 69% Specificity 83%	Emergency room patients	Used localization variables (such as zip codes) and health insurance coverage rates on ML algorithm
Saxe et al., 2017	Hospitalization data	UCLA PTSD reaction index score	163 children TE (80% training set)	SVM, LASSO, RF	AUC 0.79		Children hospitalized with injuries	Prediction model to PTSD in children
Wortwein and Scherer, 2017	Questionnaires, visual and acoustic features	PCL-C	198 subjects - 69 PTSD - 129 TC	SVM	F1 score 0.748		Veterans and non-veterans	A subset of 5 questions are required to detect symptoms of PTSD by a Virtual Human Machine
Dabek et al., 2015a	Clinical recordings	Clinical (ICD-10)[1]	89,840 subjects - 6,629 on test set	Neural network + SVM	82.35% (overall) 83.82% (PTSD)		Post-concussion patients (mTBI)	Neural network model to predict psychological conditions including PTSD

Dabek et al., 2015b	Clinical recordings	Clinical (ICD-10)[1]	89,840 subjects (training set), 16,045 subjects (testing set)	SVM	85%	AUC 86.52%	Post-concussion patients (mTBI)	Predicts PTSD within the first year following the injury
Vergyri et al., 2015	Audio recordings, questionnaires	CAPS	39 subjects - 15 PTSD - 24 TC	GB, decision tree, neural network classifiers, AdaBoost	77%		Veterans from Iraq and Afghanistan	Speech features have discriminative power for the prediction of PTSD
Galatzer-Levy et al., 2014	Medical records and telephone interviews	PSS-I	957 subjects (TE)	SVM, RF, AdaBoost, KRR-RBF, BBR	AUC 0.82 (linear SVM) [95% CI=0.78–0.86]	From ASD symptoms: AUC 0.60	ER	Forecasting non-remitting PTSD within 10 days of a traumatic event

[1] Informed by e-mail sent to the authors

AUC: area under the curve; CI: confidence interval; DS: dissociative subtype; dMRI: diffusion-weighted magnetic resonance imaging; DSM: diagnostic and statistical manual of mental disorders; ER: emergency room; FC: functional connectivity; fMRI: functional magnetic resonance imaging; GM : grey matter; HC: healthy controls; ICD: international classification of diseases; mALFF: mean amplitude of low-frequency fluctuations; mTBI: mild traumatic brain injury; sMRI: structural magnetic resonance imaging; TC: trauma-exposed controls; TE: trauma-exposed individuals; TMS: transmagnetic stimulation; WM: white matter

Instruments: CAPS: Clinician-administered PTSD Scale; CATS: Child and Adolescent Trauma Score; DTS: Davidson Trauma Scale; HTQ: Harvard Trauma Questionnaire; IES: Impact of Event Scale; PDS: Posttraumatic Diagnostic Scale; PSS-I: Posttraumatic Stress Symptoms Interview; PCL-C: Posttraumatic Checklist, Civilian; PCL-M: Posttraumatic Checklist, Military; PC-PTSD: Primary Care PTSD screen; PSS-SR: Posttraumatic Stress Symptom Scale Self-Report; SCID: Structured Clinical Interview for DSM–IV; TSQ: Trauma Screening Questionnaire

Machine learning techniques: ANN: Artificial Neural Network; BBR: Bayesian Binary Regression; BIC: Bayesian Information Criteria; CTA: Classification Tree Analysis; DBN: Deep Belief Network; GB: Gaussian Backend; GBM: Gradient Boosting Machines; GGM: Gaussian Graphycal Model; GPC: Gaussian Process Classification; KRR: Kernel Ridge Regression; k-NN: k-nearest Neighbors; LASSO: Least Absolute Shrinkage and Selection Operator; LCGA: Latent Class Growth Analysis LGMM: Latent Growth Mixture Model; LR: Linear Regression; MLP: Multi Layer Perceptron; NB: Naive Bayes; PSM: Product Score Model; RBF: Radial Basis function (kernel); RF: Random Forest; SMO: Sequencial Minimal Optimization; SVM: Single Vector Machine; TL: Transfer Learning

Table 2. Diagnostic classification studies using machine learning techniques in PTSD

First author, year	Data utilized	Diagnostic PTSD tools used	Sample size and diagnosis	ML model	Accuracy	Other measures	Event	Commentary
CLASSIFICATION STUDIES USING NEUROIMAGING AND DTI								
Salminen et al., 2019	sMRI	CAPS	97 subjects - 40 PTSD - 57 TC	SVM	69% (PTSD)	Sensitivity 58% Specificity 81%	Veterans of Iraq and Afghanistan wars	Surface area in the right posterior cingulate was selected as an important feature for classification
Rangaprakash et al., 2017	Questionnaires, fMRI, DTI	PCL-5	87 subjects (male) - 17 PTSD - 42 PTSD+ post-concussion syndrome - 28 TC	SVM	83.59%		Veterans of Iraq and Afghanistan wars	PTSD is associated with hippocampal-striatal hyperconnectivity
Zhang et al., 2016	sMRI, fMRI	CAPS, SCID	57 subjects - 17 PTSD - 20 TC - 20 HC	SVM	89.19% (PTSD vs HC) 90.00% (TC vs HC) 67.57% (PTSD vs TC)	AUC 0.90 (PTSD vs HC)	Earthquake survivors	MRI-based classification among PTSD, TC e HC
Liu et al., 2015	fMRI	CAPS	40 subjects: - 20 PTSD - 20 HC (matched)	SVM	92.5%	AUC 0.91 Sensitivity 90% Specificity 95%	Vehicle accident victims, controls from community	Limbic structure and prefrontal cortex provided the most discriminant features
Gong et al., 2014	sMRI	CAPS-IV, PCL	150 subjects - 50 PTSD - 50 TC - 50 HC	SVM	91.25% (PTSD vs HC, GM+WM) 76% (TE vs	Sensitivity 95% Specificity 87.5% (PTSD vs HC, GM+WM)	Earthquake survivors	Neuroanatomical alterations that could be used to inform the identification of

					HC, GM) 85% (TE vs HC, WM) 67% (PTSD vs TE, GM)			trauma survivors with and without PTSD
CLASSIFICATION STUDIES USING NEUROPSYCHOLOGICAL TESTS								
Breen et al., 2019	Questionnaires, sleep assessment	CAPS	60 subjects - 20 PTSD - 20 TC - 20 HC	SVM	80% (PTSD+TCxHC) 70% (PTSDxTC)	PTSD+TCxHC: AUC 0.80 Sensitivity 87% Specificity 65% PTSDxTC: AUC 0.70 Sensitivity 80% Specificity 61%	Single sexual assault (PTSD and TC groups)	Sleep characteristics were the primary features that could differentiate those with PTSD from those without
Omurca and Ekinci, 2015	Questionnaires		391 subjects - 321 PTSD - 70 TC	SMO, MLP, NB	74-79%		Heterogeneous	Defines a small subset of features to a PTSD classification
CLASSIFICATION STUDIES USING AUDIO RECORDINGS								
Banerjee et al., 2017	Audio recordings (CAPS interview)	CAPS	52 subjects - 26 PTSD - 26 HC (n=168 new testing set)	SVM, DBN, TL	74.99% (DBN/TL) 61.53% (DBN) 57.58% (SVM)		Subjects collected from youtube and hospital	Diagnosis of PTSD patients by analyzing speech signals
van der Broek et al., 2013	Audio recordings		25 subjects (female PTSD)	k-NN, SVM, MLP	89.74% (k-NN) 89.74% (SVM) 82.37% (MLP)		PTSD patients with panic attacks, agoraphobia and PD with agorapobia	Comparison of two stress elicitation methods (storytelling study and valid re-living)

Table 3. Network analysis and unsupervisioned studies using machine learning techniques in PTSD

First author, year	Data utilized	Diagnostic PTSD tools used	Sample size and diagnosis	ML model	Main results	Event	Commentary
CLUSTERING ANALYSIS							
Grisanzio et al., 2018	Questionnaires, EEG	CAPS, SCID	420 subjects (training set) - 100 MDD - 53 PD - 47 PTSD - 220 HC (n=381 new testing set)	Clustering	Clusters: normative mood, tension, anxious arousal, general anxiety, anhedonia, melancholia	Heterogeneous (outpatient and community)	Data-driven approach for identifying transdiagnostic subtypes
Galovski et al., 2016	Questionnaires	CAPS, PDS	69 subjects (PTSD)	LCGA/BIC	Trajectories: partial responders, consistent responders and initial responders	Interpersonal violence survivors	Baseline PTSD and depressive symptoms were associated with patterns of change during CPT
Ma et al., 2016	Questionnaires from phone calls and clinical interviews	CAPS-IV, PSS-I *	957 subjects (TE)	LGMM	ASD prediction: sensitivity 52.7%, specificity 69.5% PTSD prediction: sensitivity 61.6%-66.7%, specificity 69.7%-72.6% Trajectories: non-remitting, slow remitting, rapid remitting (in 10 days)	Emergency room	Early prediction of different trajectories after 10 days of trauma

NETWORK ANALYSIS

Bartels et al., 2019	Questionnaires	CATS	475 children/ adolescents 424 caregivers	LASSO/GGM	Most central symptoms were negative trauma-related cognitions and persistent negative emotional state for the self-report	Mental health clinics (USA, Germany, Norway)	Negative trauma-related cognitions, intrusive thoughts or memories and exaggerated startle response for the caregiver-report
Fried et al., 2018	Questionnaires in mental health centers	HTQ (centers 1 and 4), PSS-SR (center 2), PCL-C (center 3)	2,782 subjects (4 centers, PTSD and TC)	LASSO/GGM	R2 0.62 (Networks 2 and 4) to 0.74 (Networks 1 and 3)	Heterogeneous (clinic patients, soldiers, refugees)	Estimates regularized partial correlation networks of 16 PTSD symptoms
Sullivan et al., 2018	Survey on internet	TSQ	4,639 subjects	LASSO		Students exposed to 2007 Virginia tech shooting	Anger or intrusion likely play a crucial role in the development and maintenance of PTSD
Mitchell et al., 2017	Self-reported questionnaires (online, mail) and telephone interviews	PCL-5, SCID-5	1,377 subjects (62.9% PTSD)	LASSO	R2 0.404 (top six symptoms) R2 0.379 (ICD-11)	Veterans from Iraq and Afghanistan	Network analysis of PTSD symptoms: top 6 more associated with PCL score

3.1 Prognostic studies

Machine learning models are commonly used in prediction studies. Most seek to predict the occurrence of PTSD or non-PTSD from different types of data. Machine learning approaches involved neuroimaging attempting to identify structural, functional, and connectivity changes in the short and long term. Psychological and semi-structured questionnaires are traditional in standard statistical analysis; biological information and audio/video recordings were also found as data sources to machine learning studies.

3.1.1. Prognostic studies using neuroimaging or electro-encephalographic studies

Zandvakili et al. (2019) used a support vector machine to predict response to transmagnetic stimulation (TMS) treatment using electro-encephalographic studies on PTSD and major depressive disorder (MDD) patients, with an AUC of 0.71 (95% CI 0.54–0.91). Nicholson et al. (2018) used resting-state functional magnetic resonance imaging (fMRI) to verify whether a specific form of feature extraction (mean amplitude of low-frequency fluctuations [mALFF]) compared to amygdala maps could predict whether PTSD is associated or not with depressive symptoms, with an accuracy of approximately 89% for both categories. Twenty-two PTSD patients using paroxetine were observed by Yuan et al. (2018) using resting-state fMRI studies; a single vector machine (SVM) approach could predict long-term clinical outcome with an accuracy of 72.5% and an AUC of 0.72. Im et al. (2017) underwent a five-year follow-up study aiming to predict diagnosis and recovery of individuals exposed to trauma; structural changes in a combination of six brain regions could differentiate individuals exposed from healthy controls with an AUC of 0.73 (95% CI 0.54–0.91; $p = 0.01$). Using resting-state fMRI, Jin et al. (2017) found an accuracy of 94% in predicting PTSD from earthquake survivors with dynamic brain connectivity, compared to static connectivity. Li et al. (2016) used diffusion tension imaging (DTI) to predict PTSD in individuals who had suffered a mild TBI, with an accuracy of 75.56%. Additionally, Wang et al. (2016) used DTI with gray and white matter data for early prediction of PTSD in traffic collision survivors and found an accuracy of 79.86%. Cisler et al. (2015) observed female victims of physical and sexual violence to predict trauma memory versus neutral memory recall by fMRI studies, with an accuracy of 76%.

3.1.2. Prognostic studies using questionnaires

Rosellini et al. (2018a) used a sample of more than 23,000 earthquake survivors to assess risk factors that could predict PTSD. The use of a SuperLearner algorithm was superior to logistic regression with an AUC of 0.79 (95% CI 0.78–0.80). Karstoft et al. (2015b) used telephone-based interviews coupled with SVM to discover features that could predict PTSD within 10 days of admission to an emergency department. The authors found an AUC of 0.75 (95% CI 0.67–0.80). He et al. (2017) used text mining methods to evaluate an online survey with 150 PTSD patients compared to 150 healthy controls, to predict PTSD with an accuracy of 82% (using unigrams in a product score model) compared to Naive Bayes, SVM, and decision trees. Schalinski et al. (2016) linked adverse experiences in childhood with the development of PTSD in adult patients; physical neglect at 5 years of age was most pronounced for the development of PTSD ($R^2 = 0.58$; 95% CI 0.44–0.68). Kessler et al. (2014) used several machine learning techniques to analyze 47,466 trauma-exposed subjects from 24 countries (epidemiological data from WHO). SuperLearner algorithms gave the best AUC (0.96) to find pre-trauma predictive factors (such as rape and sexual assault), compared to random forest (AUC 0.90), Least Absolute Shrinkage and Selection Operator (LASSO), Elastic Net, and Ridge.

Reece et al. (2017) used random forest on Twitter posts to predict PTSD and depression. With more than 243,000 posts (related to 63 users diagnosed with PTSD compared to 111 healthy controls), the AUC was 0.89 for PTSD.

Data from 13,690 military personnel were used by Leightley et al. (2019) to identify probable PTSD through a range of machine learning techniques. A random forest approach found a 97% accuracy in predicting PTSD; variables such as alcohol misuse, gender, and deployment status contributed most to the model. Gradus et al. (2017) developed a mail survey to predict suicidal ideation in war veterans (AUC 0.91 male and 0.92 female), where a probable PTSD diagnosis appeared as an important variable. Karstoft et al. (2015a) observed risk factors in soldiers from combat missions in Afghanistan during pre- and post-deployment to find features that could forecast long-term PTSD (pre-deployment: AUC 0.84; 95% CI 0.81–0.87; post-deployment: AUC 0.88; 95% CI 0.85–0.91). Kobach et al. (2015) observed male ex-combatants from a demobilization camp in order to analyze predictive factors of PTSD; the number of lifetime experienced traumatic events was the main predictor ($R^2 = 0.30$) using a RF-CI approach. Marinić et al. (2007) implemented three prediction models on machine learning based on a structured interview, psychiatric scales (CAPS,

PANSS, HAMD) and a combination of both, using random forest and SVM. The CAPS+PANSS model achieved an accuracy of 80.39% in predicting PTSD in war veterans.

Situations of war and refugee victims were also studied using machine learning techniques. Augsburger and Elbert (2017) observed refugees from several localities to predict risk behaviors associated with trauma-exposed individuals. According to a study by Conrad et al. (2017), the use of random forest showed higher accuracy (77.25%) in predicting PTSD in survivors of a civil war in Uganda compared to LASSO (74.88%) and logistic regression (75.36%).

3.1.3. Prognostic studies using biological samples

The study conducted by Galatzer-Levy et al. (2017) used clinical, neuroendocrine, and demographic data to predict pathways to non-remitting PTSD after five months of follow-up from an emergency room environment (AUC 0.82; 95% CI 0.80–0.85). They found that reduced cortisol response (based on 4-hour urine collection) in the emergency room was dependent on the report of early childhood trauma exposure. Also, using blood samples, Tylee et al. (2015) sought to identify gene-expression biomarkers that could predict PTSD. A 20-exon SVM model was made, predicting PTSD with an accuracy of 90% (in a short test subset of 10 subjects, 5 cases and 5 controls). Hemmings et al. (2017) used stool samples to explore the microbiome of PTSD patients. Using a random forest approach to identify associations between bacterial taxa and PTSD, they compared data from 18 PTSD patients with 12 trauma-exposed controls and found a decreased total abundance of bacterial taxa associated with higher CAPS scores ($r = -0.387$; $p = 0.035$).

3.1.4. Prognostic studies using audio and clinical records

Some PTSD prediction studies used information directly from hospital records and clinical data. Harrington et al. (2019) proposed an algorithm to predict probability of PTSD from electronic medical records, with an AUC of 0.95. A model that could predict the disorder after three months of emergency room hospitalization was applied by Papini et al. (2018); using clinical data, localization variables, and psychological questionnaires, they achieved an accuracy of 78% (AUC 0.85; 95% CI 0.83–0.86). In a proof-of-concept study, Saxe et al. (2017) collected 105 risk factors from hospitalization data (concerning childhood development, demographics, parent symptoms, stress, magnitude of injury, candidate genes,

neuroendocrine and psychophysiological response) to develop a prediction model for PTSD in children; the authors found an AUC of 0.79 in a model with 10 variables. Dabek and Caban (2015a) used clinical records of a large set of mild TBI victims (89,840) to develop a neural network model to predict several psychological conditions, with an accuracy of 82.35% overall and 83.82% for PTSD. A replication of this same training set, but with a separated test set (of 16,045 subjects) (Dabek and Caban, 2015b), found an accuracy of 85% to predict PTSD within the first year following the injury. Galatzer-Levy et al. (2014) developed a forecasting model of non-remitting PTSD from information collected within 10 days of a traumatic event; linear SVM was the best model with an AUC of 0.82 (95% CI 0.78–0.86). From ASD symptoms only, they found an AUC of 0.60.

Three studies used machine learning techniques in the preparation and use of features based on audio recordings. Marmar et al. (2019) used audio recordings from clinical interviews to identify speech-based markers that could predict PTSD: slower, monotonous speech; less change in tonality and less activation were identified in PTSD patients with a prediction accuracy of 89.1%. Wortwein and Scherer (2017) analyzed audio and video from patient interviews by a virtual human machine designed to conduct standardized PTSD and depression screening (69 PTSD patients and 129 healthy controls). They found a subset of five questions where the most important nonverbal behaviors could detect symptoms of self-reported PTSD with an F1 score (an accuracy measure) of 0.748. Vergyri et al. (2015) studied audio recordings from war veterans and compared clinician and patient speech elements; the result of several machine learning models generated a prediction accuracy of 77% for PTSD.

3.2 Diagnostic classification studies

Five studies were found using machine learning models to identify and classify cortical areas involved in the diagnosis of PTSD. Salminen et al. (2019) used cortical and subcortical imaging to classify war veterans exposed to early life stress, with a low accuracy of 68% for PTSD diagnosis. Rangaprakash et al. (2017) combined fMRI and DTI information to identify areas related to PTSD; they found an association between hippocampal-striatal hyperconnectivity and PTSD with an accuracy of 83.59%. Zhang et al. (2016) developed a MRI-based (structural and functional) classification of earthquake survivors; they found an accuracy of 89.19% to differentiate PTSD patients from healthy controls, 90% trauma-exposed controls from healthy controls, and 67.57% PTSD patients from trauma-exposed controls. Liu et al. (2015) studied fMRI results from vehicle accident victims who developed

PTSD and compared these to healthy controls; their classification model showed an accuracy of 92.5% (with an AUC of 0.91) in distinguishing the groups. Limbic system and prefrontal cortex provided the most discriminant features. Gong et al. (2014) searched for neuro-anatomical alterations (from sMRI) that could be used to classify trauma survivors with and without PTSD; the most accurate result was a combination of gray and white matter regions to distinguish PTSD patients from healthy controls with an accuracy of 91.25%.

Breen et al. (2019) focused on sleep disturbances and biochemical markers of hyperarousal to classify individuals with PTSD from controls; they found an accuracy of 80% in distinguishing PTSD from trauma and healthy controls (AUC 0.80) with a combination of memory, sleep, and biological markers. Omurca and Ekinci (2015) conducted a study to identify features related to PTSD diagnosis. They compared the original data set with similar data using three different feature selection strategies (chi-square, PCA, and CFS); from 39 features, the best solution identified seven critical features (using a CFS strategy), classified with a Naive Bayes model, obtaining an accuracy of 78.9%.

Two classification studies were carried out using audio recordings on machine learning models. Banerjee et al. (2017) used audio recordings from CAPS interviews (with controls obtained from YouTube recordings) to classify patients with PTSD diagnosis by analyzing speech signals. A combination of Deep Belief Network with Transfer Learning resulted in an accuracy of 74.99%. Van Der Broek et al. (2013) used audio recordings from female PTSD patients to compare two stress elicitation methods (storytelling and re-living studies); from three classification methods, SVM and k-nearest neighbors have the same accuracy (89.74%), followed by SVM (82.37%), to distinguish between the studies.

3.3 Network studies and unsupervised learning

We identified three studies that used unsupervised learning techniques (clustering) and four studies that used a network analysis approach to understand the relationships between the selected features.

Grizancio et al. (2018) executed a clustering analysis to identify subtypes of a set of psychiatric disorders (depressive, panic, and posttraumatic disorders) using symptoms, behavior, and brain function data. They identified six clusters: normative mood, tension, anxious arousal, general anxiety, anhedonia, and melancholia. Questionnaires from phone calls and clinical interviews were used by Ma et al. (2016) to develop a clinical decision support system in early trauma; three outcomes were identified (non-remitting, slow

remitting, rapid remitting) with sensitivity between 0.616 and 0.667 and specificity between 0.697 and 0.726. Galovski et al. (2016) used a latent class growth analysis, a machine learning approach that identifies clusters that change over time. The authors aimed to identify patterns of change during cognitive behavioral therapy. They isolated three distinct groups based on change patterns: partial, consistent, and initial responders.

Concerning network analysis, Bartels et al. (2019) compared a group of children and adolescents with PTSD with their caregivers to identify core symptoms of the disorder. Negative trauma-related cognitions and persistent negative emotional state were found to be the most central symptoms in this sample. Fried et al. (2018) studied a set of 2,782 subjects from four different centers and analyzed the correlations among 16 PTSD symptoms; a network correlation between 0.62 and 0.74 was found when comparing the four centers. Sullivan et al. (2018) used network analysis to understand the connection and strength among PTSD symptoms from victims of a mass violence event; according to their results, intrusive thoughts had the strongest influence on other symptoms, and anger produced the shortest path (stronger connections) to all other symptoms. Studying war veterans, Mitchell et al. (2017) identified the six most central symptoms of PTSD in their sample (persistent negative emotional state, efforts to avoid external reminders, efforts to avoid thoughts or memories, inability to experience positive emotions, distressing dreams, and intrusive distressing thoughts or memories) and correlated them to PCL-5 scores. The top six symptoms produced a correlation coefficient of $R^2 = 0.404$, better than ICD-11 criteria ($R^2 = 0.379$).

3.4 Quality assessment

We proposed a quality assessment of the studies that we reviewed, given that no validated instrument exists yet to measure the quality of machine learning studies. We evaluated nine attributes that we consider relevant for assessing the quality of these studies, including methodological features, the presence of relevant information such as performance metrics, and the presence of technical detail such as handling missing data and feature selection procedures. The results are described in Supplementary Table 2.

Regarding the representativeness of the samples used in the studies included in this review, half of the articles (24 of 49 articles; 49.0%) considered a statistical sample or a larger sample representative of the described population of the study. Half of the studies (53.0%; 26 of 49 articles) used techniques to control confounding variables such as age, gender, or trauma type. An independent, blind assessment was used to measure outcomes including

PTSD in 38.7% (19 of 49 articles). The remaining studies used self-reported instruments, medical records, or non-blind interviews. The majority of the studies included a description of the machine learning techniques used (73.5%; 36 of 49 articles) and appointed the performance metrics (83.7%; 41 of 49 articles) such as accuracy, AUC, sensitivity, and specificity. Considering technical aspects, 26.5% of the studies (13 of 49 articles) provided information on how missing data were handled, mainly by exclusion or imputation; this aspect is important for the execution of machine learning techniques and could influence the results, depending on the chosen technique. Fifty studies (30.6%) used a separate, different data set to test the output model. Fourteen studies (28.5%) described how they dealt with the class imbalance problem, which is important for adjusting and interpreting performance accuracy; most studies used the same sample size of cases and controls. The use of techniques for feature selection and/or hyperparameter optimization was described in 61.2% of the studies (30 of 49). These techniques are necessary to reduce dimensionality and to ensure more robust and simpler models.

4. Discussion

We evaluated 49 articles that used machine learning techniques to assess trauma-related disorders, including ASD and PTSD. We identified several studies designed to aid in the prediction and diagnostic classification of PTSD. No study aimed at the treatment of the disorder was found.

One of the most immediate uses of robust techniques such as machine learning in psychiatry is developing predictive models of mental health disorders, especially from risk factors. Thirty-three of the 49 articles in this review used machine learning to this purpose. Different predictors will almost certainly be found in different weights in diverse populations (e.g., military veterans, disaster survivors, civilians in less-developed countries) and in different screening settings (e.g., outpatient setting, emergency rooms, medical clinics in war zones, trauma units), demanding varying machine learning strategies. Several studies conducted in specific populations were observed in our review: war veterans (13 studies: Harrington et al., 2019; Marmar et al., 2019; Salminen et al., 2019; Zandvakili et al., 2019; Gradus et al., 2017; Mitchell et al., 2017; Rangaprakash et al., 2017; Wortwein and Scherer, 2017; Karstoft et al., 2015b; Kobach et al., 2015; Tylee et al., 2015; Vergyri et al., 2015; Marinić et al., 2007), survivors of specific natural disasters (5 studies: Rosellini et al., 2018a; Yuan et al., 2018; Jin et al., 2017; Zhang et al., 2016; Gong et al., 2014), and refugees from

conflict areas (3 studies: Fried et al., 2018, partially; Augsburger and Elbert, 2017; Conrad et al., 2017) are among the most studied populations in our review, possibly reproducing limited visions of the disorder; this challenge is also observed in the current literature of PTSD (Tay et al., 2016). We highlight that studies in heterogeneous populations, which could provide a broader view of PTSD, are lacking; also, studies comparing these subpopulations may provide new insight into possible different phenotypes of PTSD.

An assortment of instruments was used to assess PTSD. Thirty from the 34 prognostic studies (35.7%) used the CAPS scale (versions IV and 5). The Clinician-adapted PTSD Scale is considered the gold standard in the assessment of PTSD, in its last version adapted to DSM-5 PTSD criteria (Weathers et al., 2018). Five studies used the PSS-I scale (Foa and Tolin, 2000) (*DSM-IV* version). Both instruments perform not only the diagnosis of PTSD but also provide a measurement of the severity of symptoms allowing the researcher to use quantitative outcomes, which is particularly interesting for the application of machine learning techniques where it is possible to consider regression (using severity measures) and binary classification outcomes (PTSD vs. non-PTSD), such as SVM and Logistic Regression. Eight studies used the PCL scale (Blanchard et al., 1996), using civilian or military versions, depending on the sample studied. Another challenge regarding assessing PTSD lies in current diagnostic systems; while *DSM-5* relies heavily on the occurrence of a specific traumatic event ("criterion A"), the new ICD-11 proposal is based on the reduction of diagnostic criteria as a way of increasing specificity by changing the prevalence and clinical characteristics of the ICD-11 disorder (Barbano et al., 2019). We remark that more studies using the advantages of machine learning techniques may be conducted to identify – or to develop – a more suitable instrument to evaluate trauma disorders that could reflect both classification systems.

While machine learning techniques were developed mainly from the theory of neural networks, it now encompasses a much more diverse set of algorithms (Liu and Salinas, 2017). In our review, SVM appears to be the preferred method in prognostic and classification studies. In its original version, SVM deals with binary classification problems, which seems to be suitable for studies that aim to predict "have or do not have" the disorder. At least half of the prognostic studies (18 of 28) and almost all classification studies (8 of 9) used SVM alone or combined with other techniques, to look for the best model, in different types of data. The objectives and methods of clustering studies were diverse and included identifying transdiagnostic subtypes of PTSD (normative mood, tension, anxious arousal, general anxiety, anhedonia, melancholia; Grisanzio et al., 2018), differentiating trajectories of patients during treatment (partial/consistent/initial responders; Galovski et al., 2016) and during 10 days after

trauma (non-/slow/rapid remitting; Ma et al., 2016). Four studies applied network analysis using LASSO technique (Bartels et al., 2019; Fried et al., 2018; Sullivan et al., 2018; Mitchell et al., 2017), aiming to provide potentially clinically meaningful pathways of interconnectivity (Sullivan et al., 2018). It also allows identification of 'central' symptoms, defined by strong correlations with a large number of other symptoms (Beard et al., 2017): Bartels et al. (2019) found that most central symptoms of trauma disorders in children and adolescents were negative trauma-related cognitions and persistent negative emotional state. Our data have shown that most of the studies included a detailed description of the chosen machine learning techniques (73.5%), which is important for understanding and reproducibility of the developed model.

A proper measurement of results is of particular interest. The accuracy of the model is evaluated using a 'testing' set that consists of observations previously 'unseen' by the model; machine learning models can be evaluated in terms of accuracy, sensivity, specificity and AUC of a receiver operating characteristic test, allowing a comparison between them (Passos et al., 2016). Of the 18 prognostic studies that reported accuracy, nine presented an excellent predictive accuracy above 80% (Leightley et al., 2019; Marmar et al., 2019; Nicholson et al., 2018; He et al., 2017; Jin et al., 2017; Dabek and Caban, 2015a, 2015b; Tylee et al., 2015; Marinić et al., 2007) and all presented at least good accuracy above 70%. Six classification studies presented an excellent classification accuracy above 80% (Breen et al., 2019; Rangaprakash et al., 2017; Zhang et al., 2016; Liu et al., 2015; Gong et al., 2014; van der Broek et al., 2013). We remark that most of the studies presented at least a 'good' accuracy measure. Most of the reviewed studies appointed performance metrics (83.7%; 41 of 49 articles).

It is noteworthy that many of the reviewed studies have a small sample size. Although machine learning can show robustness in the analysis of large data sets (the so-called 'big data'), certain techniques can work with a small sample without masking accuracy, depending on the model adjustments. Small samples in mental health studies are common because tasks and experimental protocols in different conditions are still under development and because of the costs associated with data collection involving human participants (Vabalas et al., 2019). According to our quality assessment, half of the articles considered a statistical sample or a larger sample that could represent the population of the study (49.0%); half of the articles used techniques to control confounding variables (53.0%) and less than a half performed blinded interviews (38.7%), reflecting issues to improve generalization of the results. Also, a low amount of the studies (30%) used the same sample to perform training and validating the

models; it is highly recommended that a separate data set could be used for final performance evaluation, with data not used in training or validation of the model (Baştanlar and Özuysal, 2014).

