

Original Article

Cite this article: Agne NA, Tisott CG, Ballester P, Passos IC, Ferrão YA (2022). Predictors of suicide attempt in patients with obsessive-compulsive disorder: an exploratory study with machine learning analysis. *Psychological Medicine* **52**, 715–725. <https://doi.org/10.1017/S0033291720002329>

Received: 28 December 2019

Revised: 11 May 2020

Accepted: 10 June 2020

First published online: 16 July 2020


Key words:

Algorithm; elastic net; machine learning; obsessive-compulsive disorder; predictors of suicide; suicide attempts

Author for correspondence:

Neusa Aita Agne, E-mail: neusaagne@gmail.com

Predictors of suicide attempt in patients with obsessive-compulsive disorder: an exploratory study with machine learning analysis

Neusa Aita Agne¹ , Caroline Gewehr Tisott¹, Pedro Ballester², Ives Cavalcante Passos^{3,4} and Ygor Arzeno Ferrão^{1,5}

¹Federal University of Health Sciences of Porto Alegre (UFCSA), Porto Alegre (RS), Brazil; ²Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, ON, Canada; ³Laboratory of Molecular Psychiatry, Centro de Pesquisa Experimental (CPE) and Centro de Pesquisa Clínica (CPC), Hospital de Clínicas de Porto Alegre (HCPA), Instituto Nacional de Ciência e Tecnologia Translacional em Medicina (INCT-TM), Porto Alegre (RS), Brazil; ⁴Department of Psychiatry, Universidade Federal do Rio Grande do Sul, School of Medicine, Graduate Program in Psychiatry and Behavioral Sciences, Porto Alegre, Brazil and ⁵Brazilian Research Consortium on Obsessive-Compulsive Spectrum Disorders (C-TOC), Porto Alegre, Brazil

Abstract

Background. Patients with obsessive-compulsive disorder (OCD) are at increased risk for suicide attempt (SA) compared to the general population. However, the significant risk factors for SA in this population remains unclear – whether these factors are associated with the disorder itself or related to extrinsic factors, such as comorbidities and sociodemographic variables. This study aimed to identify predictors of SA in OCD patients using a machine learning algorithm.

Methods. A total of 959 outpatients with OCD were included. An elastic net model was performed to recognize the predictors of SA among OCD patients, using clinical and socio-demographic variables.

Results. The prevalence of SA in our sample was 10.8%. Relevant predictors of SA founded by the elastic net algorithm were the following: previous suicide planning, previous suicide thoughts, lifetime depressive episode, and intermittent explosive disorder. Our elastic net model had a good performance and found an area under the curve of 0.95.

Conclusions. This is the first study to evaluate risk factors for SA among OCD patients using machine learning algorithms. Our results demonstrate an accurate risk algorithm can be created using clinical and sociodemographic variables. All aspects of suicidal phenomena need to be carefully investigated by clinicians in every evaluation of OCD patients. Particular attention should be given to comorbidity with depressive symptoms.

Introduction

Obsessive-compulsive disorder (OCD) is one of the most prevalent neuropsychiatric disorders, affecting about 2.5% of people throughout their lives (Kessler, Chiu, Demler, Merikangas, & Walters, 2005; Ruscio, Stein, Chiu, & Kessler, 2010). According to the World Health Organization (WHO), OCD was ranked one of the most debilitating and disabling diseases (Murray, Lopez, & World Health Organization, World Bank & Harvard School of Public Health, 1996). Nowadays, the WHO lists anxiety disorders, including OCD, as the sixth largest contributor to non-fatal health loss (World Health Organization, 2017). Despite the high levels of suffering, disability, and impairment associated with the disorder, high rates of comorbidities and low rate of response to treatment, suicidal behavior in patients with OCD have received less attention (American Psychiatric Association, 2013; Eisen et al., 2006; Jacoby, Leonard, Riemann, & Abramowitz, 2014).

Suicide is a major public health problem and one of the top 10 causes of death and burden of disease worldwide, across all age groups (Lozano et al., 2012; World Health Organization, 2014). Also, about 90% of suicide deaths are directly related to mental disorders (Arsenault-Lapierre, Kim, & Turecki, 2004; Hawton & van Heeringen, 2009; Nock, Hwang, Sampson, & Kessler, 2010). Suicide in OCD has been considered a rare phenomenon, with suicide death rates described in the classic literature about 1% (Coryell, 1981; Goodwin, Guze, & Robins, 1969), with some studies finding a negative association results between suicide and OCD (Kanwar et al., 2013).