New technologies and sources of information have begun to appear in PTSD research. Machine learning techniques were used to evaluate different types of applications such as text mining (He et al., 2017), genetics (Hemmings et al., 2017; Tylee et al., 2015), and outputs of a virtual human machine (Wortwein and Scherer, 2017). Different sources of data such as Twitter (Twitter (Reece et al., 2017) and YouTube (Banerjee et al., 2017) were used to feed learning models. 'Big data' structures such as epidemiological resources from the World Health Organization (Kessler et al., 2014) and local data sources with more than 10,000 patients (Leightley et al., 2019; Rosellini et al., 2018a; Dabek et al., 2015a, 2015b) used the robustness of machine learning data manipulation.

The present study evaluated the use of machine learning techniques in PTSD research. There are many reasons why these techniques might outperform standard parametric methods: for instance, in comparison to conventional regression examining the direct effect of predictors on an outcome, some algorithms can automatize identification of hidden interactions and non-linearities among features. In addition, where a conventional regression based on highly correlated independent variables might have good prediction accuracy in a given sample but perform poorly in independent samples (model overfitting), machine learning methods can be used to reduce the likelihood of overestimating prediction performance (Rosellini et al., 2018b). Machine learning techniques allow researchers to handle bigger and more complex data sets, and integrate information from very heterogeneous sources, such as audio/video recordings, and biological samples. However, some obstacles including computational power, multimodality, model validation, heterogeneity both phenotypically and etiologically, assessment of rare events, cost and non-stationary distribution of the data, phenomenological diagnosis, lack of a uniform pipeline for machine learning studies, lack of appropriate funding, and lack of interpretability, need to be addressed (Passos et al., 2019). Manipulating data is also a challenge to machine learning applications. For instance, a model performance can be altered by the nature of the data (Chekroud and Koutsouleris, 2018): a small sample size could inflate the accuracy of a machine learning model.

To promote the development of better machine learning models for PTSD diagnosis, future studies should compare the performance of experts in the research field and in clinical practice with the generated models. An instrument of quality assessment on machine learning

techniques should be developed to provide more robust and reliable models. In this sense, a method to evaluate the quality of machine learning based studies was proposed in our study; further researches are being conducted to validate this instrument. We remark that good practices to develop machine learning studies are necessary, concerning theoretical development, knowledge of different techniques, methodological rigor, care of choosing and handling suitable data, and presenting results adequately.

The recent area of personalized psychiatry has advanced toward this goal in other areas such as bipolar mood disorder (Salagre et al., 2018); machine learning may allow development of personalized interventions to prevent the transition from prodromes to the full disorder in high-risk patients (Librenza-Garcia et al., 2017). Computational psychiatry approaches combine multiple levels and types of computation with multiple types of data in an effort to improve understanding, prediction and treatment of mental illness (Huys et al., 2016). Machine learning techniques can be used to find relationships between data not found with traditional statistics; etiologic heterogeneity may be a hallmark of complex disorders including PTSD (Tylee et al., 2015), and it is suited to the purpose of these techniques. A major challenge for the future is to use the models developed from machine learning studies in clinical practice for the benefit of patients.

Acknowledgements: This study was financed in part by the Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES, Brazil, Finance Code 001) and by the Conselho Nacional de Pesquisa (CNPq, Brazil).

References

American Psychiatric Association, 2013. Diagnostic and statistical manual of mental disorders. fifth ed. Washington (DC).

Augsburger, M., Elbert, T., 2017. When do traumatic experiences alter risktaking behavior? A machine learning analysis of reports from refugees. PLoS One 12, e0177617. https://doi.org/10.1371/journal.pone.0177617

Banerjee, D., Islam, K., Mei, G., Xiao, L.M., Zhang, G.F., Xu, R., Ji, S.W., Li, J., 2017. A Deep Transfer Learning Approach for Improved Post-Traumatic Stress Disorder Diagnosis, in: 2017 17th Ieee International Conference on Data Mining. Ieee, New York, pp. 11–20.

Barbano, A.C., Van Der Mei, W.F., Bryant, R.A., Delahanty, D.L., Deroon-Cassini, T.A., Matsuoka, Y.J., Olff, M., Qi, W., Ratanatharathorn, A., Schnyder, U., Seedat, S., Kessler,

R.C., Koenen, K.C., Shalev, A.Y., 2019. Clinical implications of the proposed ICD-11 PTSD diagnostic criteria. Psychol. Med. 49, 483–490. https://doi.org/10.1017/S0033291718001101

Bartels, L., Berliner, L., Holt, T., Jensen, T., Jungbluth, N., Plener, P., Risch, E., Rojas, R., Rosner, R., Sachser, C., 2019. The importance of the DSM-5 posttraumatic stress disorder symptoms of cognitions and mood in traumatized children and adolescents: two network approaches. J Child Psychol Psychiatry 60, 545–554. https://doi.org/10.1111/jcpp.13009

Baştanlar ,Y., Özuysal, M., 2014. Introduction to Machine Learning. In: Yousef M., Allmer J. (eds) miRNomics: MicroRNA Biology and Computational Analysis. Methods in Molecular Biology (Methods and Protocols), vol 1107. Humana Press, Totowa, NJ. https://doi.org/10.1007/978-1-62703-748-8_7

Beard, C., Millner, A.J., Forgeard, M.J., Fried, E.I., Hsu, K.J., Treadway, M.T., Leonard, C.V., Kertz, S.J., Bjorgvinsson, T., 2016. Network analysis of depression and anxiety symptom relationships in a psychiatric sample. Psychol Med 46, 3359-3369. https://doi.org/10.1017/S0033291716002300

Benjet, C., Bromet, E., Karam, E.G., Kessler, R.C., McLaughlin, K.A., Ruscio, A.M., Shahly, V., Stein, D.J., Petukhova, M., Hill, E., Alonso, J., Atwoli, L., Bunting, B., Bruffaerts, R., Caldas-de-Almeida, J.M., de Girolamo, G., Florescu, S., Gureje, O., Huang, Y., Lepine, J.P., Kawakami, N., Kovess-Masfety, V., Medina-Mora, M.E., Navarro-Mateu, F., Piazza, M., Posada-Villa, J., Scott, K.M., Shalev, A., Slade, T., ten Have, M., Torres, Y., Viana, M.C., Zarkov, Z., Koenen, K.C., 2016. The epidemiology of traumatic event exposure worldwide: results from the World Mental Health Survey Consortium. Psychol. Med. https://doi.org/10.1017/S0033291715001981

Blanchard, E.B., Jones-Alexander, J., Buckley, T.C., Forneris, C.A., 1996. Psychometric properties of the PTSD checklist (PCL). Behav. Res. Ther. 34, 669–673. https://doi.org/10.1016/0005-7967(96)00033-2

Breen, M.S., Thomas, K.G.F., Baldwin, D.S., Lipinska, G., 2019. Modelling PTSD diagnosis using sleep, memory, and adrenergic metabolites: An exploratory machine-learning study. Hum Psychopharmacol 34, e2691. https://doi.org/10.1002/hup.2691

Bishop, C.M., 2006. Pattern Recognition and Machine Learning. Springer, p. 3.

Charney, M.E., Hellberg, S.N., Bui, E., Simon, N.M., 2018. Evidenced-Based Treatment of Posttraumatic Stress Disorder: An Updated Review of Validated Psychotherapeutic and Pharmacological Approaches. Harv. Rev. Psychiatry. 26, 99-115. https://doi.org/10.1097/HRP.0000000000000186

Chekroud, A.M., Koutsouleris, N., 2018. The perilous path from publication to practice. Mol Psychiatry 23, 24–25. https://doi.org/10.1038/mp.2017.227

Cisler, J.M., Bush, K., James, G.A., Smitherman, S., Kilts, C.D., 2015. Decoding the Traumatic Memory among Women with PTSD: Implications for Neurocircuitry Models of PTSD and Real-Time fMRI Neurofeedback. PLoS One 10, e0134717. https://doi.org/10.1371/journal.pone.0134717

Conrad, D., Wilker, S., Pfeiffer, A., Lingenfelder, B., Ebalu, T., Lanzinger, H., Elbert, T., Kolassa, I.T., Kolassa, S., 2017. Does trauma event type matter in the assessment of traumatic load? Eur. J. Psychotraumatol. 8, 1344079. https://doi.org/10.1080/20008198.2017.1344079

Dabek, F., Caban, J.J., 2015a. A Neural Network Based Model for Predicting Psychological Conditions, in: Brain Informatics and Health. Springer-Verlag Berlin, Berlin, pp. 252–261.

Dabek, F., Caban, J.J., 2015b. Leveraging Big Data to Model the Likelihood of Developing Psychological Conditions After a Concussion, in: Inns Conference on Big Data 2015 Program. Elsevier Science Bv, Amsterdam, pp. 265–273.

Deo, R.C., 2015. Machine Learning in Medicine. Circulation, 132, 1920-1930. https://doi.org/10.1161/CIRCULATIONAHA.115.001593

Foa, E.B., Tolin, D.F., 2000. Comparison of the PTSD Symptom Scale-Interview Version and the Clinician-Administered PTSD Scale. J. Trauma. Stress 13, 181–191. https://doi.org/10.1023/A:1007781909213

Fried, E.I., Eidhof, M.B., Palic, S., Costantini, G., Huisman-van Dijk, H.M., Bockting, C.L.H., Engelhard, I., Armour, C., Nielsen, A.B.S., Karstoft, K.I., 2018. Replicability and Generalizability of Posttraumatic Stress Disorder (PTSD) Networks: A Cross-Cultural Multisite Study of PTSD Symptoms in Four Trauma Patient Samples. Clin. Psychol. Sci. 6, 335–351. https://doi.org/10.1177/2167702617745092

Galatzer-Levy, I.R., Ma, S., Statnikov, A., Yehuda, R., Shalev, A.Y., 2017. Utilization of machine learning for prediction of post-traumatic stress: a re-examination of cortisol in the prediction and pathways to non-remitting PTSD. Transl Psychiatry 7, e0. https://doi.org/10.1038/tp.2017.38

Galatzer-Levy, I.R., Karstoft, K.I., Statnikov, A., Shalev, A.Y., 2014. Quantitative forecasting of PTSD from early trauma responses: a Machine Learning application. J Psychiatr Res 59, 68–76. https://doi.org/10.1016/j.jpsychires.2014.08.017

Galovski, T.E., Harik, J.M., Blain, L.M., Farmer, C., Turner, D., Houle, T., 2016. Identifying Patterns and Predictors of PTSD and Depressive Symptom Change During Cognitive Processing Therapy. Cognit. Ther. Res. 40, 617–626.

Gong, Q., Li, L., Tognin, S., Wu, Q., Pettersson-Yeo, W., Lui, S., Huang, X., Marquand, A.F., Mechelli, A., 2014. Using structural neuroanatomy to identify trauma survivors with and without post-traumatic stress disorder at the individual level. Psychol Med 44, 195–203. https://doi.org/10.1017/s0033291713000561

Gradus, J.L., King, M.W., Galatzer-Levy, I., Street, A.E., 2017. Gender Differences in Machine Learning Models of Trauma and Suicidal Ideation in Veterans of the Iraq and Afghanistan Wars. J Trauma Stress 30, 362–371. https://doi.org/10.1002/jts.22210

Greenhalgh, T., Howick, J., Maskrey, N., Brassey, J., Burch, D., Burton, M., Chang, H., Glasziou, P., Heath, I., Heneghan, C., Kelly, M.P., Lehman, R., Llewelyn, H., McCartney, M., Milne, R., Spence, D., 2014. Evidence based medicine: A movement in crisis? BMJ. 13, 348:g3725. https://doi.org/10.1136/bmj.g3725

Grisanzio, K.A., Goldstein-Piekarski, A.N., Wang, M.Y., Rashed Ahmed, A.P., Samara, Z., Williams, L.M., 2018. Transdiagnostic Symptom Clusters and Associations With Brain, Behavior, and Daily Function in Mood, Anxiety, and Trauma Disorders. JAMA Psychiatry 75, 201–209. https://doi.org/10.1001/jamapsychiatry.2017.3951

Hahn, T., Nierenberg, A.A., Whitfield-Gabrieli, S., 2017. Predictive analytics in mental health: Applications, guidelines, challenges and perspectives. Mol. Psychiatry. 22, 37–43. https://doi.org/10.1038/mp.2016.201

Harrington, K.M., Quaden, R., Stein, M.B., Honerlaw, J.P., Cissell, S., Pietrzak, R.H., Zhao, H., Radhakrishnan, K., Aslan, M., Gaziano, J.M., Concato, J., Gagnon, D.R., Gelernter, J., Cho, K., 2019. Validation of an Electronic Medical Record-Based Algorithm for Identifying Posttraumatic Stress Disorder in U.S. Veterans. J Trauma Stress 32, 226–237. https://doi.org/10.1002/jts.22399

He, Q., Veldkamp, B.P., Glas, C.A., de Vries, T., 2017. Automated Assessment of Patients' Self-Narratives for Posttraumatic Stress Disorder Screening Using Natural Language Processing and Text Mining. Assessment 24, 157–172. https://doi.org/10.1177/1073191115602551

Hemmings, S.M.J., Malan-Muller, S., van den Heuvel, L.L., Demmitt, B.A., Stanislawski, M.A., Smith, D.G., Bohr, A.D., Stamper, C.E., Hyde, E.R., Morton, J.T., Marotz, C.A., Siebler, P.H., Braspenning, M., Van Criekinge, W., Hoisington, A.J., Brenner, L.A., Postolache, T.T., McQueen, M.B., Krauter, K.S., Knight, R., Seedat, S., Lowry, C.A., 2017. The Microbiome in Posttraumatic Stress Disorder and Trauma-Exposed Controls: An Exploratory Study. Psychosom Med 79, 936–946. https://doi.org/10.1097/psy.0000000000000512

Huys, Q. J., Maia, T. V., Frank, M. J., 2016. Computational psychiatry as a bridge from neuroscience to clinical applications. Nat Neurosci 19, 404–413. https://doi.org/10.1038/nn.4238

Im, J.J., Kim, B., Hwang, J., Kim, J.E., Kim, J.Y., Rhie, S.J., Namgung, E., Kang, I., Moon, S., Lyoo, I.K., Park, C.H., Yoon, S., 2017. Diagnostic potential of multimodal neuroimaging in posttraumatic stress disorder. PLoS One 12, e0177847. https://doi.org/10.1371/journal.pone.0177847

Jin, C.F., Jia, H., Lanka, P., Rangaprakash, D., Li, L.J., Liu, T.M., Hu, X.P., Deshpande, G., 2017. Dynamic Brain Connectivity Is a Better Predictor of PTSD than Static Connectivity. Hum Brain Mapp 38, 4479–4496. https://doi.org/10.1002/hbm.23676.

Kannel, W.B., Doyle, J.T., McNamara, P.M., Quickenton, P., Gordon, T., 1975. Precursors of sudden coronary death. Factors related to the incidence of sudden death. Circulation 51, 606-613. https://doi.org/10.1161/01.cir.51.4.606

Karstoft, K.I., Galatzer-Levy, I.R., Statnikov, A., Li, Z., Shalev, A.Y., Ankri, Y., Freedman, S., Addesky, R., Israeli-Shalev, Y., Gilad, M., Roitman, P., 2015a. Bridging a translational gap: Using machine learning to improve the prediction of PTSD. BMC Psychiatry 15, 30. https://doi.org/10.1186/s12888-015-0399-8

Karstoft, K.I., Statnikov, A., Andersen, S.B., Madsen, T., Galatzer-Levy, I.R., 2015b. Early identification of posttraumatic stress following military deployment: Application of machine learning methods to a prospective study of Danish soldiers. J. Affect. Disord. 184, 170–175. https://doi.org/10.1016/j.jad.2015.05.057

Kessler, R.C., Rose, S., Koenen, K.C., Karam, E.G., Stang, P.E., Stein, D.J., Heeringa, S.G., Hill, E.D., Liberzon, I., McLaughlin, K.A., McLean, S.A., Pennell, B.E., Petukhova, M., Rosellini, A.J., Ruscio, A.M., Shahly, V., Shalev, A.Y., Silove, D., Zaslavsky, A.M., Angermeyer, M.C., Bromet, E.J., de Almeida, J.M., de Girolamo, G., de Jonge, P., Demyttenaere, K., Florescu, S.E., Gureje, O., Haro, J.M., Hinkov, H., Kawakami, N., Kovess-Masfety, V., Lee, S., Medina-Mora, M.E., Murphy, S.D., Navarro-Mateu, F., Piazza, M., Posada-Villa, J., Scott, K., Torres, Y., Carmen Viana, M., 2014. How well can post-traumatic stress disorder be predicted from pre-trauma risk factors? An exploratory study in the WHO World Mental Health Surveys. World Psychiatry 13, 265–274. https://doi.org/10.1002/wps.20150

Kessler, R.C., Sonnega, A., Bromet, E., Hughes, M., Nelson, C.B., 1995. Posttraumatic Stress Disorder in the National Comorbidity Survey. Arch. Gen. Psychiatry 52, 1048–1060. https://doi.org/10.1001/archpsyc.1995.03950240066012

Kilpatrick, D.G., Resnick, H.S., Milanak, M.E., Miller, M.W., Keyes, K.M., Friedman, M.J., 2013. National Estimates of Exposure to Traumatic Events and PTSD Prevalence Using DSM-IV and DSM-5 Criteria. J. Trauma. Stress 26, 537–547. https://doi.org/10.1002/jts.21848

Köbach, A., Nandi, C., Crombach, A., Bambonyé, M., Westner, B., Elbert, T., 2015. Violent offending promotes appetitive aggression rather than posttraumatic stress-a replication study with burundian ex-combatants. Front. Psychol. 6, 1755. https://doi.org/10.3389/fpsyg.2015.01755

Leightley, D., Williamson, V., Darby, J., Fear, N.T., 2019. Identifying probable post-traumatic stress disorder: applying supervised machine learning to data from a UK military cohort. J Ment Heal. 28, 34–41. https://doi.org/10.1080/09638237.2018.1521946

Li, L., Sun, G., Liu, K., Li, M., Li, B., Qian, S.W., Yu, L.L., 2016. White Matter Changes in Posttraumatic Stress Disorder Following Mild Traumatic Brain Injury: A Prospective Longitudinal Diffusion Tensor Imaging Study. Chin Med J 129, 1091–1099. https://doi.org/10.4103/0366-6999.180518

Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P.A., Clarke, M., Devereaux, P.J., Kleijnen, J., Moher, D., 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration, in: Journal of Clinical Epidemiology. pp. e1–e34. https://doi.org/10.1016/j.jclinepi.2009.06.006

Librenza-Garcia, D., Kotzian, B.J., Yang, J., Mwangi, B., Cao, B., Pereira Lima, L.N., Bermudez, M.B., Boeira, M.V., Kapczinski, F., Passos, I.C., 2017. The impact of machine learning techniques in the study of bipolar disorder: A systematic review. Neurosci. Biobehav. Rev. 80, 538–554. https://doi.org/10.1016/j.neubiorev.2017.07.004

Liu, N.T., Salinas, J., 2017. Machine Learning for Predicting Outcomes in Trauma. Shock 48, 504–510. https://doi.org/10.1097/SHK.0000000000000898

Liu, F., Xie, B., Wang, Y., Guo, W., Fouche, J.P., Long, Z., Wang, W., Chen, H., Li, M., Duan, X., Zhang, J., Qiu, M., Chen, H., 2015. Characterization of post-traumatic stress disorder using resting-state fMRI with a multi-level parametric classification approach. Brain Topogr 28, 221–237. https://doi.org/10.1007/s10548-014-0386-2

Ma, S., Galatzer-Levy, I.R., Wang, X., Fenyö, D., Shalev, A.Y., 2016. A First Step towards a Clinical Decision Support System for Post-traumatic Stress Disorders. AMIA Annu Symp Proc 2016, 837–843.

Marinić, I., Supek, F., Kovačić, Z., Rukavina, L., Jendričko, T., Kozarić-Kovačić, D., 2007. Posttraumatic stress disorder: Diagnostic data analysis by data mining methodology. Croat. Med. J. 48, 185–197.

Marmar, C.R., Brown, A.D., Qian, M., Laska, E., Siegel, C., Li, M., Abu-Amara, D., Tsiartas, A., Richey, C., Smith, J., Knoth, B., Vergyri, D., 2019. Speech-based markers for posttraumatic stress disorder in US veterans. Depress. Anxiety 36, 607–616. https://doi.org/10.1002/da.22890

McGorry, P.D., Hickie, I.B., Yung, A.R., Pantelis, C., Jackson, H.J., 2006. Clinical staging of psychiatric disorders: A heuristic framework for choosing earlier, safer and more effective interventions. Aust. N. Z. J. Psychiatry. 40, 616–622. https://doi.org/10.1111/j.1440-1614.2006.01860.x

Mitchell, K.S., Wolf, E.J., Bovin, M.J., Lee, L.O., Green, J.D., Rosen, R.C., Keane, T.M., Marx, B.P., 2017. Network models of DSM-5 posttraumatic stress disorder: Implications for ICD-11. J Abnorm Psychol 126, 355–366. https://doi.org/10.1037/abn0000252

Nash, W.P., Watson, P.J., 2012. Review of VA/DOD clinical practice guideline on management of acute stress and interventions to prevent posttraumatic stress disorder. J. Rehabil. Res. Dev. 49, 637–648. https://doi.org/10.1682/JRRD.2011.10.0194

Nicholson, A.A., Densmore, M., McKinnon, M.C., Neufeld, R.W.J., Frewen, P.A., Théberge, J., Jetly, R., Richardson, J.D., Lanius, R.A., 2018. Machine learning multivariate pattern analysis predicts classification of posttraumatic stress disorder and its dissociative subtype: A multimodal neuroimaging approach. Psychol. Med. 1–11. https://doi.org/10.1017/S0033291718002866

Omurca, S.I., Ekinci, E., 2015. An Alternative Evaluation of Post Traumatic Stress Disorder with Machine Learning Methods, in: 2015 International Symposium on Innovations in Intelligent Systems and Applications. Ieee, New York, pp. 231–237.

Ozer, E.J., Best, S.R., Lipsey, T.L., Weiss, D.S., 2008. Predictors of posttraumatic stress disorder and symptoms in adults: A meta-analysis. Psychol. Bull. Psychol. Trauma Theory Res. Pract. Policy. S(1), 3–36. https://doi.org/10.1037/1942-9681.S.1.3

Papini, S., Pisner, D., Shumake, J., Powers, M.B., Beevers, C.G., Rainey, E.E., Smits, J.A.J., Warren, A.M., 2018. Ensemble machine learning prediction of posttraumatic stress disorder screening status after emergency room hospitalization. J Anxiety Disord 60, 35–42. https://doi.org/10.1016/j.janxdis.2018.10.004

Passos, I.C., Mwangi, B., Cao, B., Hamilton, J.E., Wu, M.J., Zhang, X.Y., Zunta-Soares, G.B., Quevedo, J., Kauer-Sant'Anna, M., Kapczinski, F., Soares, J.C., 2016. Identifying a

clinical signature of suicidality among patients with mood disorders: A pilot study using a machine learning approach. J Affect Disord. 193, 109–116. https://doi.org/10.1016/j.jad.2015.12.066.

Passos, I.C., Ballester, P.L., Barros, R.C., Librenza-Garcia, D., Mwangi, B., Birmaher, B., Brietzke, E., Hajek, T., Lopez Jaramillo, C., Mansur, R.B., Alda, M., Haarman, B.C.M., Isometsa, E., Lam, R.W., McIntyre, R.S., Minuzzi, L., Kessing, L.V., Yatham, L.N., Duffy, A., Kapczinski, F., 2019. Machine Learning and Big Data Analytics in Bipolar Disorder: A Position Paper from the International Society for Bipolar Disorders Big Data Task Force. Bipolar Disord 21, 582-594. https://doi.org/10.1111/bdi.12828.

Rangaprakash, D., Deshpande, G., Daniel, T.A., Goodman, A.M., Robinson, J.L., Salibi, N., Katz, J.S., Denney Jr., T.S., Dretsch, M.N., 2017. Compromised hippocampus-striatum pathway as a potential imaging biomarker of mild-traumatic brain injury and posttraumatic stress disorder. Hum Brain Mapp 38, 2843–2864. https://doi.org/10.1002/hbm.23551

Reece, A.G., Reagan, A.J., Lix, K.L.M., Dodds, P.S., Danforth, C.M., Langer, E.J., 2017. Forecasting the onset and course of mental illness with Twitter data. Sci. Rep. 7, 13006. https://doi.org/10.1038/s41598-017-12961-9

Rosellini, A.J., Stein, M.B., Benedek, D.M., Bliese, P.D., Chiu, W.T., Hwang, I., Monahan, J., Nock, M.K., Sampson, N.A., Street, A.E., Zaslavsky, A.M., Ursano, R.J., Kessler, R.C., 2018a. Predeployment predictors of psychiatric disorder-symptoms and interpersonal violence during combat deployment. Depress. Anxiety 35, 1073–1080. https://doi.org/10.1002/da.22807

Rosellini, A. J., Dussaillant, F., Zubizarreta, J. R., Kessler, R. C., Rose, S., 2018b. Predicting posttraumatic stress disorder following a natural disaster. J Psychiatr Res 96, 15–22. https://doi.org/10.1016/j.jpsychires.2017.09.010

Salagre, E., Dodd, S., Aedo, A., Rosa, A., Amoretti, S., Pinzon, J., Reinares, M., Berk, M., Kapczinski, F.P., Vieta, E., Grande, I., 2018. Toward Precision Psychiatry in Bipolar Disorder: Staging 2.0. Front. Psychiatry 9, 641. https://doi.org/10.3389/fpsyt.2018.00641

Salminen, L.E., Morey, R.A., Riedel, B.C., Jahanshad, N., Dennis, E.L., Thompson, P.M., 2019. Adaptive Identification of Cortical and Subcortical Imaging Markers of Early Life Stress and Posttraumatic Stress Disorder. J Neuroimaging 29, 335–343. https://doi.org/10.1111/jon.12600

Saxe, G.N., Ma, S., Ren, J., Aliferis, C., 2017. Machine learning methods to predict child posttraumatic stress: A proof of concept study. BMC Psychiatry 17, 223. https://doi.org/10.1186/s12888-017-1384-1

Schalinski, I., Teicher, M.H., Nischk, D., Hinderer, E., Müller, O., Rockstroh, B., 2016. Type and timing of adverse childhood experiences differentially affect severity of PTSD, dissociative and depressive symptoms in adult inpatients. BMC Psychiatry 16, 295. https://doi.org/10.1186/s12888-016-1004-5

Schultebraucks, K., Galatzer-Levy, I.R., 2019. Machine Learning for Prediction of Posttraumatic Stress and Resilience Following Trauma: An Overview of Basic Concepts and Recent Advances. J. Trauma. Stress 32, 215–225. https://doi.org/10.1002/jts.22384

Sullivan, C.P., Smith, A.J., Lewis, M., Jones, R.T., 2018. Network analysis of PTSD symptoms following mass violence. Psychol Trauma 10, 58–66. https://doi.org/10.1037/tra0000237

Tay, A. K., Rees, S., Steel, Z., Tam, N., Soares, Z., Soares, C., & Silove, D. M., 2016. Six-year trajectories of post-traumatic stress and severe psychological distress symptoms and associations with timing of trauma exposure, ongoing adversity and sense of injustice: a latent transition analysis of a community cohort in conflict-affected Timor-Leste. BMJ Open 6, e010205. https://doi.org/10.1136/bmjopen-2015-010205

Tylee, D.S., Chandler, S.D., Nievergelt, C.M., Liu, X., Pazol, J., Woelk, C.H., Lohr, J.B., Kremen, W.S., Baker, D.G., Glatt, S.J., Tsuang, M.T., Litz, B.T., Nash, W.P., Geyer, M.A., Hammer, P.S., Larsen, G.E., O'Connor, D.T., Risbrough, V.B., Schork, N.J., Vasterling, J.J., Webb-Murphy, J.A., 2015. Blood-based gene-expression biomarkers of post-traumatic stress disorder among deployed marines: A pilot study. Psychoneuroendocrinology 51, 472–494. https://doi.org/10.1016/j.psyneuen.2014.09.024

Vabalas, A., Gowen, E., Poliakoff, E., Casson, A.J., 2019. Machine learning algorithm validation with a limited sample size. PLoS One 14, e0224365. https://doi.org/10.1371/journal.pone.0224365

van den Broek, E.L., van der Sluis, F., Dijkstra, T., 2013. Cross-validation of bimodal health-related stress assessment. Pers. Ubiquitous Comput. 17, 215–227.

Vergyri, D., Knoth, B., Shriberg, E., Mitra, V., McLaren, M., Ferrer, L., Garcia, P., Marmara, C., 2015. Speech-Based Assessment of PTSD in a Military Population using Diverse Feature Classes, in: 16th Annual Conference of the International Speech Communication Association. Isca-Int Speech Communication Assoc, Baixas, pp. 3729–3733.

Wang, S., Hu, H., Su, S.S., Liu, L.Y., Wang, Z., Wang, Q., Shen, D.G., 2016. Combination of Grey Matter and White Matter Features for Early Prediction of Posttraumatic Stress Disorder, in: Intelligent Data Engineering and Automated Learning - Ideal 2016. Springer Int Publishing Ag, Cham, pp. 560–567.

Weathers, F.W., Bovin, M.J., Lee, D.J., Sloan, D.M., Schnurr, P.P., Kaloupek, D.G., Keane, T.M., Marx, B.P., 2018. The clinician-administered PTSD scale for DSM-5 (CAPS-5): Development and initial psychometric evaluation in military veterans. Psychol. Assess. 30, 383–395. https://doi.org/10.1037/pas0000486

Wortwein, T., Scherer, S., 2017. What Really Matters-An Information Gain Analysis of Questions and Reactions in Automated PTSD Screenings, in: 2017 Seventh International Conference on Affective Computing and Intelligent Interaction. Ieee, New York, pp. 15–20.

Yehuda, R., 2004. Risk and resilience in posttraumatic stress disorder, in: Journal of Clinical Psychiatry. 65(suppl 1), pp. 29–36

Yuan, M., Qiu, C., Meng, Y., Ren, Z., Yuan, C., Li, Y., Gao, M., Lui, S., Zhu, H., Gong, Q., Zhang, W., 2018. Pre-treatment Resting-State Functional MR Imaging Predicts the Long-Term Clinical Outcome After Short-Term Paroxtine Treatment in Post-traumatic Stress Disorder. Front Psychiatry 9, 532. https://doi.org/10.3389/fpsyt.2018.00532

Zandvakili, A., Philip, N.S., Jones, S.R., Tyrka, A.R., Greenberg, B.D., Carpenter, L.L., 2019. Use of machine learning in predicting clinical response to transcranial magnetic stimulation in comorbid posttraumatic stress disorder and major depression: A resting state electroencephalography study. J Affect Disord 252, 47–54. https://doi.org/10.1016/j.jad.2019.03.077

Zhang, Q., Wu, Q., Zhu, H., He, L., Huang, H., Zhang, J., Zhang, W., 2016. Multimodal MRI-based classification of trauma survivors with and without post-traumatic stress disorder. Front. Neurosci. 10, 292. https://doi.org/10.3389/fnins.2016.00292

SUPPLEMENTARY MATERIAL

1. Search strategy

We searched PubMed, Embase and Web of Science for articles published between January 1960 and May 2019 by using the following keywords: (“Big data” OR "Artificial Intelligence" OR "Machine Learning" OR "Gaussian process" OR "Cross-validation" OR "Cross validation" OR "Crossvalidation" OR "Regularized logistic" OR "Linear discriminant analysis" OR "LDA" OR "Random forest" OR "Naïve Bayes" OR "Naive Bayes" OR "Bayesian" OR "Least Absolute selection shrinkage operator" OR "Elastic net" OR "LASSO" OR "RVM" OR "Relevance vector machine" OR "Pattern recognition" OR "Computational Intelligence" OR "Computational Intelligences" OR "Machine Intelligence" OR "Knowledge Representation" OR "Knowledge Representations" OR "Support vector" OR "SVM" OR "Pattern classification") AND ("PTSD" OR "Stress Disorder, Post Traumatic" OR "Neuroses, Posttraumatic" OR "Posttraumatic Neuroses" OR "Posttraumatic Stress Disorders" OR "Posttraumatic Stress Disorder" OR "Stress Disorder, Posttraumatic" OR "Stress Disorders, Posttraumatic" OR "Neuroses, Post-Traumatic" OR "Neuroses, Post Traumatic" OR "Post-Traumatic Neuroses" OR "Post-Traumatic Stress Disorders" OR "Post Traumatic Stress Disorders" OR "Post-Traumatic Stress Disorder" OR "Stress Disorder, Post-Traumatic" OR "Delayed Onset Post-Traumatic Stress Disorder" OR "Delayed Onset Post Traumatic Stress Disorder" OR "Acute Post-Traumatic Stress Disorder" OR "Acute Post Traumatic Stress Disorder" OR "Chronic Post-Traumatic Stress Disorder" OR "Chronic Post Traumatic Stress Disorder" OR "Stress Disorders, Acute" OR "Acute Stress Disorder" OR "Stress Disorder, Acute" OR "Acute Stress Disorders").

Supplementary Table 1. Quality assessment domains

Feature	Considerations
1. Representativeness of the sample	Was the study truly representative of the target population heterogeneity? If not, was this related to the selected sampling method, insufficient sample size or inclusion/exclusion criteria?
2. Confounding variables	Did the study control for the most relevant confounding variables (age, gender, trauma type)? If so, were covariates assessed using subjective or objective measures?
3. Outcome Assessment	How were outcome measures assessed:

	A. Independent blind assessment (✓) B. Secure record (e.g. surgical records) (✓) C. Interview not blinded, self-report or medical record D. No description
4. Machine Learning Approach	Was the machine learning algorithm used to analyze data clearly described and appropriate?
5. Performance/Accuracy	Were the following performance metrics included: A. Accuracy B. Sensitivity C. Specificity D. AUC E. PPV/NPV
6. Missing Data	Did the study describe how the authors handled missing data, including if they were inputted or removed?
7. Testing/Validation	Was the test dataset "unseen" in regard to model training? Was the model tested on a hold-out or an external dataset?
8. Class Imbalance	Did the authors address the class imbalance problem? Which method was utilized?
9. Feature Selection and hyperparameter tuning	Did the study describe both feature selection and hyperparameter tuning? Which metrics were used?

Supplementary Table 2. Quality assessment of the studies

First author, year	Sample representativeness	Control confounding	Assessment of the outcome	ML algorithm	Performance metrics	Missing data	Test unseen	Class imbalance	Feature selection + hyperparam
Augsburger and Elbert, 2017	-	-	X	-	X	-	X	-	-
Banerjee et al., 2017	-	-	-	-	X	-	X	X	-
Bartels et al., 2019	X	-	-	X	-	-	-	-	-
Breen et al., 2019	-	X	X	X	X	X	-	X	X
Cisler et al., 2015	-	X	X	X	X	-	-	-	-
Conrad et al., 2017	X	X	X	X	X	X	X	-	X
Dabek et al., 2015a	X	-	-	X	X	-	-	-	X

Dabek et al., 2015b	X	-	-	X	X	-	X	-	X
Fried et al., 2018	X	-	-	X	-	X	-	-	-
Galatzer-Levy et al., 2014	X	-	-	-	X	X	-	-	X
Galatzer-Levy et al., 2017	X	-	X	X	X	X	-	X	X
Galovski et al., 2016	-	-	-	X	-	-	-	-	-
Gong et al., 2014	-	X	-	X	X	-	-	X	-
Gradus et al., 2017	X	X	-	X	X	X	-	-	-
Grisanzio et al., 2018	-	X	-	-	-	X	X	-	X
Harrington et al., 2019	X	X	-	X	X	X	-	X	X
He et al., 2017	X	X	-	X	X	-	-	X	X
Hemmings et al., 2017	-	X	-	X	-	-	-	X	X
Im et al., 2017	-	X	X	-	X	-	-	X	-
Jin et al., 2017	X	X	X	X	X	-	X	X	X
Karstoft et al., 2015b	-	X	-	X	X	X	-	-	X
Karstoft, 2015a	X	X	-	X	X	-	-	-	X
Kessler et al., 2014	X	X	X	X	X	-	-	-	-
Kobach et al., 2015	X	-	X	X	X	X	-	-	-
Leightley et al., 2019	X	-	-	-	-	X	X	-	X

Li et al., 2016	-	-	X	-	X	-	-	-	-
Liu et al., 2015	-	X	X	X	X	-	X	X	X
Ma et al., 2016	X	X	X	-	X	-	-	-	-
Marinić et al., 2007	-	X	X	X	X	-	-	X	-
Marmar et al., 2019	-	X	-	X	X	-	-	-	X
Mitchell et al., 2017	X	-	-	X	-	-	X	-	X
Nicholson et al., 2018	X	X	-	X	X	-	-	-	X
Omurca and Ekinci, 2015	X	-	-	-	X	-	-	-	X
Papini et al., 2018	X	X	-	X	X	X	-	-	X
Rangaprakash et al., 2017	-	X	-	X	X	-	X	-	X
Reece et al., 2017	X	-	-	X	X	-	X	-	-
Rosellini et al., 2018	X	-	-	X	X	-	-	-	X
Salminen, 2019	-	-	-	X	X	-	-	X	X
Saxe et al., 2017	X	-	-	X	X	X	X	-	X
Schalinski et al., 2016	-	X	X	X	X	-	X	-	-
Sullivan et al., 2018	X	X	-	X	-	-	-	-	-
Tylee et al., 2015	-	X	X	-	X	-	X	X	X
van der	-	X	-	X	X	-	-	-	X

Broek et al., 2013									
Vergyri et al., 2015	-	-	X	-	X	-	-	-	X
Wang et al., 2016	-	X	X	-	X	-	-	-	X
Wortwein and Scherer, 2017	-	-	-	X	X	-	X	-	X
Yuan et al., 2018	-	-	X	-	X	-	-	-	-
Zandvakili et al., 2019	-	-	-	X	X	-	-	-	-
Zhang et al., 2016	-	-	X	X	X	-	-	X	X

5.3 ARTICLE 3

PREDICTION OF POSTTRAUMATIC STRESS DISORDER STAGING FROM CLINICAL AND SOCIODEMOGRAPHIC FEATURES: A MACHINE LEARNING APPROACH (IN SUBMISSION REVIEW)

Luis Francisco Ramos-Lima [1,2],
Vitoria Waikamp [1,2],
Thauana Oliveira-Watanabe [3, 4],
Mariana Recamonde-Mendoza [5,6],
Stefania Pigatto Teche [1,2],
Marcelo Feijo Mello [3,4],
Andrea Feijo Mello [3,4],
Lucia Helena Machado Freitas [1,2]

[1] Post-graduate Program in Psychiatry and Behavioral Sciences, Universidade Federal do Rio Grande do Sul, Porto Alegre/RS, Brazil
[2] Psychological Trauma Research and Treatment Program (*NET-Trauma*), Hospital de Clínicas de Porto Alegre, Porto Alegre,/RS Brazil
[3] Post-graduate Program in Medical Psychology and Psychiatry, Department of Psychiatry, Universidade Federal de São Paulo, São Paulo/SP, Brazil
[4] Program for Research and Care on Violence and PTSD (PROVE), Universidade Federal de São Paulo, São Paulo/SP, Brazil
[5] Institute of Informatics, Universidade Federal do Rio Grande do Sul, Porto Alegre/RS, Brazil
[6] Bioinformatics Core, Hospital de Clínicas de Porto Alegre, Porto Alegre/RS, Brazil

MAIN PURPOSES:

- To develop a PTSD staging prediction from a machine learning approach
- To present a novel multi-class structure, instead of dichotomous approaches previously used in machine learning applications to PTSD diagnosis (article 2 results)
- To present examples of possible treatments related to PTSD staging

Abstract

Objective: this study aimed to verify the prediction power of machine learning (ML) techniques to support posttraumatic stress disorder (PTSD) staging.

Method: we performed a naturalistic, cross-sectional study at two Brazilian centers: the Psychological Trauma Research and Treatment (NET-Trauma) Program at Universidade Federal of Rio Grande do Sul, and the Program for Research and Care on Violence and PTSD (PROVE), at Universidade Federal of São Paulo. Five supervised learning algorithms were executed: Elastic Net, Gradient Boosting Machine, Random Forest, Support Vector Machine, and C5.0, using clinical (Clinician-Administered PTSD Scale for DSM-5 [CAPS-5], Brazilian-Portuguese version) and sociodemographic features.

Results: a hundred and twelve patients were enrolled in both centers (61 from NET-Trauma and 51 from PROVE) for the study. Best performance metrics were achieved using the C5.0 algorithm to CAPS-5 15-items plus sociodemographic features, with an accuracy of 65.6% ([CI 95% 52.3%–77.3%]; $p < .001$; K = .49; F1 score = .59) for the train data set and 52.9% ([CI 95% 38.5%–67.1%]; $p = .032$; K = .31; F1 score = .47) for the test data set. We found a model using four classes suitable for the PTSD staging. The number of symptoms, CAPS-5 total score, global severity score, and presence of current/previous trauma events appear as main features to predict PTSD staging.

Conclusion: Establishing the diagnosis of PTSD has been a challenge in clinical practice and academic research. This is the first study to evaluate staging in PTSD with ML algorithms using easily accessible clinical and sociodemographic features, which may be used in future research to aid in its early detection.

Keywords: Posttraumatic Stress Disorder; Machine Learning; Classification

Introduction

Posttraumatic stress disorder (PTSD) is considered a debilitating condition, developed from exposure to traumatic events including war, urban violence, and natural disasters. The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) lists 20 diagnostic criteria for PTSD divided into four clusters of symptoms: re-experience of the traumatic event, avoidance, persistent negative thoughts or feelings, and trauma-related arousal/reactivity[1]. PTSD presents social, biological, and mental health consequences, including social isolation, chronic pain and inflammation, cardiometabolic disorders, and heightened risk of dementia[2, 3]. Although most individuals experience a traumatic event

during life, most of the exposed people do not develop PTSD. The World Health Organization (WHO) World Mental Health Survey found a lifetime prevalence of any traumatic event of 70.4%; the prevalence of PTSD in a lifetime is estimated in 5%–10% of the population (twice as common in women as in men)[2], suggesting that constitutional, biological, and sociocultural factors may be involved in the development of the disorder[4].

The impact of traumatic stress exposure could be modeled as a continuous variable, with PTSD at the more severe end of the spectrum. Post-trauma neurobiological alterations can reach a point where they result in persistent or progressive illness[5]. Determining first-line treatments may be challenging: psychopathology of PTSD involves distinct genetic, endocrine, demographic, and environmental factors that are not shared by all PTSD patients, suggesting that efficient treatment may have to address individual-specific pathways[6, 7].

The concept of staging has been used for many years as a useful tool in complex diseases, such as diabetes mellitus, cardiovascular, and neoplastic conditions. Staging attempts to define the progression of disease at particular points in time through the continuum of the illness, helping to refine diagnosis, adjust prognosis, and in choosing the best treatment according to the illness stage[8]. Some areas of psychiatry are advanced in developing specific models of staging, such as bipolar disorder, unipolar depression, and schizophrenia[9]; regarding PTSD, McFarlane et al.[10] made the first attempt to propose a staging approach. This model was developed suggesting five stages to the disorder ranging from trauma-exposed, asymptomatic presentation to severe illness, with different clinical presentations and neurobiological alterations (see Supplementary Table 1, in the Supplementary material).

Machine learning (ML), a field of computer science and a part of artificial intelligence, refers to the science by which machines (i.e., computer systems) can analyze and acquire information from data. ML can help develop sophisticated data models using advanced mathematical and statistical techniques to handle complex data sets. A recent systematic review found 49 articles that used ML techniques to diagnose PTSD[11]; several types of data were used, such as neuroimaging, biological data, neuropsychological instruments, audio/video files, and clinical recordings. Nevertheless, those studies presented a dichotomous (“have/don’t have”) classification of the disorder, not contemplating the heterogeneity of PTSD presentation.

This study aims to investigate the viability and the predictive power of ML models to support a staging approach to PTSD, based on the model proposed by McFarlane et al.[10], using clinical and sociodemographic data from an outpatient environment. Considered to be a

disorder with genetic and environmental influences, a heterogeneous population, with different prognoses, PTSD appears to be an important example of where staging could provide benefit.

Methods

Study design and participants

This naturalistic, cross-sectional study was developed at two centers: the Psychological Trauma Research and Treatment Program (NET-Trauma), from the Psychiatric Service of the Hospital de Clínicas de Porto Alegre (HCPA) / Universidade Federal do Rio Grande do Sul (UFRGS), and the Program for Research and Care on Violence and PTSD (PROVE), from the Department of Psychiatry at the Universidade Federal de São Paulo (UNIFESP).

Patients were enrolled from the screening processes of the NET-Trauma outpatient service between August 2018 and January 2020 (Sample 1) and from the PROVE outpatient service between January 2016 and March 2019 (Sample 2). The participants spontaneously sought psychiatric assistance after experiencing a traumatic event and all agreed to participate in the study. No exclusion criteria were applied. This study was approved by the Ethics Committees of the UFRGS and UNIFESP centers, and all participants signed the informed consent form.

Study protocol and staging

We used the Clinician-Administered PTSD Scale, DSM-5 edition (CAPS-5) adapted to the Brazilian-Portuguese language[12]; it evaluates the diagnostic of PTSD and the severity of the symptoms related to DSM-5 criteria. We also used a reduced version of CAPS-5 containing 15 items[13], obtained after validation of CAPS-5 Brazilian-Portuguese version. In order to allow a comparison of the prediction power of the Brazilian versions, we chose to keep both scales.

The items were assessed using a Likert scale, where 0 is considered "Absent" and 4 is considered "Extreme/incapacitating." We used the Brazilian-Portuguese Life Events Checklist, DSM-5 edition (LEC-5) to ensure the occurrence of a traumatic event (Criteria A)[14]. We also asked for sociodemographic information, previous trauma events, and present and past psychiatric comorbidities (according to clinical evaluation, based on DSM-5). We assessed the reliability of the individual's response by providing the interviewer with a specific question (CAPS-5 Item 26: Global Validity); no patients were excluded from that

criterion. Study data were collected and managed using REDCap electronic data capture tools hosted at HCPA and at PROVE/UNIFESP.

The PTSD staging model is composed of five stages[10]: Stage 0 includes trauma-exposed individuals, who are asymptomatic but at risk to develop PTSD. Stage 1 includes individuals who do not reach a full diagnosis. Stage 2 is related to the first episode of full symptoms, whereas Stage 3 refers to persistent symptoms that are relatively long-standing. Stage 4 refers to a severe clinic presentation where symptoms lasted a prolonged period. Staging was provided by agreement of two psychiatrists at the time of the first evaluation at the NET-Trauma and PROVE centers. Staging of patients from the PROVE center who initiated treatment before 2018 was provided by agreement of two psychiatrists after a later review of medical records from the first consultation. All psychiatrists had expertise in PTSD diagnosis and management. In order to determine the patient stage, they considered clinical presentations described by the staging structure (see Supplementary Table 1). Classification algorithms tend to perform better with a reduced number of classes, especially in smaller samples; in order to guarantee a better use of these ML techniques and a better interpretation of the results, we merged the second-level status from Stages 1 and 3 (1a and 1b; 3a, 3b, and 3c) to the first level. Since we did not receive any asymptomatic patients, Stage 0 was not considered.

Machine learning algorithms

A wide range of ML methods have been proposed to cover the existing variety of data and problem types. Kessler et al.[15] indicated random forest as the best model to predict PTSD after trauma exposure. Other studies indicated support vector machines (SVM) as suited to data with heterogeneous, non-normal distributions consistent with the kinds of data that PTSD researchers are interested in integrating for prediction[11, 16]. We chose five supervised learning algorithms with distinct inductive bias to train our models: Elastic Net, gradient boosting machine (GBM), random forest, radial-basis function (RBF) SVM kernel, and a cost-sensitive C5.0 (with a cost matrix to penalize misclassification). A brief description of each ML technique is presented in Supplementary Table 2. A revision of the relevant principles of ML and its limitations can be found elsewhere[17].

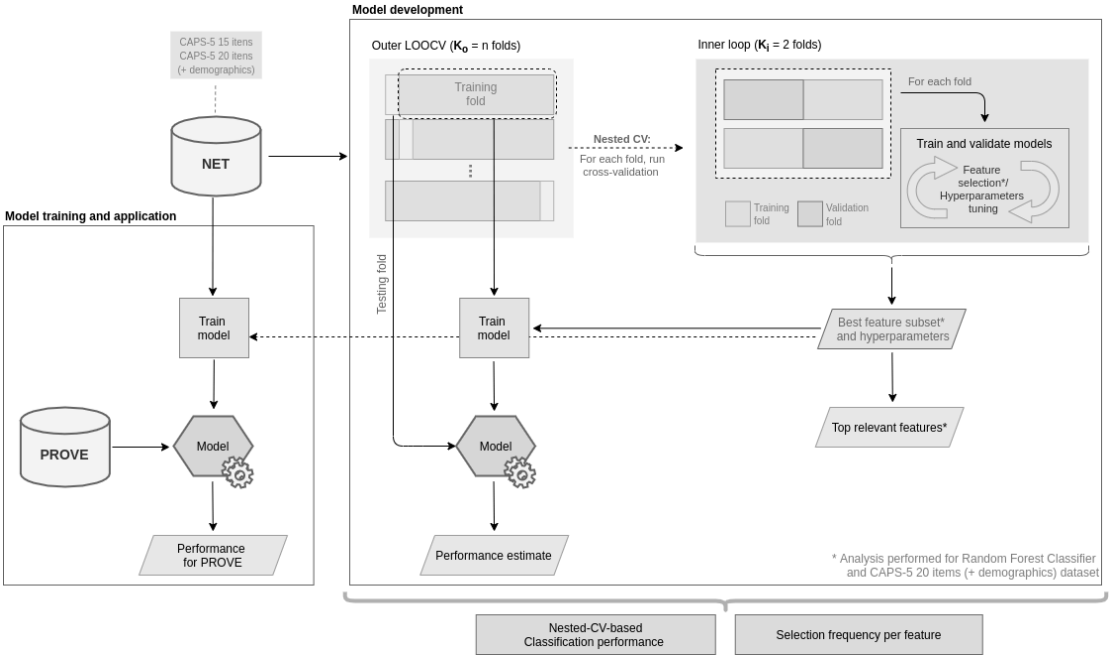
Model development

We developed four groups of models for comparison: (1) the 20-item CAPS-5; (2) the 15-item CAPS-5; (3) the 20-item CAPS-5 plus clinical and demographic data; and (4) the 15-

item CAPS-5 plus clinical and demographic data. For each group, all ML techniques were applied.

Typically, ML algorithms are implemented in two key stages[18]: first, all data are separated into "training" and "testing" sets with the former being used to "train" the ML model. Second, the performance of the model is evaluated using the "testing" set that consists (ideally) of observations previously "unseen" by the model. We chose Sample 1 to train the learning model because of the heterogeneity of the data and the presence of more cases distributed between all four stages. Sample 2 was used to evaluate the performance of the models in a fully independent set of individuals (from a similar population) to ensure external validity[19, 20]. A nested cross-validation (CV) strategy was employed to obtain an unbiased estimation of the true generalization performance, to avoid data leakage, and to provide robust parameter estimates particularly for smaller samples. We chose the down-sampling strategy as the most effective method to deal with the class imbalance in our data set. Details about methodology and sampling are described in Supplementary Item 1. Figure 1 illustrates the ML model development.

Figure 1. Model Development and Application



Although we used a fixed set of features in our analysis, we also employed a variable selection method named recursive feature elimination (RFE) implemented on the random forest algorithm to analyze top relevant features and study the predictive power of CAPS-5 20-item version.

Performance metrics

The performance results were reported in terms of area under the receiver operating characteristic curve (AUC), sensitivity, specificity, accuracy, F1 score and Cohen's Kappa score. AUC metrics were obtained by computing the macro-averaging across all four classes (Supplementary Item 2 contains a detailed description) for all algorithms except C5.0 (the use of a matrix cost with unequal values may turn class probabilities not consistent with final predicted classes, making AUC calculation unviable). F1 score takes both false positives and false negatives into account, considered a useful metric in multiclass classification with unbalanced classes. Cohen's Kappa statistics are widely used in multiclass classification to inform how well the classifier is performing compared with a random classifier: values less than 0 indicate no agreement; 0–.20 as slight, .21–.40 as fair, .41–.60 as moderate, .61–.80 as substantial, and .81–1 as almost perfect agreement[21].

Statistical analysis and ML implementation were performed using the software R version 3.6.3. Prior to any analyses, items from CAPS-5 were standardized using the *z*-score normalization for each data set separately. When necessary, categorical variables were transformed into a set of binary variables using one-hot encoding. There were no missing values in either sample. Data pre-processing and ML techniques were implemented using the *caret* package version 6.0-85 (see Supplementary Table 3). *P* values displayed for the accuracy metrics were obtained automatically through the binom.test function; it tests if the model's accuracy is better than the proportion of the data with the majority class ("no-information rate"). We followed the recommendations for ML-based studies proposed by Passos et al.[20] to describe our methodology and results (see Supplementary Table 4). The ML programmer did not participate in the enrollment and data collection phases.

Results

We included 112 patients from both centers in the study; the main profile was composed of female patients of Caucasian ethnicity, single, and reporting religiousness. Apart from educational level and employment status, none of the other sociodemographic characteristics of patients in either sample differed significantly. Half of the patients were victims of sexual assault (50.9%) and were similarly distributed between centers. The CAPS scores were similar between both centers: median 42.0 [CI 95% 25.0–48.5] from Sample 1 and 41.0 [CI 95% 35.0–47.0] from Sample 2; $p = .453$. Stages 1 to 4 represented 17.0%, 36.6%, 36.6%, and 9.8%, respectively, among all participants, maintaining similar

distributions in each center ($p = .958$). Table 1 presents the clinical and sociodemographic characteristics of participants.

Table 1. Clinical and sociodemographic information from the total sample ($n = 112$)

Variable	NET (n=61) (Sample 1)	PROVE (n=51) (Sample 2)	All	*P* value
Age, median [IQR]	35.0 [20.5-47.5]	28.0 [24.0-35.0]	32.0 [23.0-45.0]	.088
Female gender, No. (%)	52 (85.2%)	46 (90.2%)	98 (87.5%)	.430
Ethnicity, No. (%)				
Caucasian	41 (67.2%)	27 (52.9%)	68 (60.7%)	.303
African-american	19 (31.1%)	23 (45.1%)	42 (37.5%)	
Asiatic-american	1 (1.6%)	1 (2.0%)	2 (1.8%)	
Marital status, No. (%)				
Single	25 (41.0%)	28 (54.9%)	53 (47.3%)	.249
Married/engaged	23 (37.7%)	17 (33.3%)	40 (35.7%)	
Divorced/widower	13 (21.3%)	6 (11.8%)	19 (17.0%)	
Educational level, No. (%)				
Up to 4 years	1 (1.6%)	2 (3.9%)	3 (2.7%)	.049[a]
4 to 12 years	38 (62.3%)	20 (39.2%)	58 (51.8%)	
Above 12 years	22 (36.1%)	29 (56.9%)	51 (45.5%)	
Employment status, No. (%)				
Employed	22 (36.1%)	30 (58.8%)	52 (46.4%)	.028[a]
Unemployed/retired	28 (45.9%)	18 (35.3%)	46 (41.1%)	
Health licensed	11 (18.0%)	3 (5.9%)	14 (12.5%)	
Religiousness, No. (%)	43 (70.5%)	31 (60.8%)	74 (66.1%)	.280
Living alone, No. (%)	3 (4.9%)	6 (11.8%)	9 (8.0%)	.184
Current psychiatric comorbidity, No. (%)	34 (55.7%)	28 (54.9%)	62 (55.4%)	.929
Previous psychiatric disorders, No. (%)	28 (45.9%)	25 (49.0%)	53 (47.3%)	.742
Previous trauma event, No. (%) [b]	34 (55.7%)	29 (56.9%)	63 (56.3%)	.905
Types of trauma (LEC-5), No. (%)				
Fire or explosion	1 (1.6%)	0 (0%)	1 (.9%)	.190
Transportation accident	2 (3.3%)	0 (0%)	2 (1.8%)	
Physical assault	3 (4.9%)	3 (5.9%)	6 (5.4%)	
Assault with a weapon	9 (14.8%)	9 (17.6%)	18 (16.1%)	
Sexual assault	27 (44.3%)	30 (58.8%)	57 (50.9%)	
Captivity	0 (0%)	3 (5.9%)	3 (3.7%)	
Life-threatening illness or injury	4 (6.6%)	0 (0%)	4 (3.6%)	
Severe human suffering	4 (6.6%)	2 (3.9%)	6 (5.4%)	
Sudden violent death	7 (11.5%)	3 (5.9%)	10 (8.9%)	
Sudden acidental death	3 (4.9%)	1 (2.0%)	4 (3.6%)	
Serious injury, harm, or death you caused to someone else	1 (1.6%)	0 (0%)	1 (.9%)	
CAPS score, median [IQR]	42.0 [25.0-48.5]	41.0 [35.0-47.0]	41.5 [31.2-47.0]	.453
Staging, No. (%)				
1 (Undifferentiated symptoms)	11 (18.0%)	8 (15.7%)	19 (17.0%)	.958

2 (First full-episode)	23 (37.7%)	18 (35.3%)	41 (36.6%)
3 (Persistent symptoms)	21 (34.4%)	20 (39.2%)	41 (36.6%)
4 (Severe, unremitting illness)	6 (9.8%)	5 (9.8%)	11 (9.8%)

[a] $p<0.05$ [b]Excluding the current event

Prediction of PTSD staging

Models developed with CAPS-5 15-item+demographics provided most of the highest performance rates among the four data groups. Using the C5.0 algorithm, this group generated the highest accuracy rates: 65.6% ([CI 95% 52.3%–77.3%]; $p < .001$; K = .49; sensitivity = 58.6%; specificity = 87.0%; F1 score = .59) for the train set and 52.9% ([CI 95% 38.5%–67.1%]; $p = .032$; K = .31; sensitivity = 44.8%; specificity = 82.8%; F1 score = .47) for the test set. Using Elastic Net, this group also generated the best overall sensitivity (58.1%) and the best specificity rates (83.1%) for the test set. The best AUC scores for the train set were also obtained within this data group (68.6%–76.6%), the highest score reached using the random forest algorithm. The remaining data groups provided models with lower rates. The performance metrics of each ML algorithm are presented in Table 2.

Feature selection

The items caps-D5 (diminished interests) (98.4%), caps-D6 (feelings of detachment) (90.2%), caps-E1 (irritability) (86.9%), caps-B4 (negative states) (85.2%), and caps-E4 (startle response) (82.0%) were the most selected attributes. Applying RFE to the CAPS-5 20-item+demographics produced complementary results: the five most selected attributes were current psychiatric comorbidity (96.7%), global severity score (96.7%), number of symptoms (90.0%), CAPS-5 score (80.3%), and item caps-D5 (Diminished interests). RFE did not provide significantly better performance than the conventional random forest algorithm. Full RFE results are presented in Supplementary Table 5.

Class imbalance

We performed an exploratory analysis (using principal component analysis) and generated a dispersion diagram showing the four groups of data (see Supplementary Figure 2). We verified that Stages 1 and 4 are more easily identifiable; Stages 2 and 3 appear to have more homogeneity. We executed the ML algorithms with the best results (C5.0, random forest, and Elastic Net) using the data set with the best performance metrics (15-item+demographics) considering three classes: Stage 1, Stage 4, and Stages 2 and 3 combined.