Contrary to these findings, recent evidence has suggested that OCD suicide rates might be underestimated, with two systematic reviews indicating rates of suicide attempts (SA) in OCD ranging from 1% to 46.3% (Angelakis, Gooding, Tarrier, & Panagioti, 2015) and from 6% to 51.7% (Albert, De Ronchi, Maina, & Pompili, 2018). However, the results of previous studies raise some questions: suicide in OCD is connected to intrinsic factors about the psychopathology of OCD (related to the direct effects of the nuclear psychopathological aspects of the

disorder) or is the result of factors extrinsic to OCD psychopathology nucleus, such as the occurrence of psychiatric comorbidities or sociodemographic factors. To the best of our knowledge, this research field to date has mainly used traditional statistical analysis, but the psychopathological heterogeneity of OCD can potentially reduce power and obscure the findings of these clinical studies, as it is known that there are several limitations using traditional hypothesis testing methods to analyze multidimensional and heterogeneous data (Passos, Mwangi, & Kapczinski, 2019b).

In light of these findings, techniques that can analyze data from multiple biological levels at the individual level, such as machine learning, are essential (Passos *et al.*, 2019a). The use of machine learning focuses on algorithms that can analyze, learn, and extract patterns from data in a non-linear and interactive way, thus transforming data into relevant information, exceeding the human brain's ability to understand it (Greenhalgh, Howick, & Maskrey, 2014; Huys, Maia, & Frank, 2016; Passos & Mwangi, 2018; Rajkomar, Dean, & Kohane, 2019).

Machine learning has shown to be more accurate in predicting suicidality than traditional statistical analysis in several studies (de Ávila Berni *et al.*, 2018; Larsen *et al.*, 2015; Leonard Westgate, Shiner, Thompson, & Watts, 2015; Niculescu *et al.*, 2015; O'Dea *et al.*, 2015; Passos *et al.*, 2016; Simon *et al.*, 2018; Torous *et al.*, 2018) producing suicidal behavioral discriminators [areas under the curve (AUCs) = 0.60–0.80] that exceed those using isolated risk factors (AUCs = 0.58) (Franklin *et al.*, 2017). In a recent systematic review, Burke, Ammerman, and Jacobucci (2019) identified 35 articles assessing the use of machine learning techniques for predicting and identifying risk factors for suicide-related outcomes. This study found greater precision in the prediction of suicide outcomes when compared with articles that used traditional statistical methods. In addition, such studies replicated findings of known risk factors, identified new variables involved in the phenomenon of suicide, and allowed the identification of subgroups, helping to draw cut-off points for suicide risk that can be used in clinical practice (Burke *et al.*, 2019).

Currently, this type of analysis has not been used to predict the risk of SA in patients with OCD at the individual level. In this sense, our objective will be to explore sociodemographic and clinical factors associated with SA in adult OCD patients. Considering that the evaluation of a patient with OCD is extremely complex due to its heterogeneity (several presentations and symptom dimensions) (Rosario-Campos *et al.*, 2006), the precise identification of which clinical variables can predict suicidal behavior at the individual level is clinically essential for the development of preventive and therapeutic interventions.

Methods

Participants

The sample was initially formed by 1001 OCD patients from seven university centers in six Brazilian cities, which made up the Brazilian OCD Research Consortium (C-TOC). Patients were interviewed between August 2003 and August 2009 and were included in the study because they met criteria for OCD confirmed by the Structured Clinical Interview for Axis Disorders I of DSM-IV (SCID-I) (First, Spitzer, Gibbon, & Williams, 1997). Patients with psychotic disorders or any other condition that impaired their understanding of the evaluation questions were excluded (e.g. schizophrenia, autism spectrum disorder, intellectual disability), and each local Ethics Committee approved the research protocol.

The detailed description of the evaluation protocol is available elsewhere (Miguel *et al.*, 2008). For this paper, there were 42 (4.2%) patients excluded due to missing data concerning suicide information, resulting in a sample of 959 OCD patients.

Instruments

Subjects were interviewed by psychologists or psychiatrists trained in the application of the protocol and their structured interviews. The C-TOC protocol comprised sociodemographic data (sex, age, religion, marital status, occupation, socioeconomic classification, etc.), personal medical history, family psychiatric history, and other standardized instruments. Among them, the items below were relevant for this study:

The SCID-I and additional modules for tic and impulse control disorders were used to assess psychiatric comorbidities (First *et al.*, 1997). Attention deficit hyperactivity disorder (ADHD) and separation anxiety disorder were investigated through a module of the Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS) (Kaufman *et al.*, 1997).