Table 2. Overall results for each machine learning technique (Train/Test Sets)

Model/algorithm	AUC (%) $		Sensitivity (%)		Specificity (%)		Accuracy [CI 95%]		Kappa value		F1 score		*P* (Acc)	
	Train	Test	Train	Test	Train	Train	Train	Test	Train	Test	Train	Test	Train	Test
CAPS-5 20-item														
C5.0	NA	NA	33.5%	26.0%	76.3%	74.8%	34.4% [22.7%-47.7%]	35.3% [22.4%-49.9%]	.07	.00	.33	.25	.743	.762
GBM	73.7%	61.4%	56.3%	33.8%	81.7%	75.5%	44.3% [31.5%-57.6%]	33.3% [20.8%-47.9%]	.27	.04	.44	.33	.177	.842
Elastic Net	61.3%	65.1%	35.1%	31.4%	78.9%	74.0%	36.1% [24.2%-49.4%]	19.6% [9.8%-33.1%]	.15	-.01	.33	.22	.650	.999
RF	70.2%	61.6%	50.9%	30.8%	80.9%	75.5%	41.0% [28.6%-54.3%]	31.4% [19.1%-45.9%]	.24	.04	.41	.31	.343	.903
RF + RFE	71.2%	59.8%	47.5%	33.9%	79.9%	76.6%	37.7% [25.6%-51.0%]	33.3% [20.7%-47.9%]	.20	.08	.38	.34	.548	.842
RBF-SVM	65.8%	52.8%	34.3%	33.3%	76.0%	77.5%	32.8% [21.3%-46.0%]	39.2% [25.8%-53.9%]	.07	.11	.30	.35	.822	.553
CAPS-5 15-item														
C5.0	NA	NA	39.0%	20.3%	76.3%	71.5%	36.1% [24.2%-49.4%]	21.6% [11.3%-35.3%]	.09	-.12	.39	.19	.650	.998
GBM	63.0%	57.2%	42.8%	27.9%	77.2%	75.5%	29.5% [18.5%-42.6%]	35.3% [22.4%-49.9%]	.11	.03	.29	.28	.929	.762
Elastic Net	65.8%	64.1%	40.4%	30.7%	79.8%	75.6%	39.3% [27.1%-52.7%]	23.5% [12.8%-37.5%]	.19	.03	.38	.26	.443	.994
RF	68.5%	57.6%	46.2%	36.8%	79.5%	77.5%	34.4% [22.7%-47.7%]	37.3% [24.1%-51.9%]	.18	.11	.34	.34	.743	.663
RBF-SVM	65.4%	56.5%	37.7%	32.6%	77.5%	77.3%	37.7% [25.6%-51.0%]	37.3% [24.1%-51.9%]	.12	.10	.33	.34	.548	.663
CAPS-5 20-item + demographics														
C5.0	NA	NA	61.1%	43.4%	85.4%	82.0%	60.7% [47.3%-72.9%]	51.0% [36.6%-65.2%]	.43	.28	.61	.46	<.001*	.059
GBM	66.2%	70.1%	49.1%	48.5%	79.3%	77.7%	36.1% [24.2%-49.4%]	39.2% [25.8%-53.9%]	.18	.14	.36	.45	.650	.553
Elastic Net	67.5%	70.5%	44.5%	42.0%	81.1%	79.0%	42.6% [30.0%-55.9%]	37.3% [24.1%-51.9%]	.23	.16	.42	.38	.253	.663
RF	70.5%	73.8%	44.1%	54.2%	78.7%	81.2%	36.1% [24.2%-49.4%]	47.1% [32.9%-61.5%]	.16	.27	.36	.49	.650	.158
RF + RFE	75.6%	68.0%	59.3%	45.3%	83.5%	81.9%	49.2% [36.1%-62.3%]	51.0% [36.6%-65.2%]	.33	.28	.50	.48	.044*	.058
RBF-SVM	68.7%	64.6%	34.3%	39.7%	77.3%	77.1%	34.4% [22.7%-47.7%]	37.3% [24.1%-51.9%]	.10	.11	.30	.40	.743	.663
CAPS-5 15-item + demographics														
C5.0	NA	NA	58.6%	44.8%	87.0%	82.8%	65.6% [52.3%-77.3%]	52.9% [38.5%-67.1%]	.49	.31	.59	.47	<.001*	.032*
GBM	68.6%	72.0%	51.4%	45.6%	80.2%	81.0%	39.3% [27.1%-52.7%]	51.0% [36.6%-65.2%]	.21	.26	.40	.48	.443	.059
Elastic Net	71.5%	70.4%	50.0%	58.1%	82.3%	83.1%	45.9% [33.1%-59.2%]	47.1% [32.9%-61.5%]	.28	.30	.45	.47	.118	.158
RF	76.6%	72.4%	50.6%	54.3%	80.7%	80.5%	41.0% [28.6%-54.3%]	47.1% [32.9%-61.5%]	.23	.25	.41	.49	.343	.158
RBF-SVM	70.8%	63.9%	40.9%	37.8%	78.7%	77.6%	39.3% [27.1%-52.7%]	39.2% [25.8%-53.9%]	.16	.12	.36	.39	.443	.553

[a] $p<.05$ [b] AUC scores are not available for C5.0 since its implementation returns only predicted labels and not class probabilities. GBM: Gradient Boost Machine; RF: Random Forest; RF + RFE: Random Forest with Recursive Feature Extraction; RBF-SVM: Support Vector Machine, Radial Basis Function

C5.0 achieved an accuracy of 82.0% ([CI 95% 70.0%–90.6%]; $p = .054$; K = .56; F1 score = .65; sensitivity = 63.1%; specificity = 83.8%) for the train set.For the test set, accuracy was 70.6% ([CI 95% 56.2%–82.5%]; $p = .792$; K = .22; F1 score = .51; sensitivity = 47.2%; specificity = 72.6%). All other executions provided worse performance metrics. Classification performances for each stage are presented in Supplementary Table 6.

Discussion

The present study evaluated the prediction power of ML techniques in PTSD research to help differentiate PTSD among individuals exposed to trauma. Most of the features presented similar distributions between the NET-Trauma and PROVE centers (see Table 1). We highlight the majority of sexual assaults among the types of trauma reported by patients: the worldwide prevalence of sexual violence is 14.0%, but it is unequally distributed among countries[22]. Considering a low- to middle-income country such as Brazil, a higher proportion of sexual harassment victims by social permissiveness and violent behavior against women is presumed[23]. Studies observing regional risk factors and prevalence of trauma types are lacking.

We observed the best prediction performances using the data set CAPS-5 15-item+demographics. At least half of the PTSD prognostic studies using dichotomous outcomes presented accuracy under 80%[11], indicating that classification of PTSD is not a simple task. Our best result for a four-class outcome was 65.6%, and for a three-class outcome was 82.0% (train set). In order to assure generalization and external validity, we applied the resultant model to an "unseen" data set (test set) obtaining accuracy of 52.9% for the four-class model, and 70.6% for the three-class model, configuring good results for this unprecedented attempt at a multiclass approach. We obtained better accuracy with a three-class model (combining Stages 2 and 3); we consider that extremes of the PTSD spectrum are easier to detect, while Stages 2 and 3 need more clarification to ensure that they are different phenotypes of the disorder as the original staging structure indicates.

Taking this into consideration, we performed a detailed analysis of the four-class staging results from the test set (PROVE center): Stage 1, related to the "non-PTSD" patients, presented fewer symptoms and lower CAPS-5 score compared with the three "PTSD stages." Stages 2 and 3 presented a proximal CAPS-5 score median, reflecting equivalent severity of symptoms; Stage 3 may differ from Stage 2 by presenting patients with more previous and current psychiatric disorders, any drug use, and the presence of an earlier trauma event. Stage 3 patients also received higher global severity scores compared with Stage 2. All Stage 4

patients presented current psychiatric comorbidities, received the rating "4" more frequently on global severity evaluation, and appeared to have less of a support network as suggested by the "living alone" variable. Stage 4 also presented a higher CAPS-5 score compared with Stages 2 and 3, consolidating it as a different, more severe group.

The presence of current psychiatric comorbidity is one of the most relevant features in four of the five algorithms; a CAPS-5 total score appears in three of them. Age, educational level (4–12 years of study), alcohol use, the number of symptoms, and the global severity score also appear as relevant features (see Supplementary Tables 5 and 7). Of note, essential symptoms such as dissociative states and sleep disturbances did not appear as most relevant in differentiating among stages (full results are presented in Supplementary Table 8). Regarding CAPS-5 items, symptoms from different 'clusters' appear as most selected by RFE: diminished interests (caps-D5), feelings of detachment (caps-D6), negative states (caps-B4), and even irritability (caps-E1) suggest that a depressive state may also follow trauma. Co-occurrence of depression is common and associated with greater severity and impairment than PTSD alone[24]. We hypothesize that Stage 3 differentiates from Stage 2 by a different profile of symptoms related to these depressive states. Stage 3 is also characterized by the presence of previous trauma history, which supports the hypothesis that pre-trauma factors (such as previous trauma history and childhood trauma) could potentiate post-trauma alterations, leading to a persistent or progressive PTSD[5, 25]. We compiled the main results of our study considering the four-class model, along with examples of possible interventions for each stage, presented in Table 3.

Table 3. PTSD Stages, characteristics from ML models of this study, and possible treatments.

PTSD Stages	Principal characteristics	Examples of possible treatments
Stage 1 ("non-PTSD") "Preclinical stage"	• Trauma-exposed patients • Presence of previous trauma event • Low global severity score (0-1) • Lower CAPS-5 score (median 20.0) • Lower symptoms (median 6.0)	• Early interventions • Psychoeducation strategies • Goals: to detect ASD, to prevent PTSD
Stage 2 ("PTSD") "Less severe PTSD" "First episode?"	• Global severity score 2-3 • CAPS-5 score ≈ 40 • # symptoms ≈ 14	• Evidence-based, manualized therapies (CBT) focused on trauma • Consider associate pharmacotherapy to treat specific symptoms • Goal: to prevent chronic illness

Stage 3 "More severe PTSD"	• Presence of current and/or previous psychiatric comorbidities • More presence of previous trauma events than Stage 2 • Drug use • Global severity score 3 • CAPS-5 score ≈ 40 • # symptoms ≈ 15	• Evidence-based, manualized therapies (CBT) focused on trauma, longer periods • Associate pharmacotherapy for a longer period
Stage 4 "Chronic PTSD" "Unremitting PTSD	• Global severity score 4 • CAPS-5 score ≈ 50 • # symptoms ≈ 15 (more severe) • Presence of current psychiatric comorbidity	• Consider also long-term psychotherapies [a] • Consider also adjuvant pharmacotherapy [a] • Treat clinical comorbidities

Adapted from Bisson et al.[26] and Ostacher & Cifu[27].

[a] Insufficient evidence to recommend, according to Bisson et al.[26]

Although we observed important differences among Stages 2 to 4 in the multiclass approach, we also tested a dichotomous model with the best results (CAPS-5 15-item+demographics) and algorithm (C5.0), considering Stage 1 as the "non-PTSD" and Stages 2, 3, and 4 as the "PTSD" category. We obtained an accuracy of 90.1% ([CI 79.8%–96.3%]; $p = .059$; K = .71; sensitivity = 90.0%; specificity = 90.9%; F1 score = .93) for the train set, and 90.2% ([CI 78.6%–96.7%]; $p = .168$; K = .61; sensitivity = 95.3%; specificity = 62.5%; F1 score = .94) for the test set, confirming our model to be an excellent predictor, comparable with previous studies on PTSD[11].

One limitation of our study was the cross-sectional design. We understand that longitudinal studies are lacking to correlate neurobiological changes and indicate specific treatments for each clinical stage. For instance, delayed-onset PTSD is often preceded by subsyndromal symptoms, which also impact on morbidity and may be predictors of a full syndrome[28]. Of note, we have not considered the Complex PTSD (C-PTSD) diagnosis in our research, which was not detected using CAPS-5. We hypothesize that "Stage 4" may also include C-PTSD; further studies are necessary to confirm this.

Two other crucial methodological challenges are class imbalance and small sample sizes. Most of the algorithms alternated between favoring classes with a higher proportion of cases (Stages 2 and 3) or favoring classes with rare occurrences (Stages 1 and 4), altering overall accuracy. Our results initially suggested that there would be considerable homogeneity between Stages 2 and 3; a detailed revision of Stages 2 and 3 provided relevant differences. Small sample sizes may cause a higher variance in results, and impair pattern recognition due

to overfitting (i.e., when a model fits better to the train set than to new data sets)[29]. We compared the accuracies obtained in the training data: outer CV results are within or above the interquartile range of inner CV, indicating robustness in performance evaluation and reducing the possibility of overfitting (see Supplementary Figure 3). Nevertheless, we suggest that replicating the staging model in other PTSD treatment centers with larger sample sizes could better elucidate these issues and consequently improve accuracy.

We considered a strength of our study to be the use of clinical (through a validated PTSD scale, CAPS-5) and sociodemographic information to predict PTSD staging applying the advantages of ML techniques. 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[30]. To date, there is no established protocol to determine staging, so we still rely on specialist evaluation as a "gold standard." This study attempts to provide an objective method to enrich staging prediction.

Conclusions

The psychopathological heterogeneity of PTSD can potentially reduce the power and obscure the findings of clinical studies, as it is known that there are several limitations to using traditional hypothesis testing methods to analyze multidimensional and heterogeneous data[31]. Although this staging model may allow a framework for examining different clinical and biological models for PTSD and how they overlap, developing a strategy for interventions at different stages of PTSD has become a critical need. Staging links the clinical aspects to treatment selection and prediction: early successful treatments may change the prognosis and thus prevent progression to subsequent stages[32]. ML techniques can be advantageous and have increasingly been used in precision psychiatry, as they can be focused at an individual patient level[33]. The prevalence and the burden of diseases associated with PTSD highlight the importance of screening in the clinical setting, including primary care, and in heterogeneous populations at risk for trauma exposure. Future research may include other classes of information, such as biomarkers and neuroimaging studies, to adjust and detail the phenotypic presentations of PTSD in order to provide more targeted treatments.

Declaration of conflicting interests: The authors declare that there is no conflict of interest.

Funding

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Brazil, Finance Code 001) and by the Conselho Nacional de Pesquisa (CNPq, Brazil).

References

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th. ed. Washington (DC). 2013.
2. Yehuda R, Hoge CW, McFarlane AC, et al. Post-traumatic stress disorder. Nat Rev Dis Primers. 2015 Oct 8;1:15057. doi: 10.1038/nrdp.2015.57.
3. Song H, Sieurin J, Wirdefeldt K, et al. Association of Stress-Related Disorders With Subsequent Neurodegenerative Diseases. JAMA Neurol. 2020 Jun 1;77(6):700-709. doi: 10.1001/jamaneurol.2020.0117.
4. Benjet C, Bromet E, Karam EG, et al. The epidemiology of traumatic event exposure worldwide: results from the World Mental Health Survey Consortium. Psychol Med. 2016 Jan;46(2):327-43. doi: 10.1017/S0033291715001981.
5. Koek RJ, Schwartz HN, Scully S, et al. Treatment-refractory posttraumatic stress disorder (TRPTSD): a review and framework for the future. Prog Neuropsychopharmacol Biol Psychiatry. 2016 Oct 3;70:170-218. doi: 10.1016/j.pnpbp.2016.01.015.
6. Galatzer-Levy IR, Ma S, Statnikov A, et al. Utilization of machine learning for prediction of post-traumatic stress: a re-examination of cortisol in the prediction and pathways to non-remitting PTSD. Transl Psychiatry. 2017 Mar 21;7(3):e0. doi: 10.1038/tp.2017.38.
7. Ramos-Lima LF, Souza PRA, Teche SP, et al. Trauma-related disorders in a low- to middle-income country: A four-year follow-up of outpatient trauma in Brazil. Psychiatry Res. 2019 Oct;280:112525. doi: 10.1016/j.psychres.2019.112525.
8. Kapczinski F, Magalhães PV, Balanzá-Martinez V, et al. Staging systems in bipolar disorder: an International Society for Bipolar Disorders Task Force Report. Acta Psychiatr Scand. 2014 Nov;130(5):354-63. doi: 10.1111/acps.12305.
9. de la Fuente-Tomas L, Sánchez-Autet M, García-Álvarez L, et al. Clinical staging in severe mental disorders; bipolar disorder, depression and schizophrenia. Rev Psiquiatr Salud Ment. 2019 Apr-Jun;12(2):106-115. English, Spanish. doi: 10.1016/j.rpsm.2018.08.002.

10. McFarlane AC, Lawrence-Wood E, Van Hooff M, et al. The Need to Take a Staging Approach to the Biological Mechanisms of PTSD and its Treatment. Curr Psychiatry Rep. 2017 Feb;19(2):10. doi: 10.1007/s11920-017-0761-2.
11. Ramos-Lima LF, Waikamp V, Antonelli-Salgado T, et al. The use of machine learning techniques in trauma-related disorders: a systematic review. J Psychiatr Res. 2020 Feb;121:159-172. doi: 10.1016/j.jpsychires.2019.12.001.
12. Oliveira-Watanabe TT, Ramos-Lima LF, Santos RC, et al. The Clinician-Administered PTSD Scale (CAPS-5): adaptation to Brazilian Portuguese. Braz J Psychiatry. 2019 Jan-Feb;41(1):92-93. doi: 10.1590/1516-4446-2018-0136.
13. Oliveira-Watanabe TT. Reliability and internal consistency of Clinician-Administered Posttraumatic Scale (CAPS-5) in Brazilian-Portuguese [Unpublished master thesis]. São Paulo: Universidade Federal de São Paulo, 2020.
14. Lima EP, Vasconcelos AG, Berger W, et al. Cross-cultural adaptation of the Posttraumatic Stress Disorder Checklist 5 (PCL-5) and Life Events Checklist 5 (LEC-5) for the Brazilian context. Trends Psychiatry Psychother. 2016 Oct-Dec;38(4):207-215. doi: 10.1590/2237-6089-2015-0074.
15. Kessler RC, Rose S, Koenen KC, et al. How well can post-traumatic stress disorder be predicted from pre-trauma risk factors? An exploratory study in the WHO World Mental Health Surveys. World Psychiatry. 2014 Oct;13(3):265-74. doi: 10.1002/wps.20150.
16. Galatzer-Levy IR, Karstoft KI, Statnikov A, et al. Quantitative forecasting of PTSD from early trauma responses: a Machine Learning application. J Psychiatr Res. 2014 Dec;59:68-76. doi: 10.1016/j.jpsychires.2014.08.017.
17. Schultebraucks K, Galatzer-Levy IR. Machine Learning for Prediction of Posttraumatic Stress and Resilience Following Trauma: An Overview of Basic Concepts and Recent Advances. J Trauma Stress. 2019 Apr;32(2):215-225. doi: 10.1002/jts.22384.
18. Passos IC, Mwangi B, Cao B, et al. Identifying a clinical signature of suicidality among patients with mood disorders: A pilot study using a machine learning approach. J Affect Disord. 2016 Mar 15;193:109-16. doi: 10.1016/j.jad.2015.12.066.
19. Baştanlar Y, Özuysal M. Introduction to machine learning. In: Yousef, M., Allmer, J. (Eds.), miRNomics: MicroRNA Biology and Computational Analysis. Methods in Molecular Biology (Methods and Protocols), vol. 1107 Humana Press, Totowa, NJ. 2014. doi:10.1007/978-1-62703-748-8_7.

20. Passos IC, Ballester PL, Barros RC, et al. Machine learning and big data analytics in bipolar disorder: A position paper from the International Society for Bipolar Disorders Big Data Task Force. Bipolar Disord. 2019 Nov;21(7):582-594. doi: 10.1111/bdi.12828.
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics, 1977; 33(1), 159–174.
22. Kessler RC, Aguilar-Gaxiola S, Alonso J, et al. Trauma and PTSD in the WHO World Mental Health Surveys. Eur J Psychotraumatol. 2017 Oct 27;8(sup5):1353383. doi: 10.1080/20008198.2017.1353383.
23. Both LM, Favaretto TC, Freitas LHM, et al. Intimate partner violence against women: Operationalized Psychodynamic Diagnosis (OPD-2). PLoS One. 2020 Oct 1;15(10):e0239708. doi: 10.1371/journal.pone.0239708.
24. Kline AC, Cooper AA, Rytwinski NK, et al. The Effect of Concurrent Depression on PTSD Outcomes in Trauma-Focused Psychotherapy: A Meta-Analysis of Randomized Controlled Trials. Behav Ther. 2021 Jan;52(1):250-266. doi: 10.1016/j.beth.2020.04.015.
25. Kostaras P, Bergiannaki JD, Psarros C, et al. Posttraumatic stress disorder in outpatients with depression: Still a missed diagnosis. J Trauma Dissociation. 2017 Mar-Apr;18(2):233-247. doi: 10.1080/15299732.2016.1237402.
26. Bisson JI, Berliner L, Cloitre M, et al. The International Society for Traumatic Stress Studies New Guidelines for the Prevention and Treatment of Posttraumatic Stress Disorder: Methodology and Development Process. J Trauma Stress. 2019 Aug;32(4):475-483. doi: 10.1002/jts.22421.
27. Ostacher MJ, Cifu AS. Management of Posttraumatic Stress Disorder. JAMA. 2019 Jan 15;321(2):200-201. doi: 10.1001/jama.2018.19290.
28. Frueh BC, Grubaugh AL, Yeager DE, et al. Delayed-onset post-traumatic stress disorder among war veterans in primary care clinics. Br J Psychiatry. 2009 Jun;194(6):515-20. doi: 10.1192/bjp.bp.108.054700.
29. Vabalas A, Gowen E, Poliakoff E, et al. Machine learning algorithm validation with a limited sample size. PLoS One. 2019 Nov 7;14(11):e0224365. doi: 10.1371/journal.pone.0224365.
30. Greenhalgh T, Howick J, Maskrey N. Evidence Based Medicine Renaissance Group. Evidence based medicine: a movement in crisis? BMJ. 2014 Jun 13;348:g3725. doi: 10.1136/bmj.g3725.

31. Passos IC, Mwangi B, Kapczinski F. Personalized psychiatry: Big data analytics in mental Health (1st ed.)). New York, United States, 2009: Springer International Publishing Company.
32. McGorry PD, Hickie IB, Yung AR, et al. Clinical staging of psychiatric disorders: a heuristic framework for choosing earlier, safer and more effective interventions. Aust N Z J Psychiatry. 2006 Aug;40(8):616-22. doi: 10.1080/j.1440-1614.2006.01860.x.
33. Dwyer DB, Falkai P, Koutsouleris N. Machine Learning Approaches for Clinical Psychology and Psychiatry. Annu Rev Clin Psychol. 2018 May 7;14:91-118. doi: 10.1146/annurev-clinpsy-

Supplementary Material

Supplementary Item 1. Model configuration

The inner loop CV (inner-CV) is responsible for hyperparameter tuning and model selection, whereas the outer-CV is responsible for error estimation based on the model tuned in the inner-CV. We chose a leave-one-out cross-validation method (LOOCV) to the outer-CV, while the inner loop is configured as a two-fold CV. The best model configuration found during the model development process based on the analysis of Sample 1 (train set) is then applied to Sample 2 (test set) to evaluate performance.

Since we have four stages with different sizes as outcome, we must deal with a multiclass classification with imbalanced classes. In addition to using the original data (where smaller classes are underrepresented), we tested two other strategies, namely (1) down-sampling, which consists in reducing the bigger classes to the size of the smallest class, and (2) up-sampling, which consists in inflating the smaller classes by adding up randomly repeated cases to the size of the biggest class. These strategies were applied in all algorithms except C5.0, which controls the negative effect of class imbalance by using a cost-sensitive learning strategy.

Supplementary Item 2. Macro and Micro-averaging

The AUC score normally applies to binary problems. For multiclass problems, there are two possible ways to obtain 'average' scores: (1) using macro-averaging, which reduces the multiclass predictions down to multiple sets of binary predictions. It calculates the corresponding metric for each of the binary cases and then average the results together (macro-averaging reduces the problem to multiple one-vs-all comparisons); and (2) using micro-averaging, which treats the entire set of data as an aggregate result and calculates a single metric rather than *k* metrics that get averaged together (by calculating all of the true positive results for each class and using that as the numerator. and then calculating all of the true positive and false positive results for each class. and using that as the denominator). In this case, rather than each *class* having equal weight, each *observation* gets equal weight. This gives the classes with the most observations more power. We compared both methods and decided to use the macro-averaging method to compare these performance metrics. as results seemed more reliable for our imbalanced multiclass problem.

Supplementary Table 1. Staging model of PTSD.

Clinical aspects	Possible neurobiological changes
Stage 0: Trauma exposed asymptomatic but at risk	Down regulation of glucocorticoid receptor sensitivity. Increased amygdala reactivity. 5FKH genotype.
Stage 1a: Undifferentiated symptoms of mild anxiety and distress	Inflammatory cytokine activation. Decreasing response inhibition in the frontal cognitive systems.
Stage 1b: Subsyndromal distress with some behavioural and functional decline	Increased physiological reactivity to trauma-related stimuli and startle response. Prolonged autonomic arousal on provocation.
Stage 2: First episode of full-threshold symptoms that has different trajectories	Early and potentially reversible neurobiological disinhibition of frontolimbic circuitry.
Stage 3: Persistent symptoms which may fluctuate with ongoing impairment: ● 3a Incomplete remission of first episode ● 3b Recurrence or relapse of PTSD and persistent impairments ● 3c Multiple relapses or worsening following incomplete treatment response	Decreased anterior cingulate and hippocampal volume. Hypertension and metabolic syndrome.
Stage 4: Severe unremitting illness of increasing chronicity with substantial disability	High allostatic load. High levels of inflammation. Medical comorbidities. Entrenched sensitization of a range of neurobiological systems.

Adapted from McFarlane et al.[10]

Supplementary Table 2. Brief descriptions of the ML techniques.

Algorithm	Description
C5.0	C5.0 is a decision-tree, cost-sensitive learning algorithm developed for imbalanced classes. Cost-sensitive learning algorithms take the misclassification cost into account to minimize classification errors. C5.0 allows the use of a cost matrix to emphasize certain classes over others, penalizing misclassification.
Gradient Boosting Machines (GBM)	GBM is derived from AdaBoost, a Decision Tree ML technique. They develop two trees. one with equal weight to each observation and one with increased weights to those observations that were difficult to classify. With classification errors. a third tree is developed to predict the revised residuals. This process is repeated for a specific number of interactions. Final results are the weighted sum of previous predictions.
Elastic Net	Elastic Net regression allows various penalties where coefficients for collinear independent variables are shrunk toward zero (but not eliminating contributions to the predicted probability) and/or to zero (eliminating their contributions to the predicted probability).
Random Forests (RF)	RF build numerous decision trees in bootstrapped samples and generate an aggregate tree by averaging across trees (reducing overfitting). RF also performs internal feature selection and error estimation.
Support Vector Machines (SVM)	SVM treats each independent variable as a dimension in high dimensional space and attempts to identify the best hyperplane to separate the sample into classes (e.g.. cases and non-cases). SVM captures linear associations (linear kernel), but alternate algorithms can be used to capture nonlinearities (e.g. polynomial and radial basis kernels).

Supplementary Table 3. Configurations used in the parameters optimization for each ML algorithm.

Algorithm	*caret* Method	Parameters and tested values
C5.0	*c5.0*	Parameters were varied within the following pre-define sets: *trials* = {1.2.3.4.5.6.7.8.9.10.20.30.40.50}; *model* = {"tree"}; *winnow* = {FALSE};
GBM	*gbm*	Parameters were varied within the following pre-define sets: *n.minobsinnode* = {1}; *n.trees* = {100. 250. 500. 1000}; *interaction.depth* = {1. 3. 5}; *shrinkage* = {0.1}
Elastic Net	*glmnet*	*tuneLength* parameter in the caret *train* function was set to 5.
RF	*rf*	*mtry* values were manually provided to test five different values centered on the square root of the number of attributes in the training data. which is the standard *mtry* value recommended in the literature for classification tasks
RBF-SVM	*svmRadial*	*tuneLength* parameter in the caret *train* function was set to 5.

Supplementary Table 4. Description of methodological quality features.

Feature	Considerations
1. Representativeness of the sample	Was the study truly representative of the target population heterogeneity? If not. was this related to the selected sampling method. insufficient sample size or inclusion/exclusion criteria?
2. Confounding variables	Did the study control for the most relevant confounding variables (age. gender. trauma type)? If so. were covariates assessed using subjective or objective measures?
3. Outcome Assessment	How were outcome measures assessed: A. Independent blind assessment (✔) B. Secure record (e.g. surgical records) (✔) C. Interview not blinded. self-report or medical record D. No description
4. Machine Learning Approach	Was the machine learning algorithm used to analyze data clearly described and appropriate?
5. Performance/Accuracy	Were the following performance metrics included: A. Accuracy B. Sensitivity

	C. Specificity D. AUC E. PPV/NPV
6. Missing Data	Did the study describe how the authors handled missing data. including if they were inputted or removed?
7. Testing/Validation	Was the test dataset "unseen" in regard to model training? Was the model tested on a hold-out or an external dataset?
8. Class Imbalance	Did the authors address the class imbalance problem? Which method was utilized?
9. Feature Selection and hyperparameter tuning	Did the study describe both feature selection and hyperparameter tuning? Which metrics were used?

Adapted from Passos et al.[20]

Supplementary Table 5. Frequencies of CAPS-5 and demographic features for the RFE execution

CAPS-5 20-item		CAPS-5 20-item + demographics	
Attribute	**%**	**Attribute**	**%**
Caps-D5 (Diminished interest)	98.4%	Current psychiatric comorbidity	96.7%
Caps-D6 (Feelings of detachment)	90.2%	Global severity score	96.7%
Caps-E1 (Irritability)	86.9%	Amount of symptoms	90.1%
Caps-B4 (Psychological distress)	85.3%	Caps score	80.3%
Caps-E4 (Startle response)	82.0%	Caps-D5 (Diminished interest)	72.1%
Caps-C1 (Memory avoidance)	73.8%	Caps-E1 (Irritability)	57.4%
Caps-D2 (Negative beliefs)	70.5%	Caps-B4 (Psychological distress)	50.9%
Caps-B2 (Distressing dreams)	67.2%	Caps-D6 (Feelings of detachment)	45.9%
Caps-D4 (Negative states)	67.2%	Caps-E4 (Startle response)	45.9%
Caps-D3 (Distorted cognitions)	62.3%	Caps-D2 (Negative beliefs)	41.0%
Caps-B3 (Dissociative reactions)	60.7%	Lec5: life-threatening illness	41.0%
Caps-B5 (Physiological reactions)	59.0%	Caps-C1 (Memory avoidance)	39.3%
Caps-C2 (External avoidance)	59.0%	Caps-B2 (Distressing dreams)	34.5%
Caps-D7 (Reduction of positive emotions)	57.4%	Caps-D4 (Negative states)	34.4%
Caps-B1 (Recurrent memories)	50.9%	Caps-B1 (Recurrent memories)	32.8%
Caps-E3 (Hypervigilance)	49.2%	Caps-B5 (Physiological reactions)	32.8%
Caps-E2 (Recklessness)	42.6%	Caps-C2 (External avoidance)	32.8%
Caps-E5 (Concentration problems)	42.6%	Lec5: sexual assault	32.8%
Caps-E6 (Sleep disturbance)	42.6%	Caps-D7 (Reduction of positive emotions	31.1%
Caps-D1 (Dissociative amnesia)	27.9%	Age	31.1%
		Employment status: employed	31.1%
		Lec5: assault with a weapon	31.1%
		Caps-B3 (Dissociative reactions)	29.5%
		Marital status: divorced/widower	29.5%
		Religiousness	29.5%
		Presence of dissociative symptoms	29.5%
		Caps-D3 (Distorted cognitions)	27.9%
		Marital status: single	27.9%
		Previous trauma event	27.9%
		Previous psychiatric disorders	26.2%
		Drugs use	26.2%
		Marital status: married/engaged	26.2%
		Alcohol use	26.2%
		Caps-E6 (Sleep disturbance)	25.0%

Living alone	25.0%
Known aggressor	25.0%
Caps-E5 (Concentration problems)	23.0%
Educational level: up to 4y	23.0%
Ethnicity: caucasian	23.0%
Ethnicity: asiatic-american	23.0%
Lec5: sudden acidental death, serious injury, fire/explosion, transportation accident, physical assault	23.0%
Caps-E2 (Recklessness)	21.3%
Employment status: unemployed/retired	21.3%
Tobacco use	21.3%
Lec5: severe human suffering, sudden violent death	21.3%
Educational level: above 12y	19.7%
Ethnicity: african-american	19.7%
Caps-E3 (Hypervigilance)	18.0%
Educational level: 4-12y	18.0%
Any substance abuse	16.4%
Caps-D1 (Dissociative amnesia)	14.7%
Employment status: health licensed	14.7%

Supplementary Table 6. Classification performances for each stage (Recall).