- The Yale Obsessive-Compulsive Brown Scale (Y-BOCS) was used to measure the severity of OCD symptoms. It has been translated to Portuguese by Asbahr, Lotufo, Turecki, and Miguel (1996).
- The Dimensional Yale-Brown Obsessive-Compulsive Scale (DY-BOCS) (Rosario-Campos *et al.*, 2006) evaluated the presence and severity of obsessive-compulsive symptoms (OCS) according to specific six dimensions, including obsessions and related compulsions, as well as investigated avoidance and the time spent with OCS, the level of anxiety, and symptom burden. Its scores range from 0 to 5 (maximum of 15 for each dimension). The negative impact of OCS is also measured (maximum score of 30), and the therapeutic response can be evaluated according to specific dimensions. This instrument was simultaneously validated in Portuguese and English.
- The Beck Depression Inventory (BDI) (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) and the Beck Anxiety Inventory (BAI) (Beck, Epstein, Brown, & Steer, 1988) for depressive and anxious symptoms assessment, respectively. These scales were translated to Portuguese by Gorenstein and Andrade (1996).
- The Yale OCD Natural History Questionnaire was used to examine OCS onset and course, including stressful life events.

As in previous publications with similar samples (Torres *et al.*, 2011; Velloso *et al.*, 2016), we used a clinical questionnaire about suicide behaviors composed of seven questions with categorical answers ('Yes', 'No', 'I do not know'), and included: 'Have you ever thought that it was not worth living?'; 'Did you ever wish you were dead?'; 'Have you ever thought about taking your own life or committing suicide?'; 'Have you planned to take your own life or commit suicide?'; 'Have you ever tried to take your own life or commit suicide?'; 'Did you need hospitalization/treatment at that time?'; 'Have any of your relatives tried to commit suicide?'; 'Have any of your relatives ever committed suicide?'. The question about attempted suicide (SA) was chosen as the outcome in this study.

Selection of predictor variables

The variables used in the model of this study were selected through a structured search on PubMed. We searched for articles that had clinical relevance in finding risk factors for suicide in OCD. Thus, the predictor variables selected to be used in

'training' the algorithm included demographic and clinical variables related to OCD and other psychopathologies that presented comorbidly in our database. This way of selecting the predictor variables is a method already used in previous studies that used machine learning (Passos et al., 2016; Perlis, 2013). Given the extensive number of variables present in our dataset ($n = 1.680$), we suggest that readers seek Miguel et al. (2008) for a more detailed description of all the data collected. According to the meticulous C-TOC original dataset review [two authors, NAA and YAF, reviewed each variable of the dataset separately and, after a best estimate diagnosis technique (Maziade et al., 1992), a consensus was achieved], 89 remained variables were entered into the analysis. Table S1 in the Supplementary Material resumes these remaining variables.

Statistical analysis

Descriptive analyses were performed first. The Kolmogorov-Smirnov test was used to assess the normality of the data. Continuous variables were described as means and standard deviations (s.d.) and/or median and minimum-maximum values (min-max) according to their distribution (normal or non-normal). Categorical variables were described as absolute values (n) and relative values (%). Statistics Package for Social Sciences 18.0 (SPSS) was used to perform the analysis (SPSS Inc., 2009).

Machine learning analysis was performed with R software (Version R 3.3.1) (R Core Team, 2013) and R Studio (Version 0.99.902) with the R package caret (Version 6.0-73) (Kuhn, 2008). We ran experiments with the elastic net method (Zou & Hastie, 2005). We used machine learning techniques to address two problems in conventional multiple regression: (1) coefficients are unstable when high correlations exist among predictors; this is the case for the predictors included in the present study, leading to low replication of predictions in independent samples (Berk, 2008); (2) traditional regression assumes additivity, whereas the predictors considered here might have non-additive effects (DiGangi et al., 2013; Ozer, Best, Lipsey, & Weiss, 2003; Tolin & Foa, 2006). The elastic net method mitigates the first of these problems by introducing regularization factors while preserving the model interpretability provided by traditional regression. We leave using non-linear models for future work in the field as they present additional challenges for generalization and interpretation.

The elastic net is a machine learning method that comprises both feature selection characteristics with regularization and classification. In statistics and, in particular, in the fitting of linear or logistic regression models, the elastic net is a regularized regression method that linearly combines the L_1 and L_2 penalties of the lasso (least absolute shrinkage and selection operator is a regression analysis method that performs both variable selection and regularization in order to enhance the prediction accuracy and interpretability of the statistical model it produces) and ridge (it is the most commonly used method of regularization of ill-posed problems; in statistics, the method is known as ridge regression, in machine learning it is known as weight decay) methods, turning the method able to remove predictors with low impact to the outcome while regularizing for improved generalization. As our dataset is composed of several attributes, identifying the most important ones permits for a wider applicability and more practical use of our risk calculators.

To handle missing data, first, we removed all variables with more than 5% missing and imputed the remaining variables by

the variable mean, for numerical variables, or by the variable mode, for categorical variables. For each analysis, we split our data into training (75% of the whole sample) and test datasets (25%). We deployed a standard machine learning protocol with 10-fold cross-validation, feature selection, hyperparameter tuning and class imbalance correction in the training dataset (Fig. S1 in the Supplementary Material) with model selection based on the area under the receiver operating characteristic (ROC) curve. Class imbalance usually leads to very different sensitivity and specificity scores in the model; to account for this, we balanced each class predominance with class weighting. Each class was weighted inversely proportional to its frequency in the training set, or:

$$W_1 = F_0/N$$

where W_1 is the weight for the positive class and F_0 is the frequency of the negative class in the training set. Accordingly, $W_0 = 1 - W_1$. Class weighting allows us to use the whole training set instead of relying on down-sampling techniques.

We calculated individual-level predicted SA probabilities based on the elastic net algorithm, as well as ROC and AUC to evaluate prediction accuracy. Additionally, we calculated sensitivity, specificity, balanced accuracy, positive predictive value (PPV), and negative predictive value (NPV) when the cut-off is selected as 0.5. Elastic net regularization predicted probabilities were then discretized into deciles (10 groups of equal size ordered by percentiles) and cross-classified with observed SA.

From the variables selected by the elastic net algorithm to distinguish suicide attempters from non-attempters, only those with importance weighting factor higher than 20% were considered as clinically and epidemiologically relevant.

Results

Among the 959 adult patients included, there were 413 (43.1%) males and 546 females (56.9%). The mean (s.d.) age was 34.95 (12.94), with a median of 32, ranging from 9 to 82 years of age; the mean number of years studied was 14.51 (4.95), with a median of 14 and ranging from 1 to 31 years of study. Regarding marital status, 520 (54.2%) were single or without stable partners; 378 (39.4%) reported having children; 560 (58.4%) were Catholics, while 298 (31.1%) were from other religions and 101 (10.5%) said they had no religion. Of the 858 (89.5%) patients who had some religion, 521 (60.7%) admitted to being practitioners. From the occupational point of view, 415 (43.3%) were working, 166 (17.3%) were students, 102 (10.6%) were housewives, 104 (10.8%) were retired, 22 (15.6%) were unemployed, and when questioned about paid employment, 457 (47.7%) had no remuneration. The socioeconomic class, according to ABIPEME, showed that 157 (16.4%) were of class A, 369 (38.5%) of class B, 341 (35.6%) of class C, and 92 (9.6%) were of classes D and E. For clinical variables, 231 (24.1%) had never performed psychiatric or psychological treatment up to the interview. Of the 728 (75.9%) who had gone through specialized care, 675 (92.7%) used a serotonin reuptake inhibitor, 138 (19.0%) used another type of antidepressant, 365 (50.1%) used a benzodiazepine, 105 (14.4%) used a mood stabilizer, 49 (6.7%) used lithium, and 192 (26.4%) used a neuroleptic. Moreover, 609 (83.7%) had gone through psychotherapy [148 (20.3%) on cognitive-behavioral therapy for OCD]. There were 65 (8.9%) patients with a history of psychiatric hospitalization and 12 (1.7%) who had undergone electroconvulsive therapy.

From the psychopathological point of view, Table 1 shows the descriptive results of phenomena intrinsic to the symptomatological core of OCD, while Table 2 presents clinical characteristics extrinsic to the psychopathological nucleus of OCD.

Figure 1 shows the prevalence rates of suicidality in patients with OCD. Of those who committed an SA, 19 (1.99%) had a history of hospitalization – 15 (78.9%) in a general hospital and four (21.1%) in a psychiatric unit.