Model/algorithm	Train (Sample 1)				Test (Sample 2)			
	S1	S2	S3	S4	S1	S2	S3	S4
CAPS-5 20-item								
C5.0	54.5%	21.7%	23.8%	66.7%	12.5%	50.0%	15.0%	20.0%
GBM	81.8%	17.4%	42.9%	83.3%	25.0%	50.0%	20.0%	40.0%
Elastic Net	81.8%	34.8%	23.8%	0.0%	25.0%	5.6%	15.0%	80.0%
Random Forest	81.8%	21.7%	33.3%	66.7%	37.5%	55.6%	10.0%	20.0%
RBF-SVM	90.9%	13.0%	33.3%	0.0%	25.0%	33.3%	55.0%	20.0%
CAPS-5 15-item								
C5.0	54.5%	21.7%	52.4%	66.7%	12.5%	22.2%	25.0%	40.0%
GBM	81.8%	13.0%	9.5%	66.7%	25.0%	66.7%	20.0%	0.0%
Elastic Net	81.8%	39.1%	23.8%	16.7%	25.0%	27.8%	10.0%	60.0%
Random Forest	81.8%	17.4%	19.0%	66.7%	50.0%	72.2%	5.0%	20.0%
RBF-SVM	90.9%	21.7%	38.1%	0.0%	25.0%	55.6%	30.0%	20.0%
CAPS-5 20-item+demographics								
C5.0	81.8%	69.6%	42.9%	50.0%	37.5%	61.1%	55.0%	20.0%
GBM	72.7%	26.1%	14.3%	83.3%	50.0%	38.9%	25.0%	80.0%
Elastic Net	72.7%	43.5%	28.6%	33.3%	37.5%	55.6%	15.0%	60.0%
Random Forest	81.8%	30.4%	14.3%	50.0%	50.0%	66.7%	20.0%	80.0%
RBF-SVM	81.8%	17.4%	38.1%	0.0%	50.0%	38.9%	30.0%	40.0%
CAPS-5 15-item+demographics								
C5.0	81.8%	73.9%	61.9%	16.7%	37.5%	66.7%	55.0%	20.0%
GBM	72.7%	30.4%	19.0%	83.3%	50.0%	72.2%	40.0%	20.0%
Elastic Net	72.7%	39.1%	38.1%	50.0%	50.0%	72.2%	10.0%	100.0%
Random Forest	81.8%	65.2%	38.1%	0.0%	50.0%	72.2%	15.0%	80.0%
RBF-SVM	81.8%	17.4%	47.6%	16.7%	50.0%	61.1%	20.0%	20.0%

GBM: Gradient Boosting Machine; RBF-SVM: Radial-basis, Support Vector Machine

Supplementary Table 7. Top relevant features (CAPS 15-item+demographics).

Algorithm	Attribute
C5.0	Number of symptoms
	Current psychiatric comorbodity
	Global severity score
	Caps-E5 (Concentration problems)
	Lec5-life threatening illness
Gradient Boosting Machine	CAPS-5 score
	Age
	Caps-D2 (Negative beliefs)
	Current psychiatric comorbodity
	Caps-C2 (External avoidance)
Elastic Net	Lec5-life threatening illness
	Current psychiatric comorbidity
	Educational level: 4-12y
	Lec5-assault with a weapon
	Alcohol use
Support Vector Machine	Number of symptoms
	CAPS-5 score
	Global severity score
	Caps-B4
	Caps-B5 (Physiological reactions)

Supplementary Table 8: Clinical and sociodemographic information of the four predicted stages, test set (PROVE, n=51)

Variable	Stage 1	Stage 2	Stage 3	Stage 4	Total	*P* value
Female gender, No (%)	3 (100.0%)	20 (87.0%)	17 (94.4%)	6 (85.7%)	46 (90.2%)	.771
Religiousness, No (%)	1 (33.3%)	16 (69.6%)	10 (55.6%)	4 (57.1%)	31 (60.8%)	.585
Living alone, No (%)	0 (0.0%)	2 (8.7%)	1 (5.6%)	3 (42.9%)	6 (11.8%)	[a].050
Current psychiatric comorbidity, No (%)	1 (33.3%)	3 (13.0%)	17 (94.4%)	7 (100.0%)	28 (54.9%)	[a] <.000
Previous psychiatric disorders, No (%)	1 (33.3%)	5 (21.7%)	16 (88.9%)	3 (42.9%)	25 (49.9%)	[a] <.000
Any drug use, No (%)	0 (0.0%)	1 (4.3%)	10 (55.6%)	1 (14.3%)	12 (23.5%)	[a].001
[b]Previous trauma event, No (%)	2 (66.7%)	8 (34.8%)	16 (88.9%)	3 (42.9%)	29 (56.9%)	[a].005
Global severity score, No (%)						
0-1	2 (66.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	2 (4.0%)	[a].003
2	1 (33.3%)	6 (26.1%)	3 (16.7%)	0 (0.0%)	10 (19.6%)	
3	0 (0.0%)	12 (52.2%)	15 (83.3%)	4 (57.1%)	31 (60.8%)	
4	0 (0.0%)	5 (21.7%)	0 (0.0%)	3 (42.9%)	8 (15.7%)	
Number of symptoms, median [IQR]	6.0 [6.0-6.5]	14.0 [13.0-16.5]	15.0 [13.0-16.0]	15.0 [14.5-17.0]	15.0 [13.0-16.0]	.498
CAPS score, median [IQR]	20.0 [17.5-21.0]	42.0 [37.0-46.0]	40.5 [37.0-46.0]	50.0 [43.0-51.5]	41.0 [35.0-47.0]	.209

[a]$p<0.05$ [b]Excluding the current event

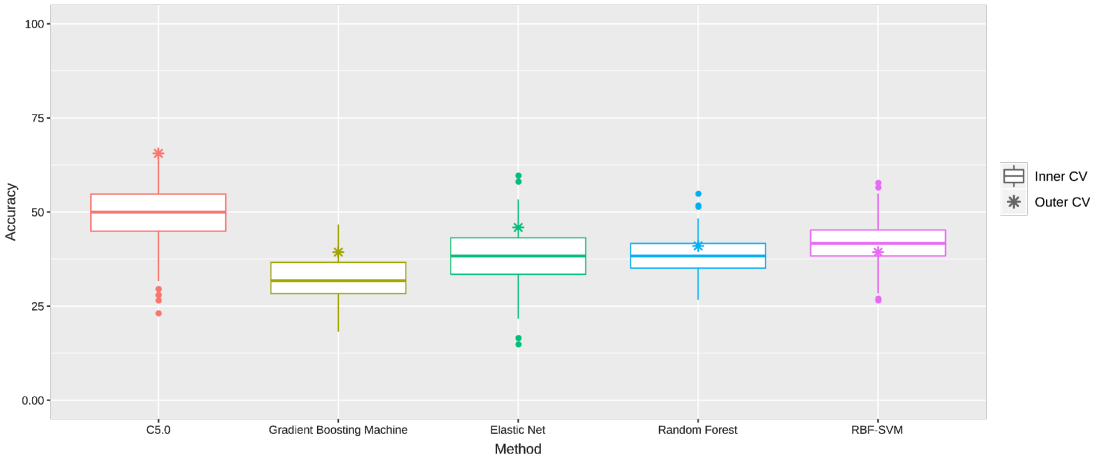
Supplementary Figure 1. Dispersion graphics.

Supplementary Figure 2. Cost matrices for C5.0 executions.

Execution	Matrix cost (predicted X real classes)				
4 classes CAPS-5 15-item+demographics CAPS-5 20-item+demographics		S1	S2	S3	S4
	S1	0	10	10	10
	S2	0	0	20	35
	S3	10	20	0	35
	S4	10	20	20	0
4 classes CAPS-5 15-item CAPS-5 20-item		S1	S2	S3	S4
	S1	0	10	10	10
	S2	0	0	20	35
	S3	0	0	0	35
	S4	0	0	0	0
3 classes CAPS-5 15-item+demographics		S1	S2-S3	S4	
	S1	0	10	10	
	S2-S3	10	0	35	
	S4	10	20	0	
2 classes CAPS-5 15-item+demographics		S1	S2-3-4		
	S1	0	10		
	S2-3-4	0	0		

Columns represented reference (real) classes and lines denote predicted classes.

Supplementary Figure 3. Performance comparison among inner and outer-CV executions.



All algorithms except C5.0 considered down-sampling method.

6 FINAL CONSIDERATIONS AND CONCLUSION

This research aims to reinforce the role of ML and the technologies involved in the development of knowledge in the area of health sciences, especially its application in psychiatry.

As pointed out in the final article of this thesis (see **article 3**), one limitation of our study was the cross-sectional design. Longitudinal studies are lacking for correlating neurobiological changes and indicating specific treatments for each clinical stage. Longitudinal studies can also provide an observation of the staging evaluation throughout time, allowing an improvement of our understanding of PTSD evolution and prognosis. Another presentation of PTSD is Complex PTSD (C-PTSD) [84]. C-PTSD was not included in this research because of the lack of instruments up to the beginning of this study (CAPS-5 was not designed to diagnose C-PTSD; recently, the ICD-11 Trauma Questionnaire was released in Brazilian Portuguese [85]). It is hypothesized that "Stage 4" may also include C-PTSD; further studies are necessary to confirm this.

The use of clinical (CAPS-5) and sociodemographic information to predict PTSD staging was considered a strength of this study, besides using the advantages of ML techniques. Those kinds of data are easy to obtain in an outpatient practice but difficult to combine via traditional statistical approaches. In the field of PTSD, clinical heterogeneity can be a significant factor, not always taken into account in research. Staging links the clinical aspects to treatment selection and prediction: early successful treatments may change the prognosis and thus prevent progression to subsequent stages [27].

Considering that a doctoral thesis involves synthesis of the available information in order to propose new knowledge, an understanding of the targeted population was initially needed. An article analyzing a 4-year follow-up of victims of a traumatic event was developed (see **article 1**). Also, it was necessary to adapt and to validate the instrument to diagnose and measure the severity of symptoms regarding PTSD, CAPS-5 (see **additional articles 1 and 2** in annex). In order to propose the most suitable ML techniques to develop a staging model, a systematic review of these techniques applied to the field of PTSD was performed (see **article 2**). As the main result of these efforts, a staging model for PTSD was developed (see **article 3**).

PTSD staging definitions were proposed by McFarlane et al. [29]; no previous studies using PTSD staging were found. Reflecting possible outpatient environments, stages 2 and 3 were more representative in the samples; stage 1 (sub-syndromic PTSD) and stage 4 (severe

PTSD, C-PTSD?) appeared in fewer occurrences. Staging allows correlation with specific treatment schemes: one of the most relevant results of the final article is a treatment proposition based on staging (see **article 3**). Further investigation is needed to confirm and improve this proposition.

Clinical observation and proper knowledge of diagnostic criteria are of major importance to determine PTSD. Technology becomes an ally to health professionals in their research and practice throughout many fields; in psychiatry, its use is still incipient. Considering psychiatric disorders with genetic and environmental influences, a heterogeneous population, with different treatments and prognoses, PTSD emerges as an essential example of where staging could be of benefit. Its implementation, though, could be a tough job. ML techniques can be advantageous, as they can operate complex relationships among variables including non-linear patterns, and can be focused at an individual patient level [86].

Anyone can experience situations of stress throughout life. Most individuals are able to overcome these situations in a given period of time; often, the removal of the stressor stimulus already causes stress to disappear. Some stressful situations are essentially traumatic, and tend to last throughout life. They bring scars that are difficult to heal, with much suffering for these victims. PTSD is a disorder that can impact the individual's life in several ways, affecting personal, professional and health aspects [19]. This research is dedicated to those who have suffered these scars.

Future research may include longitudinal studies and the use of other sorts of information, such as biomarkers and neuroimaging studies, to adjust and detail the phenotypic presentations of PTSD in order to provide more targeted treatments.

REFERENCES

[1] Schestatsky S, Shansis F et al. A evolução histórica do conceito de A evolução histórica do conceito de estresse pós-traumático estresse pós-traumático. Rev Bras Psiquiatr 25(supl 1): 8-11. 2003. Rev. Bras. Psiquiatr., São Paulo , v. 25, supl. 1, p. 8-11, June 2003. doi:10.1590/S1516-44462003000500003.

[2] Lifton RJ, Olson E. The human meaning of total disaster. The Buffalo Creek experience. Psychiatry. 1976;39(1):1-18. doi:10.1080/00332747.1976.11023872

[3] American Psychiatric Association (APA). Diagnostic and statistical manual of mental disorders. 3rd ed. Washington (DC), 1980.

[4] American Psychiatric Association (APA). Diagnostic and statistical manual of mental disorders. 5th ed. Washington (DC), 2013.

[5] Yehuda R. Risk and resilience in posttraumatic stress disorder. J Clin Psychiatry. 2004;65 Suppl 1:29-36.

[6] Ribeiro WS, Mari Jde J, Quintana MI, et al. The impact of epidemic violence on the prevalence of psychiatric disorders in Sao Paulo and Rio de Janeiro, Brazil. PLoS One. 2013;8(5):e63545. Published 2013 May 8. doi:10.1371/journal.pone.0063545

[7] Kessler RC, Sonnega A, Bromet E, Hughes M, Nelson CB. Posttraumatic stress disorder in the National Comorbidity Survey. Arch Gen Psychiatry. 1995;52(12):1048-1060. doi:10.1001/archpsyc.1995.03950240066012

[8] Friedman MJ, Resick PA, Keane TM Handbook of PTSD: Science and Practice. Guilford Press; New York, NY, USA: 2007. PTSD: Twenty-five years of progress and challenges; pp. 376–405.

[9] Ozer EJ, Best SR, Lipsey TL, Weiss DS. Predictors of posttraumatic stress disorder and symptoms in adults: a meta-analysis. Psychol Bull. 2003;129(1):52-73. doi:10.1037/0033-2909.129.1.52

[10] Kilpatrick DG, Resnick HS, Milanak ME, Miller MW, Keyes KM, Friedman MJ. National estimates of exposure to traumatic events and PTSD prevalence using DSM-IV and DSM-5 criteria. J Trauma Stress. 2013;26(5):537-547. doi:10.1002/jts.21848

[11] Shalev AY, Gevonden M, Ratanatharathorn A, et al. Estimating the risk of PTSD in recent trauma survivors: results of the International Consortium to Predict PTSD (ICPP). World Psychiatry. 2019;18(1):77-87. doi:10.1002/wps.20608

[12] Pupo MC, Serafim PM, de Mello MF. Health-related quality of life in posttraumatic stress disorder: 4 years follow-up study of individuals exposed to urban violence. Psychiatry Res. 2015;228(3):741-745. doi:10.1016/j.psychres.2015.05.030

[13] Cordero MI, Moser DA, Manini A, et al. Effects of interpersonal violence-related post-traumatic stress disorder (PTSD) on mother and child diurnal cortisol rhythm and cortisol reactivity to a laboratory stressor involving separation. Horm Behav. 2017;90:15-24. doi:10.1016/j.yhbeh.2017.02.007

[14] Kostaras P, Bergiannaki JD, Psarros C, Ploumbidis D, Papageorgiou C. Posttraumatic stress disorder in outpatients with depression: Still a missed diagnosis. J Trauma Dissociation. 2017;18(2):233-247. doi:10.1080/15299732.2016.1237402

[15] Post RM, Weiss SR. Sensitization and kindling phenomena in mood, anxiety, and obsessive-compulsive disorders: the role of serotonergic mechanisms in illness progression. Biol Psychiatry. 1998;44(3):193-206. doi:10.1016/s0006-3223(98)00144-9

[16] Galatzer-Levy IR, Karstoft KI, Statnikov A, Shalev AY. Quantitative forecasting of PTSD from early trauma responses: a Machine Learning application. J Psychiatr Res. 2014;59:68-76. doi:10.1016/j.jpsychires.2014.08.017

[17] Ressler KJ, Mercer KB, Bradley B, et al. Post-traumatic stress disorder is associated with PACAP and the PAC1 receptor [published correction appears in Nature. 2011 Sep 1;477(7362):120]. Nature. 2011;470(7335):492-497. doi:10.1038/nature09856

[18] Michopoulos V, Norrholm SD, Jovanovic T. Diagnostic Biomarkers for Posttraumatic Stress Disorder: Promising Horizons from Translational Neuroscience Research. Biol Psychiatry. 2015;78(5):344-353. doi:10.1016/j.biopsych.2015.01.005

[19] Yehuda R, Hoge CW, McFarlane AC, et al. Post-traumatic stress disorder. Nat Rev Dis Primers. 2015;1:15057. Published 2015 Oct 8. doi:10.1038/nrdp.2015.57

[20] Sartory G, Cwik J, Knuppertz H, et al. In search of the trauma memory: a meta-analysis of functional neuroimaging studies of symptom provocation in posttraumatic stress disorder (PTSD). PLoS One. 2013;8(3):e58150. doi:10.1371/journal.pone.0058150

[21] Lanius RA, Vermetten E, Loewenstein RJ, et al. Emotion modulation in PTSD: Clinical and neurobiological evidence for a dissociative subtype. Am J Psychiatry. 2010;167(6):640-647. doi:10.1176/appi.ajp.2009.09081168

[22] Koch SB, van Zuiden M, Nawijn L, Frijling JL, Veltman DJ, Olff M. Aberrant resting-state brain activity in posttraumatic stress disorder: a meta-analysis and systematic review. Depress Anxiety. 2016;33(7):592-605. doi:10.1002/da.22478

[23] Yehuda R, Lehrner A. Intergenerational transmission of trauma effects: putative role of epigenetic mechanisms. World Psychiatry. 2018;17(3):243-257. doi:10.1002/wps.20568

[24] VA/DOD Clinical Practice Guideline for the Management of Posttraumatic Stress Disorder and Acute Stress Disorder: Clinician Summary. Focus (Am Psychiatr Publ). 2018;16(4):430-448. doi:10.1176/appi.focus.16408

[25] Charney ME, Hellberg SN, Bui E, Simon NM. Evidenced-Based Treatment of Posttraumatic Stress Disorder: An Updated Review of Validated Psychotherapeutic and Pharmacological Approaches. Harv Rev Psychiatry. 2018;26(3):99-115. doi:10.1097/HRP.0000000000000186

[26] Nash WP, Watson PJ. Review of VA/DOD Clinical Practice Guideline on management of acute stress and interventions to prevent posttraumatic stress disorder. J Rehabil Res Dev. 2012;49(5):637-648. doi:10.1682/jrrd.2011.10.0194

[27] McGorry PD, Hickie IB, Yung AR, Pantelis C, Jackson HJ. Clinical staging of psychiatric disorders: a heuristic framework for choosing earlier, safer and more effective interventions. Aust N Z J Psychiatry. 2006;40(8):616-622. doi:10.1080/j.1440-1614.2006.01860.x

[28] Galatzer-Levy IR, Ma S, Statnikov A, Yehuda R, Shalev AY. Utilization of machine learning for prediction of post-traumatic stress: a re-examination of cortisol in the prediction and pathways to non-remitting PTSD. Transl Psychiatry. 2017;7(3):e0. Published 2017 Mar 21. doi:10.1038/tp.2017.38

[29] McFarlane AC, Lawrence-Wood E, Van Hooff M, Malhi GS, Yehuda R. The Need to Take a Staging Approach to the Biological Mechanisms of PTSD and its Treatment. Curr Psychiatry Rep. 2017;19(2):10. doi:10.1007/s11920-017-0761-2

[30] Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ. 1996;312(7023):71-72. doi:10.1136/bmj.312.7023.71

[31] Greenhalgh T, Howick J, Maskrey N; Evidence Based Medicine Renaissance Group. Evidence based medicine: a movement in crisis?. BMJ. 2014;348:g3725. Published 2014 Jun 13. doi:10.1136/bmj.g3725

[32] Vasterling JJ, Aslan M, Proctor SP, et al. Longitudinal Examination of Posttraumatic Stress Disorder as a Long-Term Outcome of Iraq War Deployment. Am J Epidemiol. 2016;184(11):796-805. doi:10.1093/aje/kww151

[33] Hines LA, Sundin J, Rona RJ, Wessely S, Fear NT. Posttraumatic stress disorder post Iraq and Afghanistan: prevalence among military subgroups. Can J Psychiatry. 2014;59(9):468-479. doi:10.1177/070674371405900903

[34] Silva TLG, Mello PG, Silveira KAL, Wolffenbüttel L, Lobo BOM, Bicca CHM, et al. Primeiros Socorros Psicológicos: relato de intervenção em crise em Santa Maria. Rev. bras. psicoter. 2013;15(1):93-104 (Portuguese)

[35] Tay AK, Rees S, Steel Z, et al. Six-year trajectories of post-traumatic stress and severe psychological distress symptoms and associations with timing of trauma exposure, ongoing adversity and sense of injustice: a latent transition analysis of a community cohort in conflict-affected Timor-Leste. BMJ Open. 2016;6(2):e010205. Published 2016 Feb 23. doi:10.1136/bmjopen-2015-010205

[36] Huys QJ, Maia TV, Frank MJ. Computational psychiatry as a bridge from neuroscience to clinical applications. Nat Neurosci. 2016;19(3):404-413. doi:10.1038/nn.4238

[37] Liu NT, Salinas J. Machine Learning for Predicting Outcomes in Trauma. Shock. 2017;48(5):504-510. doi:10.1097/SHK.0000000000000898

[38] Passos IC, Mwangi B, Cao B, et al. Identifying a clinical signature of suicidality among patients with mood disorders: A pilot study using a machine learning approach. J Affect Disord. 2016;193:109-116. doi:10.1016/j.jad.2015.12.066

[39] Schultebraucks K, Galatzer-Levy IR. Machine Learning for Prediction of Posttraumatic Stress and Resilience Following Trauma: An Overview of Basic Concepts and Recent Advances. J Trauma Stress. 2019;32(2):215-225. doi:10.1002/jts.22384

[40] Bishop CM. Pattern Recognition and Machine Learning. Springer; 2006.

[41] Deo RC. Machine Learning in Medicine. Circulation. 2015;132(20):1920-1930. doi:10.1161/CIRCULATIONAHA.115.001593

[42] Rosellini AJ, Dussaillant F, Zubizarreta JR, Kessler RC, Rose S. Predicting posttraumatic stress disorder following a natural disaster. J Psychiatr Res. 2018;96:15-22. doi:10.1016/j.jpsychires.2017.09.010

[43] Karstoft KI, Galatzer-Levy IR, Statnikov A, Li Z, Shalev AY; members of Jerusalem Trauma Outreach and Prevention Study (J-TOPS) group. Bridging a translational gap: using machine learning to improve the prediction of PTSD. BMC Psychiatry. 2015;15:30. Published 2015 Mar 16. doi:10.1186/s12888-015-0399-8

[44] Karstoft KI, Statnikov A, Andersen SB, Madsen T, Galatzer-Levy IR. Early identification of posttraumatic stress following military deployment: Application of machine

learning methods to a prospective study of Danish soldiers. J Affect Disord. 2015;184:170-175. doi:10.1016/j.jad.2015.05.057

[45] Conrad D, Wilker S, Pfeiffer A, et al. Does trauma event type matter in the assessment of traumatic load?. Eur J Psychotraumatol. 2017;8(1):1344079. Published 2017 Jul 6. doi:10.1080/20008198.2017.1344079

[46] Kessler RC, Rose S, Koenen KC, et al. How well can post-traumatic stress disorder be predicted from pre-trauma risk factors? An exploratory study in the WHO World Mental Health Surveys. World Psychiatry. 2014;13(3):265-274. doi:10.1002/wps.20150

[47] Köbach A, Nandi C, Crombach A, Bambonyé M, Westner B, Elbert T. Violent Offending Promotes Appetitive Aggression Rather than Posttraumatic Stress-A Replication Study with Burundian Ex-Combatants. Front Psychol. 2015;6:1755. Published 2015 Dec 8. doi:10.3389/fpsyg.2015.01755

[48] Tylee DS, Chandler SD, Nievergelt CM, et al. Blood-based gene-expression biomarkers of post-traumatic stress disorder among deployed marines: A pilot study. Psychoneuroendocrinology. 2015;51:472-494. doi:10.1016/j.psyneuen.2014.09.024

[49] Nicholson AA, Densmore M, McKinnon MC, et al. Machine learning multivariate pattern analysis predicts classification of posttraumatic stress disorder and its dissociative subtype: a multimodal neuroimaging approach. Psychol Med. 2019;49(12):2049-2059. doi:10.1017/S0033291718002866

[50] Badura-Brack AS, Heinrichs-Graham E, McDermott TJ, et al. Resting-State Neurophysiological Abnormalities in Posttraumatic Stress Disorder: A Magnetoencephalography Study. Front Hum Neurosci. 2017;11:205. Published 2017 Apr 25. doi:10.3389/fnhum.2017.00205

[51] Li Y, Hou X, Wei D, et al. Long-Term Effects of Acute Stress on the Prefrontal-Limbic System in the Healthy Adult. PLoS One. 2017;12(1):e0168315. Published 2017 Jan 3. doi:10.1371/journal.pone.0168315

[52] Zhang Q, Wu Q, Zhu H, et al. Multimodal MRI-Based Classification of Trauma Survivors with and without Post-Traumatic Stress Disorder. Front Neurosci. 2016;10:292. Published 2016 Jun 24. doi:10.3389/fnins.2016.00292

[53] Liu F, Xie B, Wang Y, et al. Characterization of post-traumatic stress disorder using resting-state fMRI with a multi-level parametric classification approach. Brain Topogr. 2015;28(2):221-237. doi:10.1007/s10548-014-0386-2

[54] Chekroud AM, Koutsouleris N. The perilous path from publication to practice. Mol Psychiatry. 2018;23(1):24-25. doi:10.1038/mp.2017.227

[55] Salagre E, Dodd S, Aedo A, et al. Toward Precision Psychiatry in Bipolar Disorder: Staging 2.0. Front Psychiatry. 2018;9:641. Published 2018 Nov 29. doi:10.3389/fpsyt.2018.00641

[56] Kapczinski F, Magalhães PV, Balanzá-Martinez V, et al. Staging systems in bipolar disorder: an International Society for Bipolar Disorders Task Force Report. Acta Psychiatr Scand. 2014;130(5):354-363. doi:10.1111/acps.12305

[57] McGorry PD, Hickie IB, Yung AR, Pantelis C, Jackson HJ. Clinical staging of psychiatric disorders: a heuristic framework for choosing earlier, safer and more effective interventions. Aust N Z J Psychiatry. 2006;40(8):616-622. doi:10.1080/j.1440-1614.2006.01860.x

[58] Fava GA, Kellner R. Staging: a neglected dimension in psychiatric classification. Acta Psychiatr Scand. 1993;87(4):225-230. doi:10.1111/j.1600-0447.1993.tb03362.x

[59] de la Fuente-Tomas L, Sánchez-Autet M, García-Álvarez L, et al. Clinical staging in severe mental disorders; bipolar disorder, depression and schizophrenia. Rev Psiquiatr Salud Ment. 2019;12(2):106-115. doi:10.1016/j.rpsm.2018.08.002

[60] Berk M, Hallam KT, McGorry PD. The potential utility of a staging model as a course specifier: a bipolar disorder perspective. J Affect Disord. 2007;100(1-3):279-281. doi:10.1016/j.jad.2007.03.007

[61] Kapczinski F, Dias VV, Kauer-Sant'Anna M, et al. Clinical implications of a staging model for bipolar disorders. Expert Rev Neurother. 2009;9(7):957-966. doi:10.1586/ern.09.31

[62] Duffy A. Toward a comprehensive clinical staging model for bipolar disorder: integrating the evidence. Can J Psychiatry. 2014;59(12):659-666. doi:10.1177/070674371405901208

[63] Fava GA, Tossani E. Prodromal stage of major depression. Early Interv Psychiatry. 2007;1(1):9-18. doi:10.1111/j.1751-7893.2007.00005.x

[64] Hetrick SE, Parker AG, Hickie IB, Purcell R, Yung AR, McGorry PD. Early identification and intervention in depressive disorders: towards a clinical staging model. Psychother Psychosom. 2008;77(5):263-270. doi:10.1159/000140085

[65] Lieberman JA, Perkins D, Belger A, et al. The early stages of schizophrenia: speculations on pathogenesis, pathophysiology, and therapeutic approaches [published correction appears in Biol Psychiatry 2002 Feb 15;51(4):346]. Biol Psychiatry. 2001;50(11):884-897. doi:10.1016/s0006-3223(01)01303-8

[66] Singh SP, Cooper JE, Fisher HL, et al. Determining the chronology and components of psychosis onset: The Nottingham Onset Schedule (NOS). Schizophr Res. 2005;80(1):117-130. doi:10.1016/j.schres.2005.04.018

[67] McGorry PD, Nelson B, Goldstone S, Yung AR. Clinical staging: a heuristic and practical strategy for new research and better health and social outcomes for psychotic and related mood disorders. Can J Psychiatry. 2010;55(8):486-497. doi:10.1177/070674371005500803

[68] Agius M, Goh C, Ulhaq S, McGorry P. The staging model in schizophrenia, and its clinical implications. Psychiatr Danub. 2010;22(2):211-220.

[69] Koek RJ, Schwartz HN, Scully S, et al. Treatment-refractory posttraumatic stress disorder (TRPTSD): a review and framework for the future. Prog Neuropsychopharmacol Biol Psychiatry. 2016;70:170-218. doi:10.1016/j.pnpbp.2016.01.015

[70] Calhoun PS, Wiley M, Dennis MF, Means MK, Edinger JD, Beckham JC. Objective evidence of sleep disturbance in women with posttraumatic stress disorder. J Trauma Stress. 2007;20(6):1009-1018. doi:10.1002/jts.20255

[71] Grassi-Oliveira R, Ashy M, Stein LM. Psychobiology of childhood maltreatment: effects of allostatic load?. Braz J Psychiatry. 2008;30(1):60-68. doi:10.1590/s1516-44462008000100012

[72] Stein MB, Walker JR, Hazen AL, Forde DR. Full and partial posttraumatic stress disorder: findings from a community survey. Am J Psychiatry. 1997;154(8):1114-1119. doi:10.1176/ajp.154.8.1114

[73] Fernandes BS, Williams LM, Steiner J, Leboyer M, Carvalho AF, Berk M. The new field of 'precision psychiatry'. BMC Med. 2017;15(1):80. Published 2017 Apr 13. doi:10.1186/s12916-017-0849-x

[74] Gentes EL, Schry AR, Hicks TA, et al. Prevalence and correlates of cannabis use in an outpatient VA posttraumatic stress disorder clinic. Psychol Addict Behav. 2016;30(3):415-421. doi:10.1037/adb0000154

[75] Schnyder U, Schäfer I, Aakvaag HF, et al. The global collaboration on traumatic stress. Eur J Psychotraumatol. 2017;8(sup7):1403257. Published 2017 Nov 30. doi:10.1080/20008198.2017.1403257

[76] Blake DD, Weathers FW, Nagy LM, et al. The development of a Clinician-Administered PTSD Scale. J Trauma Stress. 1995;8(1):75-90. doi:10.1007/BF02105408

[77] Weathers FW, Bovin MJ, Lee DJ, et al. The Clinician-Administered PTSD Scale for DSM-5 (CAPS-5): Development and initial psychometric evaluation in military veterans. Psychol Assess. 2018;30(3):383-395. doi:10.1037/pas0000486

[78] Gray MJ, Litz BT, Hsu JL, Lombardo TW. Psychometric properties of the life events checklist. Assessment. 2004;11(4):330-341. doi:10.1177/1073191104269954

[79] Weathers FW, Litz BT, Keane TM, Palmieri PA, Marx BP, Schnurr PP. The PTSD Checklist for DSM-5 (PCL-5) [Internet]. 2013. Available at https://www.ptsd.va.gov/professional/assessment/te-measures/life_events_checklist.asp. Accessed in 19 Jan 2019.

[80] Lima EP, Vasconcelos AG, Berger W, et al. Cross-cultural adaptation of the Posttraumatic Stress Disorder Checklist 5 (PCL-5) and Life Events Checklist 5 (LEC-5) for the Brazilian context. Trends Psychiatry Psychother. 2016;38(4):207-215. doi:10.1590/2237-6089-2015-0074

[81] Baştanlar Y, Ozuysal M. Introduction to machine learning. Methods Mol Biol. 2014;1107:105-128. doi:10.1007/978-1-62703-748-8_7

[82] Olivera AR. Comparação de algoritmos de aprendizagem de máquina para construção de modelos preditivos de diabetes não diagnosticado [manuscrito].
Dissertação (Mestrado) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação. Porto Alegre/RS, 2014.

[83] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-174.

[84] World Health Organization. (2018). The ICD-11 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines. Geneva: World Health Organization. Available at: http://www.who.int/classifications/icd/en/. Accessed in May 05th, 2020.

[85] Rocha J, Rodrigues V, Santos E, et al. The first instrument for complex PTSD assessment: psychometric properties of the ICD-11 Trauma Questionnaire. Braz J Psychiatry. 2020;42(2):185-189. doi:10.1590/1516-4446-2018-0272

[86] Dwyer DB, Falkai P, Koutsouleris N. Machine Learning Approaches for Clinical Psychology and Psychiatry. Annu Rev Clin Psychol. 2018;14:91-118. doi:10.1146/annurev-clinpsy-032816-045037

ANNEXES

ANNEX 1: CLINICIAN-ADMINISTERED PTSD SCALE (VERSION DSM-5)

ADAPTED TO BRAZILIAN-PORTUGUESE

D. CAPS-5

Record ID _____

CLINICIAN-ADMINISTERED PTSD SCALE 5

Para o restante da entrevista, gostaria que você mantivesse em mente o (EVENTO) enquanto pergunto sobre os diferentes problemas que pode ter causado a você. Você pode ter tido algum desses problemas antes, mas para esta entrevista iremos focar apenas no MÊS PASSADO. Para cada problema irei perguntar se ocorreu durante o mês passado, e caso afirmativo, o quanto e com que frequencia isso incomodou você.

Critério B: Presença de um (ou mais) dos seguintes sintomas intrusivos associados ao(s) evento(s) traumático(s), começando depois de sua ocorrência.

1. (B1) Lembranças intrusivas angustiantes, recorrentes e involuntárias do(s) evento(s) traumático(s). Nota: Em crianças acima de 6 anos de idade, pode ocorrer brincadeira repetitiva na qual temas ou aspectos do evento traumático são expressos.

No mês passado você teve quaisquer memórias indesejadas do (EVENTO), enquanto você estava acordado(a), por isso, sem contar os sonhos?
[Classifique 0=Ausente, se apenas durante os sonhos]

Como acontece de você começar a lembrar do (EVENTO)?

[Se não estiver claro:] (Estas memórias são indesejadas ou você pensa sobre (EVENTO) propositalmente?) [Classifique 0=Ausente, a menos que percebido de forma involuntária e intrusiva]

O quanto essas memórias incomodam você?

Você é capaz de deixá-las de lado e pensar em outra coisa?

B1. Angústia

- ◯ 0 - Mínima
- ◯ 1 - Claramente Presente
- ◯ 2 - Pronunciada
- ◯ 3 - Extrema

B1. Quantas vezes você já teve essas lembranças no mês passado?

_____
(Número de vezes que teve lembranças no mês passado)

B1. Pontuação:

Dimensões chave para classificação = frequência / intensidade da angústia.

Moderado = ao menos 2 X mês / angústia claramente presente, alguma dificuldade em dispensar as memórias.

Grave= ao menos 2 X semana / angústia pronunciada, dificuldade considerável em dispensar as memórias.

◯ 0 - Ausente
◯ 1 - Leve / sublimiar
◯ 2 - Moderado / limiar
◯ 3 - Grave / marcadamente elevado
◯ 4 - Extremo / incapacitante

2. (B2) Sonhos angustiantes recorrentes nos quais o conteúdo e/ou sentimento do sonho estão relacionados com o evento traumático. Nota: Em crianças, pode haver pesadelos sem conteúdo identificável.

No mês passado, você teve algum sonho angustiante sobre (EVENTO)?

Descreva um sonho típico. (O que acontece?)
[Se não estiver claro:] (Esses sonhos acordam você?)

[Se sim:] (O que você sentiu ao acordar? Quanto tempo você levou para voltar a dormir?)

[Se a resposta for não voltar a dormir:] (Quanto tempo de sono você perde?)

O quanto esses sonhos incomodam a você?

B2. Angústia

◯ 0 - Mínima
◯ 1 - Claramente Presente
◯ 2 - Pronunciada
◯ 3 - Extrema

B2. Com que frequência você teve esses sonhos no ultimo mês?

_____
(número de vezes que apresentou o sonho)

B2. Pontuação:

Dimensões chave para classificação = frequência / intensidade da angústia

Moderado = ao menos 2 X mês / angústia claramente presente, menos de 1 hora de sono perdido

Grave = ao menos 2 X semana / angústia pronunciada, mais de 1 hora de sono perdido

◯ 0 - Ausente
◯ 1 - Leve / sublimiar
◯ 2 - Moderado / limiar
◯ 3 - Grave / marcadamente elevado
◯ 4 - Extremo / incapacitante

3. (B3) Reações dissociativas (p. ex.: flashbacks) nas quais o indivíduo sente ou age como se o evento traumático estivesse ocorrendo novamente. (Essas reações podem ocorrer em um continuum, com a expressão mais extrema na forma de uma perda completa de percepção do ambiente ao redor.) Nota: Em crianças a reencenarão específica do trauma pode ocorrer na brincadeira.

No mês passado houve vezes em que você de repente agiu ou sentiu como se o (EVENTO) estivesse realmente acontecendo de novo?

[Se não estiver claro:] (Isto é diferente do que pensar ou sonhar à respeito - agora estamos perguntando sobre flashbacks, quando você se sente como se estivesse realmente de volta a hora do (EVENTO), realmente revivendo o acontecido.)

O quanto parece como se (EVENTO) estivesse acontecendo novamente? (Você está confuso sobre onde realmente você está?)

O que você faz enquanto isso está acontecendo?
(Outras pessoas percebem o seu comportamento? O que eles dizem?)

Quanto tempo isso dura?

B3. Dissociação	◯ 0 - Mínima ◯ 1 - Claramente Presente ◯ 2 - Pronunciada ◯ 3 - Extrema
B3. Com que frequência isso aconteceu no ultimo mês?	_____ (número de flashbacks no último mês)
B3. Pontuação: Dimensões chave para classificação = frequência / intensidade da dissociação Moderado = ao menos 2 X mês / qualidade dissociativa claramente presente, pode manter alguma percepção do ambiente ao redor mas revive o evento de maneira distinta dos pensamentos e memórias. Grave = ao menos 2 X semana / qualidade dissociativa pronunciada, relata vívido reviver, e.x., com imagens, sons e cheiros.	◯ 0 - Ausente ◯ 1 - Leve / sublimiar ◯ 2 - Moderado / limiar ◯ 3 - Grave / marcadamente elevado ◯ 4 - Extremo / incapacitante

4. (B4) Sofrimento psicológico intenso ou prolongado ante a exposição a sinais internos ou externos que simbolizam ou se assemelham a algum aspecto do evento traumático.

No mês passado, você ficou emocionalmente abalado quando algo lhe lembrou (EVENTO)?

Quais os tipos de lembranças que deixaram você chateado?

O quanto essas lembranças incomodaram você?

Você é capaz de se acalmar quando isso acontece? (Quanto tempo leva?)

B4. Angústia

◯ 0 - Mínima
◯ 1 - Claramente Presente
◯ 2 - Pronunciada
◯ 3 - Extrema

B4. Com que frequência isso aconteceu no último mês?

_____

B4. Pontuação:

Dimensões chave para classificação = frequência / intensidade da angústia

Moderado = ao menos 2 X mês / angústia claramente presente, alguma dificuldade de recuperação.

Grave = ao menos 2 X semana / angústia pronunciada, dificuldade considerável de recuperação.

◯ 0 - Ausente
◯ 1 - Leve / sublimiar
◯ 2 - Moderado / limiar
◯ 3 - Grave / marcadamente elevado
◯ 4 - Extremo / incapacitante

5. (B5) Reações fisiológicas intensas a sinais internos ou externos que simbolizam ou se assemelham a algum aspecto do evento traumático.

No mês passado você teve alguma reação física quando algo o fez lembrar o (EVENTO)?

Você pode dar alguns exemplos? (Seu coração dispara ou sua respiração muda? E quanto a suar ou sentir-se realmente tenso ou trêmulo?)

Quais os tipos de recordação que desencadearam essas reações?

Quanto tempo você levou para se recuperar?

B5. Reação fisiológica

◯ 0 - Mínima
◯ 1 - Claramente Presente
◯ 2 - Pronunciada
◯ 3 - Extrema

B5. Com que frequência isso aconteceu no ultimo mês?

_____
(Número de vezes)

B5. Pontuação: Dimensões chave para classificação = frequência / intensidade de excitação fisiológica Moderado = ao menos 2 X mês / reação claramente presente, alguma dificuldade de recuperação. Grave= ao menos 2 X semana / reação pronunciada, dificuldade considerável de recuperação	◯ 0 - Ausente ◯ 1 - Leve / sublimiar ◯ 2 - Moderado / limiar ◯ 3 - Grave / marcadamente elevado ◯ 4 - Extremo / incapacitante

Critério C: Evitação persistente de estímulos associados ao evento traumático, começando após a ocorrência do evento, conforme evidenciado por um ou ambos dos seguintes aspectos:

6. (C1) Evitação ou esforços para evitar recordações, pensamentos ou sentimentos angustiantes acerca de ou associados de perto com o evento traumático.

No mês passado você tentou evitar pensamentos ou sentimentos sobre (EVENTO)?

Quais tipos de pensamentos ou sentimentos você evita?

Quanto esforço você faz para tentar evitar esses pensamentos ou sentimentos? (Que tipo de coisas você faz?)

C1. Evitação	◯ 0 - Mínima ◯ 1 - Claramente Presente ◯ 2 - Pronunciada ◯ 3 - Extrema
C1. Com que frequência isso aconteceu no ultimo mês?	_____ (Número de vezes que aconteceu no último mês)
C1. Pontuação: Dimensões chave para classificação = frequência / intensidade da evitação Moderado = ao menos 2 X mês / evitação claramente presente. Grave = ao menos 2 X semana / evitação pronunciada.	◯ 0 - Ausente ◯ 1 - Leve / sublimiar ◯ 2 - Moderado / limiar ◯ 3 - Grave / marcadamente elevado ◯ 4 - Extremo / incapacitante

7. (C2) Evitação ou esforços para evitar lembranças externas (pessoas, lugares, atividades, objetos, situações) que despertem recordações, pensamentos ou sentimentos angustiantes acerca de ou associados de perto ao evento traumático

No mês passado, você tentou evitar coisas que lembrassem você do (EVENTO), como certas pessoas, lugares ou situações?

Quais tipos de coisas você evitou?

Quanto esforço você fez para evitar essas lembranças? (Você tem que se planejar ou mudar suas atividades para evita-las?)

[Se não estiver claro:] (De forma geral, o quanto isso representa um problema para você? Como as coisas seriam diferentes se você não tivesse que evitar essas recordações?)

C2. Evitação

◯ 0 - Mínima
◯ 1 - Claramente Presente
◯ 2 - Pronunciada
◯ 3 - Extrema

C2. Com que frequência isso aconteceu no ultimo mês?

_____
(Número de vezes que aconteceu no último mês)

C2. Pontuação:

Dimensões chave para classificação = frequência / intensidade da evitação

Moderado = ao menos 2 X mês / evitação claramente presente.

Grave = ao menos 2 X semana / evitação pronunciada.

◯ 0 - Ausente
◯ 1 - Leve / sublimiar
◯ 2 - Moderado / limiar
◯ 3 - Grave / marcadamente elevado
◯ 4 - Extremo / incapacitante

Critério D: Alterações negativas em cognições e humor associadas ao evento traumático, começando ou piorando depois da ocorrência de tal evento, conforme evidenciado por dois ou mais dos seguintes aspectos:

8. (D1) Incapacidade de recordar um aspecto importante do evento traumático (geralmente em decorrência a amnesia dissociativa, e não a outros fatores, como traumatismo craniano, álcool ou drogas).

No mês passado, você teve dificuldade para se lembrar de alguma parte importante do (EVENTO)? (Você sente que há lacunas na sua memória sobre o [EVENTO]?)

Quais partes você teve dificuldade em se lembrar?

Você sente que você deveria ser capaz de lembrar dessas coisas?

[Se não estiver claro:] (Por que você acha que não pode? Você sofreu traumatismo craniano durante [EVENTO]? Você estava inconsciente? Você estava sob o efeito de álcool ou drogas?) [Classifique 0=Ausente se devido ao traumatismo craniano ou perda de consciêcia ou intoxicação durante o evento]

[Se não estiver claro:] (Este esquecimento é normal? Ou você acha que possa ter bloqueado porque seria doloroso demais se lembrar?) [Classifique 0=Ausente se devido apenas ao esquecimento normal]

Você seria capaz de se lembrar dessas coisas se você tentar?

D1. Dificuldade de lembrar

◯ 0 - Mínima
◯ 1 - Claramente Presente
◯ 2 - Pronunciada
◯ 3 - Extrema

D1. No mês passado, quantas partes importantes do (EVENTO) você teve dificuldade em se lembrar? (Quais partes você ainda se lembra?)

_____
(Quantidade de aspectos importantes não lembrados)

D1. Pontuação:

Dimensões chave para classificação = quantidade de eventos que não lembrou / intensidade da incapacidade de recordar

Moderado = ao menos um aspecto importante/ incapacidade de recordar claramente presente, alguma recordação é possível com esforço.

Grave = vários aspectos importantes/ incapacidade de recordar marcada, pouca recordação mesmo com esforço

◯ 0 - Ausente
◯ 1 - Leve / sublimiar
◯ 2 - Moderado / limiar
◯ 3 - Grave / marcadamente elevado
◯ 4 - Extremo / incapacitante

9. (D2) Crenças ou expectativas negativas persistentes e exageradas a respeito de si mesmo, dos outros e do mundo. (p.ex. , , ,).

No mês passado, você teve fortes crenças negativas sobre si mesmo(a), outras pessoas ou sobre o mundo?

Você pode me dar alguns exemplos? (E quanto a acreditar em coisas como , , , ?)

O quão fortes são essas crenças? (O quão convencido você está de que essas crenças são realmente verdade? Você pode ver formas alternativas de pensar sobre isso?)

D2. Convicção

◯ 0 - Mínima
◯ 1 - Claramente Presente
◯ 2 - Pronunciada
◯ 3 - Extrema

D2. Com que frequência no ultimo mês você se sentiu assim?

_____
(% do tempo)

D2. Essas crenças começaram ou ficaram piores depois do [EVENTO]? (Você acha que elas são relacionadas ao [EVENTO]? Como?)

Relação com o trauma

◯ 0 - Definitiva
◯ 1 - Provável
◯ 2 - Improvável

D2. Pontuação:

Dimensões chave para classificação = frequência / intensidade das crenças

Moderado = parte do tempo (20-30%) / expectativas negativas exageradas claramente presentes, alguma dificuldade em considerar crenças mais realistas.

Grave = grande parte do tempo (50-60%) / expectativas negativas exageradas pronunciadas, considerável dificuldade em considerar crenças mais realistas.

◯ 0 - Ausente
◯ 1 - Leve / sublimiar
◯ 2 - Moderado / limiar
◯ 3 - Grave / marcadamente elevado
◯ 4 - Extremo / incapacitante

10. (D3) Cognições distorcidas persistentes a respeito das causas ou consequências do evento traumático que levam o individuo a culpar a si mesmo ou a outros.

No mês passado você se culpou pelo (EVENTO) ou pelo que aconteceu como resultado dele? Me conte mais sobre isso. (Em que sentido você se considera causador [EVENTO]? É por causa de alguma coisa que você fez? Ou alguma coisa que você acha que você deveria ter feito e não fez? É por causa de alguma coisa sobre você em geral?)

E quanto a culpar alguém mais pelo (EVENTO) ou pelo que aconteceu como resultado disso? Me conte mais sobre isso. (Em que sentido você considera [OUTROS] como causadores [EVENTO]? É por causa de algo que eles fizeram? Ou alguma coisa que você acha que eles deveriam ter feito, mas não fizeram?)

Quanto você culpa (SI MESMO OU OS OUTROS)?

O quão convencido você está de que (VOCÊ OU OS OUTROS) são realmente responsáveis pelo o que aconteceu? (Outras pessoas concordam com você? Você consegue ver outra maneira de pensar sobre isso?)

[Classificação 0=Ausente se culpa somente do perpetrador, por exemplo alguém que causou deliberadamente o evento e tinha a intenção de prejudicar]

D3. Convicção

◯ 0 - Mínima
◯ 1 - Claramente Presente
◯ 2 - Pronunciada
◯ 3 - Extrema

D3. Com que frequência no ultimo mês você se sentiu assim?

_____
(% do tempo)

D3. Pontuação:

Dimensões chave para classificação = frequência / intensidade do sentimento de culpa

Moderado = parte do tempo (20-30%) / culpa distorcida claramente presente, alguma dificuldade em considerar crenças mais realistas.

Grave = grande parte do tempo (50-60%) / culpa distorcida pronunciada, grande dificuldade em considerar crenças mais realistas

◯ 0 - Ausente
◯ 1 - Leve / sublimiar
◯ 2 - Moderado / limiar
◯ 3 - Grave / marcadamente elevado
◯ 4 - Extremo / incapacitante

11. (D4) Estado emocional negativo persistente (p. ex., medo, pavor, raiva, culpa ou vergonha).

No mês passado, você teve algum forte sentimento negativo como: medo, horror, raiva, culpa ou vergonha?

Você pode dar alguns exemplos? (Quais sentimentos negativos você teve?)

Quão fortes são esses sentimentos negativos?

Quão bem você é capaz de lidar com eles?

D4. Emoções negativas

◯ 0 - Mínima
◯ 1 - Claramente Presente
◯ 2 - Pronunciada
◯ 3 - Extrema

D4. Com que frequência no ultimo mês você se sentiu assim?

_____
(% do tempo)

D4. Esses sentimentos negativos começaram ou ficaram piores depois do [EVENTO]? (Você acha que elas são relacionadas ao [EVENTO]? Como?)

Relação com o trauma

◯ 0 - Definitiva
◯ 1 - Provável
◯ 2 - Improvável

D4. Pontuação:

Dimensões chave para classificação = frequência / intensidade das emoções negativas

Moderado = parte do tempo (20-30%) / emoções negativas claramente presentes, alguma dificuldade em lidar com elas.

Grave = grande parte do tempo (50-60%) / emoções negativas pronunciadas, considerável dificuldade em lidar com elas.

◯ 0 - Ausente
◯ 1 - Leve / sublimiar
◯ 2 - Moderado / limiar
◯ 3 - Grave / marcadamente elevado
◯ 4 - Extremo / incapacitante

12. (D5) Interesse ou participação bastante diminuída em atividades significativas.

No mês passado, você se interessou menos por atividades que você costumava gostar?

Por quais tipos de coisas você perdeu interesse ou não faz mais com a frequência que costumava fazer? (Algo mais?)

Por quê? [Classifique 0=Ausente se diminuiu a participação devido à falta de oportunidade, incapacidade física ou mudança do desenvolvimento apropriado nas atividades preferidas]

Quão forte é a sua perda de interesse? (Você ainda apreciaria [ATIVIDADES] uma vez começada?)

D5. Perda de interesse	◯ 0 - Mínima ◯ 1 - Claramente Presente ◯ 2 - Pronunciada ◯ 3 - Extrema
D5. Em geral, no mês passado, por quantas das suas atividades habituais você esteve menos interessado(a)?	_____ (% das atividades)
D5. Quais tipos de coisas você ainda gosta de fazer? Essa perda de interesse começou ou ficou pior depois [EVENTO]? (Você acha que está relacionado ao [EVENTO]? Como?) Relação com o trauma	◯ 0 - Definitiva ◯ 1 - Provável ◯ 2 - Improvável
D5. Pontuação: Dimensões chave para classificação = frequência / intensidade da perda de interesse Moderado = algumas atividades (20-30%) / perda de interesse claramente presente, mas ainda algum prazer nas atividades. Grave = muitas atividades (50-60%) / perda de interesse pronunciada, pouco interesse em participar de atividades	◯ 0 - Ausente ◯ 1 - Leve / sublimiar ◯ 2 - Moderado / limiar ◯ 3 - Grave / marcadamente elevado ◯ 4 - Extremo / incapacitante

13. (D6) Sentimentos de distanciamento e alienação em relação aos outros.

No mês passado você se sentiu distante ou isolado(a) das outras pessoas? Conte mais sobre isso.

Quão forte são esses sentimentos de distanciamento ou isolamento dos outros? (De quem você se sente mais próximo? Com quantas pessoas você se sente confortável em falar sobre questões pessoais?)

D6. Perda de interesse	◯ 0 - Mínima ◯ 1 - Claramente Presente ◯ 2 - Pronunciada ◯ 3 - Extrema
D6. Durante quanto tempo você se sentiu assim no mês passado?	_____ (% do tempo)
D6. Essa sensação de estar distante ou isolado começou ou piorou após o (EVENTO)? (Você acha que tem relação com [EVENTO]? Como?) Relação com trauma	◯ 0 - Definitiva ◯ 1 - Provável ◯ 2 - Improvável
D6. Pontuação: Dimensões chave para classificação = frequência / intensidade de alienação ou distanciamento Moderado = parte do tempo (20-30%) / sensação de distanciamento claramente presente, mas ainda sente alguma conexão interpessoal. Grave = grande parte do tempo (50-60%) / pronunciada sensação de distanciamento ou alienação da maior parte das pessoas, pode sentir-se próximo a apenas uma ou duas pessoas.	◯ 0 - Ausente ◯ 1 - Leve / sublimiar ◯ 2 - Moderado / limiar ◯ 3 - Grave / marcadamente elevado ◯ 4 - Extremo / incapacitante

14. (D7) Incapacidade persistente de sentir emoções positivas (p.ex., incapacidade de vivenciar sentimentos de felicidade, satisfação ou amor).

No mês passado, houve momentos em que você teve dificuldades em experimentar sentimentos positivos
como amor ou felicidade?

Conte-me mais sobre isso. (Quais sentimentos são difíceis de experimentar?)

Qual tem sido sua dificuldade em experimentar sentimentos positivos? (Você ainda é capaz de experimentar algum sentimento positivo?)

D7. Redução de emoções positivas	◯ 0 - Mínima ◯ 1 - Claramente Presente ◯ 2 - Pronunciada ◯ 3 - Extrema
D7. Durante quanto tempo você se sentiu assim no mês passado?	_____ (% do tempo)

D7. Essa dificuldade de experimentar sentimentos positivos começou ou ficou pior depois [EVENTO]? Você acha que isso está relacionado ao [EVENTO]? Como?

Relação com o trauma

◯ 0 - Definitiva
◯ 1 - Provável
◯ 2 - Improvável

D7. Pontuação:

Dimensões chave para classificação = frequência / intensidade da redução de emoções positivas

Moderado = parte do tempo (20-30%) / redução de experiências emocionais positivas claramente presente, mas ainda capaz de vivenciar algumas emoções positivas.

Grave = grande parte do tempo (50-60%) / redução pronunciada da experiência em toda gama de emoções positivas.

◯ 0 - Ausente
◯ 1 - Leve / sublimiar
◯ 2 - Moderado / limiar
◯ 3 - Grave / marcadamente elevado
◯ 4 - Extremo / incapacitante

Critério E: Alterações marcantes na excitação e na reatividade associadas ao evento traumático, começando ou piorando após o evento, conforme evidenciado por dois ou mais dos seguintes aspectos:

15. (E1) Comportamento irritadiço ou surtos de raiva (com pouca ou nenhuma provocação) geralmente expressos sob a forma de agressão verbal ou física em relação a pessoas e objetos.

No mês passado houve momentos em que você se sentiu especialmente irritado ou com raiva e demonstrou isso em seu comportamento?

Você pode dar alguns exemplos? (Como você demonstra isso? Você levanta a voz ou grita? Atira ou bate coisas? Empurra ou acerta outras pessoas?)

E1. Agressão

◯ 0 - Mínima
◯ 1 - Claramente Presente
◯ 2 - Pronunciada
◯ 3 - Extrema

E1. Quantas vezes no mês passado?

_____
(Número de vezes)

E1. Esse comportamento começou ou ficou pior depois [EVENTO]? (Você acha que está relacionado ao [EVENTO])? Como?

Relação com o trauma

◯ 0 - Definitiva
◯ 1 - Provável
◯ 2 - Improvável

E1. Pontuação: Dimensões chave para classificação = frequência / intensidade do comportamento agressivo Moderado = ao menos 2 X mês / agressividade claramente presente, principalmente verbal. Grave = ao menos 2 X semana /agressividade pronunciada, ao menos alguma agressão física.	◯ 0 - Ausente ◯ 1 - Leve / sublimiar ◯ 2 - Moderado / limiar ◯ 3 - Grave / marcadamente elevado ◯ 4 - Extremo / incapacitante

16. (E2) Comportamento imprudente ou autodestrutivo.

No mês passado houve vezes em que você tenha se colocado em risco ou feito coisas que poderiam te prejudicar?

Você pode dar alguns exemplos?

Quanto risco você correu? (O quão perigosos são esses comportamentos? Você se feriu ou se prejudicou de alguma forma?)

E2. Risco	◯ 0 - Mínima ◯ 1 - Claramente Presente ◯ 2 - Pronunciada ◯ 3 - Extrema
E2. Com que frequência você tem assumido esse tipo de risco no ultimo mês?	_____ (Número de vezes)
E2. Esse comportamento começou ou ficou pior depois [EVENTO]? (Você acha que isso está relacionado ao [EVENTO]? Como?) Relação com o trauma	◯ 0 - Definitiva ◯ 1 - Provável ◯ 2 - Improvável
E2. Pontuação: Dimensões chave para classificação = frequência / grau do risco Moderado = ao menos 2 X mês / risco claramente presente, pode ter se prejudicado. Grave = ao menos 2 X semana / risco pronunciado, dano real ou alta probabilidade de dano.	◯ 0 - Ausente ◯ 1 - Leve / sublimiar ◯ 2 - Moderado / limiar ◯ 3 - Grave / marcadamente elevado ◯ 4 - Extremo / incapacitante

17. (E3) Hipervigilância.

No mês passado, você ficou especialmente em alerta ou vigilante, mesmo quando não havia ameaça ou perigo específico? (Você se sente como se tivesse que ficar em constante alerta?)

Você pode dar alguns exemplos? (Que tipo de coisas você faz quando esta em alerta ou vigilante?)

[Se não estiver claro:] (O que faz você reagir assim?
Você sente que está em perigo ou ameaçado de alguma forma? Você se sente dessa forma mais do que a maioria das pessoas se sentiriam na mesma situação?)

E3. Hipervigilância.	◯ 0 - Mínima ◯ 1 - Claramente Presente ◯ 2 - Pronunciada ◯ 3 - Extrema
E3. Durante quanto tempo você se sentiu assim no mês passado?	_____ (% do tempo)
E3. Ficar muito alerta ou vigilante começou ou ficou pior depois do [EVENTO]? (Você acha que está relacionado ao [EVENTO]? Como?) Relação com o trauma	◯ 0 - Definitiva ◯ 1 - Provável ◯ 2 - Improvável
E3. Pontuação: Dimensões chave para classificação = frequência / intensidade da hipervigilância Moderado = alguns momentos (20-30%) / hipervigilância claramente presente, p.ex., estar vigilante em publico, atenção altamente voltada para ameaças. Grave = grande parte do tempo (50-60%) / pronunciada hipervigilancia, p.ex., procurar por perigo no ambiente, fazer rituais de segurança, exagerada preocupação com a segurança de si/familiares/casa.	◯ 0 - Ausente ◯ 1 - Leve / sublimiar ◯ 2 - Moderado / limiar ◯ 3 - Grave / marcadamente elevado ◯ 4 - Extremo / incapacitante

18. (E4) Resposta de sobressalto exagerada.

No mês passado você teve qualquer reação de forte sobressalto?

Quais tipos de coisas assustou você?

Quão fortes são essas reações de sobressalto? (Quão fortes são se comparado a como a maioria das pessoas reagiriam? Você faz algo que outras pessoas percebam?)

Quanto tempo leva para você se recuperar?

E4. Sobressalto

◯ 0 - Mínima
◯ 1 - Claramente Presente
◯ 2 - Pronunciada
◯ 3 - Extrema

E4. Quantas vezes isso aconteceu no mês passado?

_____
(Número de vezes)

E4. Essas reações de sobressalto começaram ou ficaram piores depois do [EVENTO]? (Você acha que elas estão relacionadas ao [EVENTO]? Como?)

Relação com o trauma

◯ 0 - Definitiva
◯ 1 - Provável
◯ 2 - Improvável

E4. Pontuação:

Dimensões chave para classificação = frequência / intensidade do sobressalto

Moderado = ao menos 2 X mês / sobressalto claramente presente, recuperação com alguma dificuldade.

Grave = ao menos 2 X semana / sobressalto pronunciado, excitação mantida, dificuldade considerável de recuperação.

◯ 0 - Ausente
◯ 1 - Leve / sublimiar
◯ 2 - Moderado / limiar
◯ 3 - Grave / marcadamente elevado
◯ 4 - Extremo / incapacitante

19. (E5) Problemas com a concentração.

No mês passado você teve qualquer problema de concentração?

Você poderia me dar alguns exemplos?

Você é capaz de se concentrar se você realmente tentar?

E5. Problema de concentração

◯ 0 - Mínima
◯ 1 - Claramente Presente
◯ 2 - Pronunciada
◯ 3 - Extrema

E5. Durante quanto tempo no mês passado você teve problemas de concentração?

_____
(% de tempo)

E5. Esses problemas de concentração começaram ou ficaram piores depois [EVENTO]? (Você acha que eles estão relacionados ao [EVENTO]? Como?)

Relação com o trauma

◯ 0 - Definitiva
◯ 1 - Provável
◯ 2 - Improvável

E5. Pontuação:

Dimensões chave para classificação = frequência / intensidade dos problemas de concentração

Moderado = parte do tempo (20-30%) / problema de concentração claramente presente, alguma dificuldade, mas consegue se concentrar com esforço.

Grave = grande parte do tempo (50-60%) / pronunciado problema de concentração, dificuldade considerável mesmo com esforço.

◯ 0 - Ausente
◯ 1 - Leve / sublimiar
◯ 2 - Moderado / limiar
◯ 3 - Grave / marcadamente elevado
◯ 4 - Extremo / incapacitante

20. (E6) Perturbação do sono (p.ex., dificuldade para iniciar ou manter o sono, ou sono agitado).

Durante o mês passado, você teve problemas para começar a dormir ou manter o sono?

Quais tipos de problemas? (Quanto tempo você levou para dormir? Quantas vezes você acordou durante a noite? Você acordou mais cedo do que pretendia?)

Qual o total de horas que você dorme em cada noite?

Quantas horas você acha que você deveria estar dormido?

E6. Perturbação do sono

◯ 0 - Mínima
◯ 1 - Claramente Presente
◯ 2 - Pronunciada
◯ 3 - Extrema

E6. Quantas vezes no mês passado você teve problemas para dormir?

_____
(Número de vezes)

E6. Esses problemas para dormir começaram ou ficaram piores após (EVENTO)? (Você acha que eles estão relacionados ao [EVENTO]? Como?)

Relação com o trauma

◯ 0 - Definitiva
◯ 1 - Provável
◯ 2 - Improvável

E6. Pontuação:

Dimensões chave para classificação = frequência / intensidade da perturbação do sono

Moderado = ao menos 2 X mês / perturbação do sono claramente presente, latência claramente maior ou clara dificuldade em continuar dormindo, 30-90 minutos perdidos de sono.

Grave = ao menos 2 X semana / pronunciada perturbação do sono, latência consideravelmente maior ou marcada dificuldade em continuar dormindo, 90 min a 3 horas de sono perdido.

◯ 0 - Ausente
◯ 1 - Leve / sublimiar
◯ 2 - Moderado / limiar
◯ 3 - Grave / marcadamente elevado
◯ 4 - Extremo / incapacitante

Critério F: A perturbação (Critérios B, C, D e E) dura mais de um mês

21. Início dos sintomas

[Se não estiver claro:] Quando você começou a apresentar (SINTOMAS DE TEPT) que você me contou? (Quanto tempo após o trauma os sintomas começaram? Mais de seis meses?)

_____ (Total # meses de atraso do início)

21. Com início tardio (> ou igual a 6 meses)?

◯ Yes
◯ No

22. Duração dos sintomas

[Se não estiver claro:] Por quanto tempo no total duraram (SINTOMAS DE TEPT)?

_____ (Total # meses de duração)

22. Mais de 1 mês de duração?

◯ Yes
◯ No

Critério G: A perturbação causa sofrimento clinicamente significativo e prejuízo social, profissional ou em outras áreas importantes na vida do indivíduo.

23. Sofrimento subjetivo

No geral, no mês passado, o quanto você ficou incomodado com esses (SINTOMAS DE TEPT) que você narrou? [Considere angústia relatada nos itens anteriores]?

◯ 0 - Nenhum pouco
◯ 1 - Leve, angústia mínima
◯ 2 - Moderado angústia claramente presente, mas ainda administrável
◯ 3 - Grave angústia considerável
◯ 4 - Extremo, angústia incapacitante

24. Prejuízo no funcionamento social.

No mês passado, esses (SINTOMAS DE TEPT) afetaram suas relações com as outras pessoas? Como? [Considere prejuízo no funcionamento social relatado nos itens anteriores]

◯ 0 - Nenhum impacto negativo
◯ 1 - Impacto leve, prejuízo mínimo no funcionamento social
◯ 2 - Impacto moderado, comprometimento definitivo, mas muitos dos aspectos do funcionamento social ainda intactos
◯ 3 - Impacto grave, prejuízo acentuado, alguns aspectos do funcionamento social ainda intactos
◯ 4 - Impacto extremo, pouco ou nenhum funcionament social

25. Prejuízo ocupacional ou em outra importante área do funcionamento.

[Se não estiver claro:] Você está trabalhando atualmente?

[Se sim:] No mês passado, ter esses (SINTOMAS DE TEPT) afetaram seu trabalho ou sua capacidade de trabalhar? Como? [Considere história relatada de trabalho, incluindo número e duração dos trabalhos, bem como a qualidade das relações de trabalho. Se o funcionamento pré-mórbido não estiver claro, pergunte sobre as experiências antes do trauma. Para trauma em crianças ou adolescentes, avalie o desempenho escolar pré-trauma e a possível presença de problemas de comportamento]

[Se não:] Esses (SINTOMAS DE TEPT) afetaram qualquer outra parte importante da sua vida? [Se for apropriado sugira exemplos como familiar, trabalhos domésticos, escolares, voluntários, etc.] Como?

- ◯ 0 - Nenhum impacto negativo
- ◯ 1 - Impacto leve, prejuízo mínimo no funcionamento ocupacional / outra importante área de funcionamento
- ◯ 2 - Impacto moderado, comprometimento definitivo, mas com muitos aspectos ainda intactos do funcionamento ocupacional / outra importante área de funcionamento
- ◯ 3 - Impacto grave, prejuízo acentuado, poucos aspectos ainda intactos do funcionamento ocupacional / outra importante área de funcionamento
- ◯ 4 - Impacto extremo, pouco ou nenhum funcionament ocupacional / outra importante área de funcionamento

Classificação Global

26. Validade Global.

Estima a validade total de respostas. Considere fatores como o cumprimento da entrevista, o estado mental (por exemplo: problemas de concentração, compreensão dos itens, dissociação), e evidência de esforços para exagerar ou minimizar sintomas.

- ◯ 0 - Excelente, não há razão para suspeitar de respostas inválidas
- ◯ 1 - Boa presença de fatores que podem afetar negativamente a validade
- ◯ 2 - Justa presença de fatores que definitivamente reduzem a validade
- ◯ 3 - Pobre validade substancialmente reduzida
- ◯ 4 - Respostas inválidas gravemente prejudicado pelo estado mental ou possível esforço para exagerar ou minimizar sintomas deliberadamente

27. Gravidade Global.

Estima a gravidade geral de sintomas do TEPT. Considere o grau de sofrimento subjetivo, o grau de comprometimento funcional, observações do comportamento na entrevista e juízo em relação ao estilo de informação.

- ◯ 0 - Nenhum sintoma clinicamente significativo , nenhuma angústia e sem comprometimento funciona
- ◯ 1 - Leve angústia ou comprometimento funcional mínimo
- ◯ 2 - Moderado, angústia definitiva ou prejuízo funcional, mas funções satisfatórias com esforço
- ◯ 3 - Grave, angústia considerável ou incapacidade
- ◯ funcionamento limitado mesmo com esforço
- ◯ 4 - Extremo, sofrimento ou angústia acentuada em duas ou mais áreas importantes do funcionamento

28. Melhora Global.

Avalie a melhoria global total desde a classificação anterior. Classifique o grau de mudança e, segundo seu julgamento, se foi ou não, devido ao tratamento.

- ◯ 0 - Assintomático
- ◯ 1 - Melhora considerável
- ◯ 2 - Melhora moderada
- ◯ 3 - Leve melhora
- ◯ 4 - Nenhuma melhora
- ◯ 5 - Informação insuficiente

Especifique se com sintomas dissociativos (subtipo): Os sintomas do indivíduo satisfazem os critérios de TEPT, e, além disso, em resposta ao estressor, o indivíduo tem sintomas persistentes ou recorrentes de:

29. (1) Despersonalização: Experiências persistentes ou recorrentes de sentir-se separado e como se fosse um observador externo dos processos mentais ou do corpo (p.ex., sensação de estar em um sonho, sensação de irrealidade de si mesmo ou do corpo ou como se estivesse em câmera lenta).

No mês passado houve vezes em que você sentiu como se você estivesse separado de si mesmo, como se você estivesse assistindo a si mesmo de fora ou observando seus pensamentos e sentimentos, como se você fosse outra pessoa?

[Se não:] (E quanto a sensação de como se você estivesse em um sonho, mesmo acordado? Sensação como se algo em você não fosse real? Sentindo como se o tempo estivesse passando mais lentamente?)

Conte mais sobre isso.

Quão forte é essa sensação? (Você perde a noção de onde você realmente está ou sobre o que está acontecendo?)

O que você faz enquanto isso está acontecendo? (Outras pessoas percebem o seu comportamento? O que eles dizem?)

Por quanto tempo isso dura?

[Se não estiver claro:] (Isso foi devido a efeitos de álcool ou drogas? E quanto a alguma condição médica como convulsões?)

[Classifique 0=Ausente se devido aos efeitos de substância ou outra condição médica]

29. Dissociação.

- ◯ 0 - Mínima
- ◯ 1 - Claramente Presente
- ◯ 2 - Pronunciada
- ◯ 3 - Extrema

29. Quantas vezes isso aconteceu no mês passado?

_____
(Número de vezes)

29. Pontuação:

Dimensões chave para classificação = frequência / intensidade da dissociação

Moderado = ao menos 2 X mês / qualidade dissociativa claramente presente porém transitória, retém algum senso de realidade de si e consciência do ambiente.

Grave = ao menos 2 X semana / qualidade dissociativa pronunciada, marcada sensação de distanciamento e irrealidade.

◯ 0 - Ausente
◯ 1 - Leve / sublimiar
◯ 2 - Moderado / limiar
◯ 3 - Grave / marcadamente elevado
◯ 4 - Extremo / incapacitante

30. (2) Desrealização: Experiências persistentes ou recorrentes de irrealidade do ambiente ao redor (p.ex., o mundo ao redor do indivíduo é sentido como irreal, onírico, distante ou distorcido).

No mês passado, houve vezes em que as coisas ao seu redor pareciam irreais ou muito estranhas e desconhecidas?

[Se não:] (As coisas acontecendo ao seu redor pareciam com um sonho ou uma cena de cinema?
Pareciam distantes ou distorcidas?)

Conte mais à respeito.

Quão forte é essa sensação? (Você perde a noção de onde você realmente está ou o que está realmente acontecendo?)

O que você faz enquanto isso está acontecendo? (Outras pessoas percebem o seu comportamento? O que eles dizem?)

Quanto tempo isso dura?

[Se não estiver claro:] (Isso foi devido a efeitos de álcool ou drogas? E quanto a condição médica como: convulsões?)

[Classifique 0=Ausente se devido aos efeitos de substância ou outra condição médica]

30. Dissociação.

◯ 0 - Mínima
◯ 1 - Claramente Presente
◯ 2 - Pronunciada
◯ 3 - Extrema

30. Quantas vezes isso aconteceu no mês passado?

_____
(número de vezes)

30. Pontuação:

Dimensões chave para classificação = frequência / intensidade da dissociação

Moderado = ao menos 2 X mês / qualidade dissociativa claramente presente porém transitória, retém algum senso realista do ambiente.

Grave = ao menos 2 X semana / qualidade dissociativa pronunciada, marcada sensação de irrealidade .

◯ 0 - Ausente
◯ 1 - Leve / sublimiar
◯ 2 - Moderado / limiar
◯ 3 - Grave / marcadamente elevado
◯ 4 - Extremo / incapacitante

CAPS-5 SUMÁRIO

Nome:_____ ID#:_____ Entrevistador:_____ Estudo:_____ Data:_____

A. Exposição a episódio concreto ou ameaça de morte, lesão grave ou violência sexual	
Critério A atendido?	**0 = NÃO 1 = SIM**

B. Sintomas de intrusão (necessidade de 1 para o diagnóstico)	*Mês Passado*	
	Gravidade	*Sx (Grv ≥ 2)?*
(1) B1 – Memórias intrusivas		**0 = NÃO 1 = SIM**
(2) B2 – Sonhos angustiantes		**0 = NÃO 1 = SIM**
(3) B3 – Reações dissociativas		**0 = NÃO 1 = SIM**
(4) B4 – Sofrimento psicológico		**0 = NÃO 1 = SIM**
(5) B5 – Reações fisiológicas		**0 = NÃO 1 = SIM**
B subtotal	***B Grv =***	***# B Sx =***

C. Sintomas de evitação (necessidade 1 para o diagnóstico)	*Mês Passado*	
	Gravidade	*Sx (Grv ≥ 2)?*
(6) C1 – Evitação de memórias, pensamentos, sentimentos		**0 = NÃO 1 = SIM**
(7) C2 – Evitação de lembranças externas		**0 = NÃO 1 = SIM**
C subtotal	***C Grv =***	***# C Sx =***

D. Cognições e sintomas de humor (necessidade 2 para o diagnóstico)	*Mês Passado*	
	Gravidade	*Sx (Sev ≥ 2)?*
(8) D1 – Incapacidade de recordar aspecto importante do evento		**0 = NÃO 1 = SIM**
(9) D2 – Crenças negativas ou expectativas exageradas		**0 = NÃO 1 = SIM**
(10) D3 – Cognições distorcidas conduzindo para a culpa		**0 = NÃO 1 = SIM**
(11) D4 – Estado emocional negativo persistente		**0 = NÃO 1 = SIM**
(12) D5 – Diminuição do interesse ou participação em atividades		**0 = NÃO 1 = SIM**
(13) D6 – Distanciamento ou estranhamento dos outros		**0 = NÃO 1 = SIM**
(14) D7 – Incapacidade persistente de experimentar emoções positivas		**0 = NÃO 1 = SIM**
D subtotal	***D Sev =***	***# D Sx =***

E. Excitação e reatividade de sintomas (necessidade 2 para o diagnóstico)	*Mês Passado*	
	Gravidade	*Sx (Sev ≥ 2)?*
(15) E1 – Comportamento irritadiço ou surtos de raiva		**0 = NÃO 1 = SIM**
(16) E2 – Comportamento imprudente ou autodestrutivo		**0 = NÃO 1 = SIM**
(17) E3 – Hipervigilância		**0 = NÃO 1 = SIM**
(18) E4 – Resposta de sobressalto exagerada		**0 = NÃO 1 = SIM**
(19) E5 – Problemas de concentração		**0 = NÃO 1 = SIM**
(20) E6 – Perturbação do sono		**0 = NÃO 1 = SIM**
E subtotal	***E Sev =***	***# E Sx =***

TEPT total	*Mês Passado*	
	Total Grav.	*Total # Sx*
Soma de subtotais (B+C+D+E)		

F. Duração do distúrbio	*Atual*
(22) Duração do distúrbio ≥ 1 mês?	**0 = NÃO 1 = SIM**

G. Sofrimento ou prejuízo (necessidade 1 para o diagnóstico)	*Mês Passado*	
	Gravidade	*Cx (Sev ≥ 2)?*
(23) Sofrimento subjetivo		**0 = NÃO 1 = SIM**
(24) Prejuízo no funcionamento social		**0 = NÃO 1 = SIM**
(25) Prejuízo no funcionamento ocupacional		**0 = NÃO 1 = SIM**
G subtotal	*G Grav. =*	*# G Cx =*

Classificação Global	*Mês Passado*
(26) Validade global	
(27) Gravidade global	
(28) Melhora global	

Sintomas dissociativos (necessidade 1 para o subtipo)	*Mês Passado*	
	Gravidade	*Sx (Sev ≥ 2)?*
(29) 1 – Despersonalização		**0 = NÃO 1 = SIM**
(30) 2 – Desrealização		**0 = NÃO 1 = SIM**
Subtotal Dissociativo	*Diss Sev =*	*# Diss Sx =*

Diagnóstico de TEPT	*Mês Passado*
TEPT PRESENTE – TODOS CRITÉRIOS (A-G) ATENDIDOS?	**0 = NÃO 1 = SIM**
Com sintomas dissociativos	**0 = NÃO 1 = SIM**
(21) Com início tardio (≥ 6 meses)	**0 = NÃO 1 = SIM**

ANNEX 2: LIFE EVENTS CHECKLIST (VERSION DSM-5)

ADAPTED TO BRAZILIAN-PORTUGUESE

C. LEC-5

Record ID _____

Critério A: Exposição a episódio concreto ou ameaça de morte, lesão grave ou violência sexual em uma (ou mais) das seguintes formas:
1. Vivenciar diretamente o(s) evento(s) traumático(s)
2. Testemunhar pessoalmente o evento ocorrido com outras pessoas
3. Saber que o evento ocorreu com algum membro da família próximo, ou a algum amigo próximo. Nos casos de ameaça real de morte de um membro da família ou amigo, o(s) evento(s) deve(m) ter sido violento(s) ou acidental(is)
4. Ser exposto de forma repetida ou extrema a detalhes aversivos do evento traumático (p.ex. socorristas coletando restos mortais; policiais repetidamente expostos a detalhes de abuso infantil)

Nota: O critério A4 não se aplica à exposição por meio de mídia eletrônica, televisão, filmes ou fotografias, a menos que tal exposição esteja relacionada ao trabalho.

[LIFE EVENTS CHECKLIST 5]
Abaixo estão listados uma série de acontecimentos difíceis ou estressantes que algumas vezes acontecem com as pessoas. Para cada evento marque uma ou mais alternativas nos quadros à direita para indicar se: (a) aconteceu com você pessoalmente; (b) testemunhou isso acontecer com outra pessoa; (c) ficou sabendo / acontecendo com familiares ou amigos próximos; (d) foi exposto através de seu trabalho (por exemplo: paramédico, policial, militar ou outro tipo de socorrista); (e) não está certo se isto se encaixa a você ou (f) não se aplica a você.

Certifique-se de considerar toda sua vida (infância até a vida adulta) se você passou por alguns dos eventos citados na lista.

1. Desastre natural (ex: enchente, furacão, tornado, terremoto)
- ◯ Aconteceu comigo
- ◯ Testemunhei
- ◯ Fiquei sabendo
- ◯ Parte do meu trabalho
- ◯ Não tenho certeza
- ◯ Não se aplica

2. Incêndio ou explosões
- ◯ Aconteceu comigo
- ◯ Testemunhei
- ◯ Fiquei sabendo
- ◯ Parte do meu trabalho
- ◯ Não tenho certeza
- ◯ Não se aplica

3. Acidente com meio de transporte (ex: acidente de carro, barco, trem, avião)	◯ Aconteceu comigo ◯ Testemunhei ◯ Fiquei sabendo ◯ Parte do meu trabalho ◯ Não tenho certeza ◯ Não se aplica
4. Grave acidente de trabalho, em casa ou durante atividade recreativa.	◯ Aconteceu comigo ◯ Testemunhei ◯ Fiquei sabendo ◯ Parte do meu trabalho ◯ Não tenho certeza ◯ Não se aplica
5. Exposição à substância tóxica (p.ex. produtos químicos perigosos, radiação)	◯ Aconteceu comigo ◯ Testemunhei ◯ Fiquei sabendo ◯ Parte do meu trabalho ◯ Não tenho certeza ◯ Não se aplica
6. Agressão física (p.ex. ser atacada, apanhar, levar socos e chutes, ser espancada)	◯ Aconteceu comigo ◯ Testemunhei ◯ Fiquei sabendo ◯ Parte do meu trabalho ◯ Não tenho certeza ◯ Não se aplica
7. Assalto com arma (p.ex. ser baleada, esfaqueada, ameaçada com arma, faca, bomba)	◯ Aconteceu comigo ◯ Testemunhei ◯ Fiquei sabendo ◯ Parte do meu trabalho ◯ Não tenho certeza ◯ Não se aplica
8. Agressão sexual (estupro, tentativa de estupro, qualquer tipo de ato sexual feito através da força ou ameaça de dano)	◯ Aconteceu comigo ◯ Testemunhei ◯ Fiquei sabendo ◯ Parte do meu trabalho ◯ Não tenho certeza ◯ Não se aplica
9. Outra experiência sexual indesejada ou desconfortável	◯ Aconteceu comigo ◯ Testemunhei ◯ Fiquei sabendo ◯ Parte do meu trabalho ◯ Não tenho certeza ◯ Não se aplica
10. Combate ou exposição a zona de guerra (como militar ou civil)	◯ Aconteceu comigo ◯ Testemunhei ◯ Fiquei sabendo ◯ Parte do meu trabalho ◯ Não tenho certeza ◯ Não se aplica

11. Cativeiro (p.ex. ser sequestrada, mantida como refém, prisioneira de guerra)	◯ Aconteceu comigo ◯ Testemunhei ◯ Fiquei sabendo ◯ Parte do meu trabalho ◯ Não tenho certeza ◯ Não se aplica
12. Ameaça à vida através de doença ou lesão	◯ Aconteceu comigo ◯ Testemunhei ◯ Fiquei sabendo ◯ Parte do meu trabalho ◯ Não tenho certeza ◯ Não se aplica
13. Grave sofrimento humano	◯ Aconteceu comigo ◯ Testemunhei ◯ Fiquei sabendo ◯ Parte do meu trabalho ◯ Não tenho certeza ◯ Não se aplica
14. Morte violenta repentina (p.ex. homicídio, suicídio)	◯ Aconteceu comigo ◯ Testemunhei ◯ Fiquei sabendo ◯ Parte do meu trabalho ◯ Não tenho certeza ◯ Não se aplica
15. Morte súbita acidental	◯ Aconteceu comigo ◯ Testemunhei ◯ Fiquei sabendo ◯ Parte do meu trabalho ◯ Não tenho certeza ◯ Não se aplica
16. Ferimentos graves, danos ou morte que você causou a outra pessoa	◯ Aconteceu comigo ◯ Testemunhei ◯ Fiquei sabendo ◯ Parte do meu trabalho ◯ Não tenho certeza ◯ Não se aplica
17. Qualquer outro evento ou experiência muito estressante	◯ Aconteceu comigo ◯ Testemunhei ◯ Fiquei sabendo ◯ Parte do meu trabalho ◯ Não tenho certeza ◯ Não se aplica
Se outro evento, qual?	_____

Vou lhe perguntar sobre o questionário de experiências estressantes que você preencheu. Primeiro vou pedir para me contar um pouco sobre o evento que você disse que foi o pior para você. Então vou perguntar como esse evento pode ter afetado você durante o mês passado. Em geral não vou precisar de muitas informações - apenas o suficiente para entender qualquer problema que você possa ter tido. Por favor, me avise se você sentir que está se aborrecendo conforme forem feitas as perguntas, para que possamos ir com calma e falar a respeito delas. Também me avise se você tiver alguma dúvida ou não entender algo. Você tem alguma pergunta antes de iniciarmos?

O evento que você disse ter sido o pior foi (EVENTO). Eu gostaria que você descrevesse brevemente o que aconteceu.

O que aconteceu? (Quantos anos você tinha? Como você se envolveu? Quem mais estava envolvido? Alguém foi seriamente ferido ou morto? A vida de alguém estava em perigo? Quantas vezes isso aconteceu?)

Tipo de exposição:	◯ Experienciou ◯ Testemunhou ◯ Ficou sabendo ◯ Exposto a detalhes aversivos
Ameaça à vida?	◯ Não ◯ Sim (de si) ◯ Sim (do outro)
Lesão grave?	◯ Não ◯ Sim (de si) ◯ Sim (do outro)
Violência sexual?	◯ Não ◯ Sim (de si) ◯ Sim (do outro)
Preenche o Critério A?	◯ Não ◯ Provável ◯ Sim

(SE PREENCHER O CRITÉRIO A, SEGUIR PARA INSTRUMENTO CAPS5)

ANNEX 3: FREE INFORMED CONSENT TERM

(TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO)

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Nº do projeto CAAE: 58511716.5.0000.5327

Título do Projeto: PROGRAMA DE DESENVOLVIMENTO DO NÚCLEO INTEGRADO DE ESTUDOS E TRATAMENTO DE TRAUMA PSÍQUICO (NET-TRAUMA-HCPA)

Você está sendo convidado a participar de uma pesquisa cujo objetivo é avaliar como as pessoas reagem quando são vítimas de uma situação traumática como aconteceu com você. Esta pesquisa está sendo realizada pelo Serviço de Psiquiatria (Ambulatório do Núcleo Integrado de Estudos e Tratamento de Trauma Psíquico) do Hospital de Clínicas de Porto Alegre (HCPA).

Se você aceitar participar da pesquisa, os procedimentos envolvidos em sua participação são os seguintes: a partir das informações fornecidas em uma entrevista completa a ser realizada na primeira consulta, um formulário de pesquisa estruturado será preenchido; podendo ser complementado após com dados de consulta ao prontuário eletrônico. Poderá também ser proposta a gravação em meio digital do áudio da consulta, de acordo com sua concordância, cujas informações também complementarão a pesquisa.

Os possíveis desconfortos decorrentes da participação na pesquisa são o tempo de resposta às perguntas (correspondentes ao tempo da consulta a ser realizada) e a possibilidade de constrangimento ao responder às perguntas sobre sua vida pessoal, evento traumático e dados sobre sua saúde.

A participação na pesquisa não trará benefícios diretos aos participantes; porém, contribuirá para o aumento do conhecimento sobre o assunto estudado, podendo beneficiar futuros pacientes.

Sua participação na pesquisa é totalmente voluntária, ou seja, não é obrigatória. Caso você decida não participar, ou ainda, desistir de participar e retirar seu consentimento, não haverá nenhum prejuízo ao atendimento que você recebe ou possa vir a receber na instituição.

Não está previsto nenhum tipo de pagamento pela sua participação na pesquisa e você não terá nenhum custo com respeito aos procedimentos envolvidos.

Caso ocorra alguma intercorrência ou dano, resultante de sua participação na pesquisa, você receberá todo o atendimento necessário, sem nenhum custo pessoal.

Os dados coletados durante a pesquisa serão sempre tratados confidencialmente. Os resultados serão apresentados de forma conjunta, sem a identificação dos participantes, ou seja, o seu nome não aparecerá na publicação dos resultados.

Caso você tenha dúvidas, poderá entrar em contato com o pesquisador responsável Profa Dra Lucia Helena Freitas / Dra Stefania Teche, pelo telefone (51) 33598294, ou com o Comitê de Ética em Pesquisa do Hospital de Clínicas de Porto Alegre (HCPA), pelo telefone (51) 33597640, ou no 2º andar do HCPA, sala 2227, de segunda à sexta, das 8h às 17h.

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Esse Termo é assinado em duas vias, sendo uma para o participante e outra para os pesquisadores.

_____
Nome do participante da pesquisa

_____
Assinatura

_____
Nome do pesquisador que aplicou o Termo

_____
Assinatura

Local e Data: _____

ANNEX 4: ADDITIONAL ARTICLE 1

(LETTER TO THE EDITOR) THE CLINICIAN-ADMINISTERED PTSD SCALE (CAPS-5): ADAPTATION TO BRAZILIAN PORTUGUESE

Thauana T Oliveira-Watanabe [1],
Luis F Ramos-Lima [2],
Roberta C Santos [1],
Marcelo F Mello [1],
Andrea F Mello [1]

[1] Departamento de Psiquiatria, Universidade Federal de São Paulo (UNIFESP), São Paulo, SP, Brazil.
[2] Departamento de Psiquiatria, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil.

Braz J Psychiatry. 2019;41(1):92-93. doi:10.1590/1516-4446-2018-0136

Establishing the diagnosis of posttraumatic stress disorder (PTSD) has always been a challenge in clinical practice, as well as in academic research. Since this diagnosis was first published in DSM-III,[1] several of its criteria have been modified and updated, reflecting current understanding of the disorder.

PTSD is currently considered a debilitating condition that develops from exposure to traumatic events such as actual or threatened death, actual or threatened serious injury, or actual or threatened sexual violence. One can develop PTSD symptoms by direct exposure (e.g., witnessing a traumatic event; learning that a relative or close friend was exposed to trauma) or by indirect exposure to aversive details of the event, usually in the course of professional duties. The DSM-5 lists 20 diagnostic criteria[2] divided into four symptom clusters: re-experience of the traumatic event; avoidance; persistent negative thoughts or feelings; and trauma-related arousal and reactivity.

The Clinician-Administered PTSD Scale (CAPS) is the non-self-administered scale most widely used for PTSD assessment in clinical and research scenarios. It assesses diagnostic status and symptom severity, and was developed in 1989 at the U.S. Department of Veterans Affairs National Center for PTSD.[3] To reflect recent changes in the definition and diagnostic criteria of PTSD, the CAPS has been adapted to the DSM-5 criteria,[4] and has demonstrated good psychometric properties when compared to its previous version. Even though the CAPS-5 is available in English, there is still no DSM-5-based, clinician-administered structured interview in the Brazilian Portuguese language to measure presence and severity of PTSD symptoms. In this letter, we describe the process of cross-cultural adaptation of the CAPS-5 for use in Brazil.

For the cross-cultural adaptation process, we used a formal, structured methodology to ensure conceptual, semantic, and operational equivalence.[5] The original scale was translated into Brazilian Portuguese by two native Brazilian translators, experts in English, and both first versions merged by one of the authors of this study (RCS, bilingual and qualified in use of the previous version). Back-translation was performed by a native English speaker who is fluent in Portuguese and has extensive experience with psychological instruments. Then, an expert team evaluated the equivalence of the instrument to review cultural differences. A pilot study of this version of the instrument was conducted with five individuals who sought treatment at PROVE, a specialized outpatient PTSD clinic of the Universidade Federal de São Paulo (UNIFESP) Department of Psychiatry. The operational equivalence process was conducted by the expert team to analyze some discrepancies found when the target population completed the instrument, and a final version was proposed.

It is our opinion that incorporation of the CAPS-5 as a diagnostic instrument in the context of Brazilian violence is critical. A reliability study to assess the internal consistency of the final version of this instrument, after the cross-cultural adaptation process, is already ongoing. An important step to follow is validation of the translated version, which will allow it to be widely used in Brazil.

Disclosure: The authors report no conflicts of interest.

Acknowledgements: This study received financial support from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP; grant 2014/12559-5) and from Conselho Nacional de Desenvolvimento Cientifico e Tecnológico (CNPq; grant 303389/2016-8).

References

1. American Psychiatric Association . Diagnostic and Statistical Manual of Mental Disorders, Third Edition (DSM-III) Washington: American Psychiatric Publishing; 1980.
2. American Psychiatric Association . Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) Arlington: American Psychiatric Publishing; 2013.
3. Blake DD, Weathers FW, Nagy LM, Kaloupek DG, Gusman FD, Charney DS, et al. The development of a clinician-administered PTSD scale. J Trauma Stress. 1995;8:75–90.
4. Weathers FW, Bovin MJ, Lee DJ, Sloan DM, Schnurr PP, Kaloupek DG, et al. The clinician-administered PTSD scale for DSM-5 (CAPS-5): development and initial psychometric evaluation in military veterans. Psychol Assess. 2018;30:383–95.
5. Reichenheim ME, Moraes CL. [Operationalizing the cross-cultural adaptation of epidemiological measurement instruments] Rev Saude Publica. 2007;41:655–73.

ANNEX 5: ADDITIONAL ARTICLE 2

VALIDATION OF THE BRAZILIAN-PORTUGUESE VERSION OF THE CLINICIAN-ADMINISTERED POSTTRAUMATIC STRESS DISORDER SCALE-5 (CAPS-5)

Thauana T. Oliveira-Watanabe [1],
Luis Francisco Ramos-Lima [2],
Cecilia Zylberstajn [1],
Vinicius Calsavara [1],
Bruno M. Coimbra [1],
Mariana R. Maciel [1],
Lucia Helena M. Freitas [2],
Marcelo F. Mello [1],
Andrea F. Mello [1]

[1] Federal University of São Paulo, Brazil,
[2] Federal University of Rio Grande do Sul, Brazil

(IN SUBMISSION REVIEW)

Abstract

Objectives: The aim of this study was to validate CAPS-5 for the Brazilian-Portuguese language on a sample of 128 individuals from two centers (from the cities of São Paulo and Porto Alegre) who have been recently exposed to a traumatic event.

Methods: We performed a reliability analysis between interviewers (with a subset of 32 individuals), an internal consistency analysis, and a confirmatory factorial analysis for the validation study.

Results: The inter-rater reliability of the total PTSD symptom severity score was high (intraclass correlation coefficient=.994, 95% CI [.987–.997], p<.001). Cohen's Kappa for individual items ranged between .759 and 1. Cronbach's alpha coefficients indicated high internal consistency for the CAPS-5 full scale (α=.826) and an acceptable level of internal consistency for the four symptom clusters. The confirmatory factorial analysis for the 20-item original CAPS-5 did not fit the data well. A 15-item model with better results was then established by excluding the following CAPS-5 items: dissociative amnesia, recklessness, distorted cognitions, irritability, and hypervigilance.

Conclusion: The model with only 15 items provided a good fit to the data with high internal consistency (α=.835).

Contribution to the field

Post-traumatic stress disorder (PTSD) is a highly prevalent disorder around the world, in Brazil between 8.7%, in Rio de Janeiro, and 11.7%, in São Paulo, therefore bringing a big burden to its public health system. It is important to have instruments to properly diagnosis and evaluate people with this condition. The gold standard instrument for this purpose is the clinician administered PTSD scale (CAPS-5) that has not yet been validated to Brazilian-Portuguese. This study intends to validate the CAPS-5 for the Brazilian-Portuguese language. We believe that our study makes a significant contribution to the literature because establishing a version of a Brazilian-Portuguese recognized instrument to evaluate PTSD symptomatology is extremely important for researchers to better understand trauma, mainly related to urban violence in the socio-cultural context of Brazil.

Funding statement

All sources of funding was received from Fundação de Apoio à Pesquisa do Estado de São Paulo (FAPESP) Thematic Project number 2014/12559-5. Additional grant support came

from CNPq 303389/2016-8. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001

Author contribution statement

TT Oliveira-Watanabe was the first author of this paper and LF Ramos-Lima shared the first author contributions for this manuscript. C Zylberstajn and BM Coimbra were involved with the data collection. V Calsavara had the data analyzed. MR Maciel reviewed the article and helped with the discussion and the data collection. LH Freitas, MF Mello and AF Mello guided all the research process and reviewed the paper.

INTRODUCTION

Post-traumatic stress disorder (PTSD) is a severe psychiatric condition developed after exposure to a traumatic event (1). Since 1980, when PTSD was first included in the third version of the Diagnostic and Statistical Manual of Mental Disorders (DSM-III), the definition has been changed and updated in the versions that followed (2). A traumatic event is required for the diagnosis of PTSD, and this has been highlighted in DSM-5, as PTSD is no longer classified among Anxiety Disorders, but in a specific category of Trauma and Stressor-Related Disorders (3). Other changes to the classification of PTSD have restricted what qualifies as a traumatic event and have split the symptoms into four clusters: reexperience, avoidance, negative thoughts and cognitions, and hyperarousal.

Due to the burden of traumatic events, the World Health Survey Consortium conducted a study, which demonstrated that 70.4% of the respondents of all countries had experienced at least one traumatic event in their lifetime (4). In Brazil, this number is even higher: approximately 80% of the Brazilian population has experienced a traumatic event, especially related to urban violence (5). This estimate raises great concern to the Brazilian public health system; an epidemiological study demonstrated that 10.2% of the trauma-exposed population in São Paulo and 8.7% in Rio de Janeiro develop PTSD (2). PTSD can cause a poorer quality of life, which consequently burdens health and social public services (3).

Due to the high prevalence of PTSD in Brazil, its proper assessment and diagnosis is crucial. Currently, the instrument recognized as the gold standard for evaluating PTSD is the clinician-administered PTSD scale (CAPS-5), a structured diagnostic interview to be applied by clinicians. The original scale (6), developed in English, has been last updated to match the DSM-5 PTSD diagnostic criteria (7). The scale has demonstrated high internal consistency,

inter-rater reliability, and test-retest reliability. CAPS-5 also demonstrated good diagnostic correspondence with CAPS-IV (7) and has already been validated in other languages, such as Dutch and German (8, 9).

Adapting an instrument to a specific language has great significance, not only because of the language itself, but also because of the impact that differences in culture, beliefs, and behaviors have on understanding mental health (10). CAPS-5 was validated in the United States on predominantly male veterans (7), a very specific population that greatly differs from the PTSD population in the public health system of Brazil (11).

Brazilian epidemiological studies have revealed that women tend to be more often diagnosed with PTSD than men (12, 13). Ribeiro et al. (2013) evaluated the conditional risk of developing PTSD following a traumatic event and found a three-fold increased risk of developing the disorder in females compared to that in males (15.9% in females [95% CI 14.2–17.6] vs. 5.1% in males [95% CI 4.0–6.2]). These findings highlight the importance of adapting the original version of CAPS-5 to the Brazilian-Portuguese language and validating the instrument to better conduct research in Brazil.

The aim of this study was to validate CAPS-5 for the Brazilian-Portuguese language. Previously, our research team performed a cross-cultural adaptation process with a formal and structured methodology to ensure conceptual, semantic, and operational equivalence (14). In order to complete the validation process, we performed a reliability analysis between interviewers, evaluated internal consistency, and conducted a confirmatory factorial analysis (CFA). We hypothesized that the instrument would exhibit good inter-rater reliability and internal consistency, based on psychometric measurements obtained in previous studies performed in other countries (7, 8, 9).

METHODS

The CAPS-5 instrument

The scale assesses the diagnostic criteria based on DSM-5 and the intensity of the PTSD symptoms. It has 30 questions, 20 of which correspond to each DSM-5 diagnostic criterion. The first question refers to the existence of a traumatic experience (Criteria A); the original scale recommends another instrument to evaluate the occurrence of traumatic events, usually the Life Events Checklist (LEC-5) (15). The LEC-5 is a self-report measure designed to recognize potentially traumatic events in a respondent's lifetime. We also used LEC-5 to evaluate Criteria A, adapted to Brazilian-Portuguese in a previous study (16).

The 20 symptom-related questions were divided into four groups: intrusion symptoms (five items, Criterion B), avoidance questions (two items, Criterion C), negative alterations in cognition and mood (seven items, Criterion D), and hyperarousal (six items, Criterion E). Concerning other DSM-5 criteria, one question refers to how long the identified traumatic event lasted (Criteria F) and three questions to the impact of the disturbance on functionality (Criteria G). Three final questions regarding the interviewer's impression on the patient are presented, long with two questions regarding the presence of dissociative symptoms.

Cross-cultural adaptation

The cross-cultural adaptation process was performed within the Program for Research and Care on Violence and PTSD (PROVE) at the Department of Psychiatry of The Federal University of São Paulo (UNIFESP) according to the model proposed by Reichenheim and Moraes (2007). This process included the following steps: translation from English to Portuguese, a back-translation to the original language, a revision by an expert team, a pilot study using the adapted version, and an equivalence evaluation with the original version. This adaptation was published previously (14).

Participants

This study was conducted in two centers: at UNIFESP and at the Clinical Hospital of Porto Alegre (HCPA), which belongs to the Federal University of Rio Grande do Sul (UFRGS). The complete sample was composed of three groups of patients.

The first group (Sample 1) consisted of sexually assaulted women included in a randomized clinical trial, which is part of a thematic project sponsored by Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP), conducted at PROVE-UNIFESP. The patients were enrolled between January 2016 and March 2019. The adapted CAPS-5 was applied during the screening process to select the participants for the study, together with other instruments concerning the thematic project. The scale was administrated by two trained professionals (one psychiatrist and one psychologist). All patients experienced the trauma up to 6 months before the assessment.

The second group (Sample 2) was enrolled from the screening process of PROVE outpatient's service. The participants spontaneously sought for psychiatric assistance after experiencing a traumatic event. The screening process was conducted by two professionals (one psychiatrist and one psychologist), who also applied the adapted version of CAPS-5. The patients were enrolled between March 2018 and February 2019.

The third group (Sample 3) was enrolled at the Psychological Trauma Research and Treatment Program (NET-Trauma) outpatient service from HCPA-UFRGS. The screening process was similar to the PROVE outpatient center; patients suffering different types of trauma agreed to participate in the study. They were assessed between August 2018 and February 2019. A summary of the three samples is presented in Table 1.

All the professionals who applied the instrument were trained for CAPS-5 use. In every case, the LEC-5 scale was applied to ensure that the DSM-5 Criteria A for PTSD was fulfilled. This study was conducted with approval from the Ethics Committees of both UNIFESP and HCPA-UFRGS). All participants signed the informed consent form.

Data analysis

Reliability between interviewers. We compared the results of two independent interviewers (one psychiatrist and one psychologist), who had administered CAPS-5 to the same participants. For the reliability evaluation between interviewers, 32 participants were selected from Sample 1 (n=15 participants) and 2 (n=17 participants). We calculated the Cohen's kappa coefficient, considering a confidence interval of 95% and the 20 items of the scale as ordinal variables. The kappa coefficient varies between 1 and -1, which indicates complete agreement or complete disagreement, and a value of 0 indicates a random result (17). This coefficient was applied to all of the 20 items of the scale corresponding to DSM-5 symptoms. We also used the intraclass correlation coefficient to evaluate the total PTSD severity score (18).

Internal consistency. We combined the three samples in order to obtain the minimal number of participants acceptable for a good psychometric analysis; the final sample comprised 128 participants. The Cronbach's alpha coefficient was used to determine internal consistency, which is considered good when the value is > 0.80 and most inter-item correlations are in the recommended range of moderate magnitude (.15–.50) (19).

CFA. We used the final sample (with 128 participants) to elucidate whether the CAPS-5 in the Brazilian context should have the same structure as the original CAPS-5, validated in the American context. The factor structure of the adapted CAPS-5 was examined using CFA. Items were treated as ordinal variables, and parameters were estimated using the maximum-likelihood estimator method, which provides good performance for small samples and a robust chi-square. The model fit was evaluated using chi-square under the degree of freedom

ratio (X2/df): values < 3 are considered acceptable for the model; Goodness of Fit Index (GFI), Tucker-Lewis Index, Comparative Fit Index (CFI): values < .90 indicate a lack of fit, values between .90 and .95 indicate a reasonable fit, and values between .95 and 1.00 indicate a good fit; Root Mean Square Error of Approximation (RMSEA): values ≤ .06 indicate a close fit; and Standardized Root Mean Square Residual (SRMR): values < .08 indicate a well-fitting model (20, 21). For comparative analysis, we performed the chi-square, One-Way ANOVA and Kruskal-Wallis for categorical and continuous variables. Demographic information was missing for up to six patients, depending on the variable, in Sample 2. Otherwise, no other missing data were detected. The significance level of the tests was fixed at .05. All statistical analyses were performed using SPSS version 21 (IBM Corporation, Armonk, NY, USA). The SPSS AMOS was used to perform CFA.

RESULTS

The sample included 128 patients: 97 patients from PROVE (Sample 1 and Sample 2) and 31 patients from HCPA-UFRGS (Sample 3). Most of the patients were female (93.8%), single (63.5%), employed (69.9%), and religious (66.9%). CAPS-5 total severity score and total amount of symptoms did not differ among samples.

Table 1. Sociodemographic data from three samples (n=128).

Variables	Sample 1 (n=80)	Sample 2 (n=17)	Sample 3 (n=31)	All	*P* value
Age (mean, sd)	25.5 (6.7)	37.8 (11.1)	35.9 (16.2)	29.7 (11.6)	[a] <.001
Female gender	80 (100%)	14 (82.4%)	26 (83.9%)	120 (93.8%)	[a].001
Ethnicity					
Caucasian	32 (40.0%)	7 (43.8%)	19 (61.3%)	58 (45.7%)	[a] <.001
African-American	46 (57.5%)	8 (50.0%)	5 (16.1%)	59 (46.5%)	
Asiatic-American	2 (2.5%)	1 (6.3%)	7 (22.6%)	10 (7.9%)	
Marital status					
Single	55 (68.8%)	9 (60.0%)	16 (51.6%)	80 (63.5%)	[a].006
Married/engaged	23 (28.7%)	4 (26.7%)	7 (22.6%)	34 (27.0%)	
Divorced/widower	2 (2.5%)	2 (13.3%)	8 (25.8%)	12 (9.5%)	
Educational level					
< 4 years	0 (0%)	0 (0%)	2 (6.5%)	2 (1.6%%)	[a].001
4−12 years	39 (48.8%)	2 (16.7%)	22 (71.0%)	63 (51.2%)	
> 12 years	41 (51.2%)	10 (83.3%)	7 (22.6%)	58 (47.2%)	
Employment status					

Employed	63 (78.8%)	8 (66.7%)	15 (48.4%)	86 (69.9%)	[a].011
Unemployed/retired	13 (16.3%)	4 (33.3%)	10 (32.3%)	27 (22.0%)	
Health licensed	4 (5.0%)	0 (0%)	6 (19.4%)	10 (8.1%)	
Religious	55 (68.8%)	7 (53.8%)	21 (67.7%)	83 (66.9%)	.567

Inter-rater Reliability

To estimate the inter-rater reliability, we considered the 32 individuals from Samples 1 and 2 who had CAPS-5 performed by two interviewers. The inter-rater reliability of the total PTSD symptom severity score was high (intraclass correlation coefficient=.994, 95% CI [.987–.997], p<.001). Cohen's kappa for each item was evaluated to determine if there was agreement between the two raters. Among the 20 items from the CAPS-5 scale, we found total agreement in four (B5/Physiological distress; C1/Memory avoidance; E1/Irritability; E2/Recklessness). We found Kappa between 0.759-0.8 in five items (D2/Distressing dreams, k=.792; C2/External avoidance, k=.795; D2/Negative beliefs, k=.768; D4/Negative states, k=.759; E5/Concentration problems, k=.770). The remaining 15 symptoms resulted in a kappa value > .8, indicating an "almost perfect" agreement between raters.

Table 2: Inter-rater reliability coefficients of CAPS-5 (n=36).

Factor	**Item**	**Kappa**
B Intrusion	B1 Recurrent memories	.880 *
	B2 Distressing dreams	.792*
	B3 Dissociative reactions	.953*
	B4 Psychological distress	.897*
	B5 Physiological reactions	1*
C Avoidance	C1 Memory avoidance	1*
	C2 External avoidance	.795*
D Negative cognitions	D1 Dissociative amnesia	.842*
	D2 Negative beliefs	.768*
	D3 Distorted cognitions	.960*
	D4 Negative states	.759*
	D5 Diminished interest	.813*
	D6 Feelings of detachment	.952*
	D7 Reduction of positive emotions	.956*
E Hyperarousal	E1 Irritability	1*
	E2 Recklessness	1*
	E3 Hypervigilance	.857*
	E4 Startle response	.871*
	E5 Concentration problems	.770*
	E6 Sleep disturbance	.823*

*p<.001

Internal Consistency

Cronbach's alpha coefficients indicated 172 high internal consistency for the CAPS-5 full scale (α=.826) and an acceptable level of internal consistency (22) for the four symptom clusters: B/Intrusion (α=.631), C/Avoidance (α=.404), D/Negative cognitions (α=.701), and E/Hyperarousal (α=.537). Two symptoms, D1/Dissociative amnesia and E2/Recklessness, had low item-total correlations (.025 and .095, respectively). The range of item-total correlations for the remaining 18 symptoms was .317−.613, with a mean of .438. By excluding these two items, the Cronbach's alpha coefficient for the full scale increased to .842.

Most inter-item correlations were in the recommended range of .15−.50 (19), with a mean of .193 across all 20 symptoms. The symptoms D1/Dissociative amnesia and E2/Recklessness exhibited low inter-item correlations, with values of .152 and .189, respectively. The mean inter-item correlation for the remaining 18 symptoms was .233.

CFA

CFA with the maximum-likelihood estimation method was conducted to determine whether the factor structure indicated by the original scale could be confirmed. We performed CFA for the 20-item original CAPS-5 scale and for the 18-item model with the exclusion of two items (D1 and E2), suggested by the internal consistency analysis. The fit indices for the 18-item model were X2/gl=1.441, GFI=.861, CFI=.878, TFI=.855, RMSEA=.059, and SRMR=.076, supporting a reasonable fit to the data. We concluded that the 20-item and 18-item CAPS-5 models did not fit the data adequately well. In order to improve the performance of the instrument, we analyzed all factor loads from each item from the 18-item model. We found three items with low factor loads: D3/Distorted cognitions (.388), E1/Irritability (.305), and E3/Hypervigilance (.400). All other items had factor loads > .40. We then performed a third CFA of the 15 remaining items. The 15-item model exhibited a good fit to the data (X2/gl=1.248, GFI=.910, 198 CFI=.948, TFI=.951, RMSEA=.044, and SRMR=.063).

Table 3: CFA fit statistics for 20-item, 18-item, and 15-item CAPS-5 models.

Fit index	Level of good fit[1]	20-item	18-item	15-item
X2/df	< 3	1.350	1.441	1.248
GFI	> 0.9	.854	.861	.910
CFI	> 0.9	.878	.878	.948
TFI	> 0.9	.858	.855	.951
RMSEA [90% CI]	< 0.06	.052 [.033−.069]	.059 [.039−.077]	.044 [.000−.069]
SRMR	< 0.08	.076	.076	.063

[1] (Brown, 2015; Hu & Bentler, 1999)

Observing the final 15-item model, all items were found to have significant loadings onto their respective latent constructs. The standardized regression weights (factor loadings) for all items were > .3, corresponding to good-magnitude loadings (20). The occurrence of factors with item reduction (two items in factor D/Negative cognitions and three items in factor E/Hyperarousal) may explain the relatively poor loadings. The item D5/Diminished interest exhibited a high factor load: 81% of the factor variance was accounted by this item, suggesting that D5/Diminished interest is a strong indicator of negative cognition. All other factor loadings ranged between .40 and .65.

Table 4: Parameter estimates for the 15-item CAPS-5 model.

Factor	**Item**	**Estimate**	**SE**	**FL**
B Intrusion	B1	1.00		.45
	B2	1.88	.47*	.52
	B3	1.99	.51*	.50
	B4	1.04	.26*	.42
	B5	2.13	.49*	.62
C Avoidance	C1	1.00		.48
	C2	1.41	.36*	.54
D Negative cognitions	D2	1.00		.65
	D4	.52	.12*	.45
	D5	1.26	.18*	.81
	D6	.94	.16*	.60
	D7	1.03	.19*	.58
E Hyperarousal	E4	1.00		.40
	E5	1.01	.25*	.41
	E6	1.12	.26*	.47

* p<.001.

Estimate: unstandardized regression weights; SE: standard error of the unstandardized regression weights; FL: factor loadings (standardized regression weights).

We calculated the new Cronbach's alpha coefficients for the 15-item model and found improvement compared with the full scale. The 15-item model exhibited high internal consistency (α=.835), and the two factors with item reduction maintained acceptable levels of internal consistency: D/Negative cognitions (α=.747) and E/Hyperarousal (α=.403).

DISCUSSION

The present study describes the development of the Brazilian-Portuguese version of the CAPS-5 scale, a unique instrument for clinicians to evaluate PTSD in a structured manner in Brazil. Our research team has already translated the original English scale using a structured method published elsewhere (14). The present study now demonstrates a high inter-rater reliability for all 20 items and the total severity score. It is important to emphasize that the raters were experienced professionals in attending PTSD patients. Further studies with less experienced health-care professionals are necessary to determine the consistency of our results.

The present study has demonstrated an adequate internal consistency for the four clusters of symptoms and high internal consistency for the full scale. These findings are in line with previous scale validations for other languages (8, 9) as well as the validation of the original CAPS-5 version (7). Lower Cronbach's alpha coefficients for the cluster C/Avoidance have been reported before and can be explained by the existence of only two items in this cluster; similar results were found in the original English version and in the Dutch translated version (7, 9). On observing the items separately, we found that two items had a low item-total correlation (D1/Dissociative amnesia and E2/Recklessness), consistent with the findings of the original English version. According to the original CAPS-5 validation process (7), these findings may be attributable to a very infrequent endorsement of these two symptoms, corroborating the hypothesis that these items may be important but relatively rare symptoms of PTSD, or it may be that they are simply not representative of the PTSD construct.

A CFA was conducted to verify adequate fit to the data. Due to the similar validation process of CAPS-5 in previous studies and the existence of consolidated constructs (clusters) that explain PTSD, we decided to perform only a CFA instead of evaluating constructs from the entire scale (with exploratory factor analysis). Previous results from internal consistency analysis were used to suggest items for exclusion; the 15-item model provided the best fit to the data, compared to the full scale and 18-item model. We postulated that a 15-item model for CAPS-5 could maintain adequate results without compromising diagnostic capacity.

The use of heterogeneous data from different sources is a strength of the present study and is in contrast to the validation of the original scale, which was based on symptoms observed in only war veterans. It would be of interest for future studies to perform a more complete evaluation of the scale construct. A challenge observed in this study was the maintenance of the original structure and constructs/factors in the CFA: the occurrence of

factors composed of few items can explain the relatively low factor loads. Other strong points are the well-trained professionals that were able to diagnose PTSD properly, only one research team undertaking the entire adaptation process of the scale, and strong inclusion criteria for the participants. Our study presents an important aspect that must be considered: although we were able to evaluate individuals from different institutions and sources, a significant number of participants were women who experienced a traumatic event related to sexual violence. Indeed, the high prevalence of sexual assault victims in Brazil is a prevailing reason that individuals seek treatment in outpatient services that deal with traumatic events. Future studies should assess the consistency of results by comparing the 15-item scale with the 20-item scale for different types of traumatic events.

CONCLUSION

In developing countries, PTSD is mostly related to urban violence that has a high frequency of traumatic events, such as robbery, kidnapping, sexual assault, rape, witnessing shootings, and other life-threatening situations. This is frequently related to complex PTSD diagnoses; thus, we must ensure that the CAPS-5 is an efficient instrument to detect this reality. Establishing a validated version of a Brazilian- Portuguese diagnostic instrument to evaluate PTSD symptomatology is extremely important for researchers to better understand these issues in this socio-cultural context.

REFERENCES

1. American Psychiatric Association. DSM-5 Task Force. (2013). Diagnostic and statistical manual of mental disorders : DSM-5. American Psychiatric Pub.

2. Ribeiro, W. S., Mari, J. de J., Quintana, M. I., Dewey, M. E., Evans-Lacko, S., Vilete, L. M. P., … Andreoli, S. B. (2013). The Impact of Epidemic Violence on the Prevalence of Psychiatric Disorders in Sao Paulo and Rio de Janeiro, Brazil. PLOS ONE, 8(5), e63545. https://doi.org/10.1371/JOURNAL.PONE.0063545

3. Atwoli, L., Stein, D. J., Koenen, K. C., & McLaughlin, K. A. (2015). Epidemiology of posttraumatic stress disorder: prevalence, correlates and consequences. Current Opinion in Psychiatry, 28(4), 307–311. https://doi.org/10.1097/YCO.0000000000000167

4. Benjet, C., Bromet, E., Karam, E. G., Kessler, R. C., McLaughlin, K. A., Ruscio, A. M., … Koenen, K. C. (2016). The epidemiology of traumatic event exposure worldwide: results from the World Mental Health Survey Consortium. Psychological Medicine, 46(2), 327–343. https://doi.org/10.1017/S0033291715001981

5. Teche, S. P., Barros, A. J. S., Rosa, R. G., Guimarães, L. P., Cordini, K. L., Goi, J. D., … Freitas, L. H. (2017). Association between resilience and posttraumatic stress disorder among Brazilian victims of urban violence: a cross-sectional case-control study. Trends in Psychiatry and Psychotherapy, 39(2), 116–123. https://doi.org/10.1590/2237-6089-2016-0070

6. Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Gusman, F. D., Charney, D. S., & Keane, T. M. (1995). The development of a Clinician-Administered PTSD Scale. Journal of Traumatic Stress, 8(1), 75–90. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7712061

7. Weathers, F. W., Bovin, M. J., Lee, D. J., Sloan, D. M., Schnurr, P. P., Kaloupek, D. G., … Marx, B. P. (2018). The Clinician-Administered PTSD Scale for DSM–5 (CAPS-5): Development and initial psychometric evaluation in military veterans. Psychological Assessment, 30(3), 383–395. https://doi.org/10.1037/pas0000486

8. Boeschoten, M. A., Van der Aa, N., Bakker, A., Ter Heide, F. J. J., Hoofwijk, M. C., Jongedijk, R. A., … Olff, M. (2018). Development and Evaluation of the Dutch Clinician-Administered PTSD Scale for DSM-5 (CAPS-5). European Journal of Psychotraumatology, 9(1), 1546085. https://doi.org/10.1080/20008198.2018.1546085

9. Müller-Engelmann, M., Schnyder, U., Dittmann, C., Priebe, K., Bohus, M., Thome, J., … Steil, R. (2018). Psychometric Properties and Factor Structure of the German Version of the Clinician-Administered PTSD Scale for DSM-5. Assessment, 1073191118774840. https://doi.org/10.1177/1073191118774840

10. Reichenheim, M. E., & Moraes, C. L. (2007, August). Operacionalização de adaptação transcultural de instrumentos de aferição usados em epidemiologia. Revista de Saude Publica. https://doi.org/10.1590/s0034-89102006005000035

11. Ramos-Lima, L. F., Souza, P. R. A. de, Teche, S. P., & Freitas, L. H. M. (2019). Trauma-related disorders in a low- to middle-income country: A four-year follow-up of outpatient trauma in Brazil. Psychiatry Research, 280. https://doi.org/10.1016/j.psychres.2019.112525

12. Andrade, L. H., Wang, Y.-P., Andreoni, S., Silveira, C. M., Alexandrino-Silva, C., Siu, E. R., … Viana, M. C. (2012). Mental Disorders in Megacities: Findings from the São Paulo Megacity Mental Health Survey, Brazil. PLoS ONE, 7(2), e31879. https://doi.org/10.1371/journal.pone.0031879

13. Seedat, S., Scott, K. M., Angermeyer, M. C., Berglund, P., Bromet, E. J., Brugha, T. S., … Kessler, R. C. (2009). Cross-national associations between gender and mental disorders in the World Health Organization World Mental Health Surveys. Archives of General Psychiatry, 66(7), 785–795. https://doi.org/10.1001/archgenpsychiatry.2009.36

14. Oliveira-Watanabe, T. T., Ramos-Lima, L. F., Santos, R. C., Mello, M. F., & Mello, A. F. (2019). The Clinician-Administered PTSD Scale (CAPS-5): adaptation to Brazilian Portuguese. Revista Brasileira de Psiquiatria (Sao Paulo, Brazil : 1999), 41(1), 92–93. https://doi.org/10.1590/1516-4446-2018- 0136

15. Gray, M. J., Litz, B. T., Hsu, J. L., & Lombardo, T. W. (2004). Psychometric Properties of the Life Events Checklist. Assessment, 11(4), 330–341. https://doi.org/10.1177/1073191104269954

16. Lima, E. de P., Vasconcelos, A. G., Berger, W., Kristensen, C. H., Nascimento, E. do, Figueira, I., & Mendlowicz, M. V. (2016). Cross-cultural adaptation of the Posttraumatic Stress Disorder Checklist 5 (PCL-5) and Life Events Checklist 5 (LEC-5) for the Brazilian context. Trends in Psychiatry and Psychotherapy, 38(4), 207–215. https://doi.org/10.1590/2237-6089-2015-0074

17. Andreoli, S., Blay, S. L., de Almeida Filho, N., de Jesus Mari, J., Torres de Miranda, C. T., da Silva Freire Coutinho, E., … D'Arrigo Busnello, E. (2001). Reliability of diagnostic instruments: investigating the psychiatric DSM-III checklist applied to community samples. Cadernos de Saúde Pública /Ministério Da Saúde, Fundação Oswaldo Cruz, Escola Nacional de Saúde Pública, 17(6), 1393–1402. https://doi.org/10.1590/S0102-311X2001000600010

18. Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin, 86(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

19. Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. Psychological Assessment, 7(3), 309–319. https://doi.org/10.1037/1040-3590.7.3.309

20. Brown, T. (2015). Confirmatory factor analysis for applied research. Guilford publications.

21. Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling, 6(1), 1–55. https://doi.org/10.1080/10705519909540118

22. Gorenstein, C., Wang, Y. P., Hungerbühler, I. (2016) Fundamentos de mensuração em saúde mental; Instrumentos de avaliação em saúde mental. Artmed Ed. Porto Alegre.