Machine learning analysis

The model showed an accuracy of 85% in distinguishing individual attempters from non-attempters. The elastic net algorithm selected 18 variables that could predict SA: (1) previous suicide plans; (2) history of suicidal thoughts; (3) lifetime depressive episode; (4) lifetime intermittent explosive disorder (IED); (5) lifetime substance use/dependence disorder; (6) lower socioeconomic class; (7) lifetime anorexia disorder; (8) lifetime ADHD; (9) mulatto ethnicity; (10) lifetime kleptomania disorder; (11) presence of any sensory phenomena; (12) lifetime simple phobia; (13) lifetime panic/agoraphobia disorder; (14) familial history of OCD; (15) familial history of alcohol dependence; (16) familial history of psychosis; (17) familial history of SA; and (18) having no occupation. Thus, the most relevant (higher than 20% of importance weighting factor) predictor variables that remained in the model were: (1) previous suicide plans, (2) history of suicidal thoughts, (3) presence of lifetime depressive episode, and (4) presence of lifetime IED. Figure 2 shows the absolute weighting factors of each clinical variable that remained in the model normalized by the largest factor.

Sensitivity and specificity of the model were 84.61% and 87.32%, respectively, with a balanced accuracy of 85.97% and significant at $\chi^2 p < 0.0001$. The PPV was 44.89% and the NPV was 97.89%, with an AUC of 0.95, obtained by ROC curve analysis (Fig. 3).

All samples were ranked based on their probability of belonging to the positive class and separated in deciles. The percentage of positive and negative class samples was then analyzed at each decile based on their expected outcome. This showed that over half of the OCD sample which had an SA were recognized to be at the highest risk of SA, and 85% were categorized as at highest or second-to-highest risk of SA deciles. Likewise, 100% of the sample was categorized in the first three deciles of predictive risk of having SA. Figure 4 shows the concentration risk for SA based on output probability model, illustrating how suitable to real-world scenario the model would be. A perfectly calibrated model (e.g. the predicted probability translates into the exact SA chance) would have a smoother curve across the deciles, which was not the case. However, the high prevalence of SA among the two first deciles shows that high probabilities predicted by the model concentrated the most cases of SA. Therefore, instead of relying only on the probabilities predicted by the model, this chart can be used to provide a more realistic suicide risk for individuals, since it is grounded on previous out-of-sample predictions.

Discussion

Describing suicidality in patients with OCD

In our study, 104 (10.8%) patients with OCD attempted suicide throughout their lives, a similar rate to the 10.3% average found in the systematic review by Angelakis et al. (2015) and identical to the median of 10.8% (mean of 14.2%) in clinical samples

Table 1. Descriptive results of the psychopathological variables intrinsic to the phenomenology of obsessive-compulsive disorder

	Mean (s.d.)	Median (minimum–maximum)
Age of (in years)		
Compulsions onset	13.2 (7.8)	11 (3–60)
Obsessions onset	13.5 (8.0)	12 (1–59)
Interference of OCS	22.3 (106)	20 (4–64)
Specific treatment sought	30.3 (12.7)	28 (6–75)
OCD diagnosis	31.9 (12.4)	29.5 (9–75)
Duration of OCD (in years)	23.2 (12.9)	22 (0–67)
Untreated gap of time (in years)	17.6 (13.0)	15 (0–58)
Severity of symptoms dimensions (DY-BOCS)		
Aggression/catastrophes	5.34 (5.0)	5 (0–15)
Sexual/religious/moral	4.36 (4.9)	2 (0–15)
Contamination/cleaning	6.22 (5.2)	7 (0–15)
Symmetry/ordering	7.39(4.6)	8 (0–15)
Hoarding	3.19 (4.1)	0 (0–15)
Miscellaneous	7.56 (4.7)	9 (0–15)
Total DY-BOCS Score	21.25 (6.2)	22 (0–30)
Severity of symptoms (Y-BOCS)		
Obsessions score	12.75 (3.9)	13 (0–20)
Compulsions score	12.89 (4.2)	13 (0–20)
Total Y-BOCS Score	25.60 (7.5)	26 (7–40)
Insight severity (BABS)	6.88 (5.5)	6 (0–24)
<i>n</i>		%
Symptoms dimension prevalence (DY-BOCS)		
Aggression/catastrophes	644	67.2
Sexual/religious/moral	554	57.8
Contamination/cleaning	705	73.5
Symmetry/ordering	832	86.8
Hoarding	511	53.3
Miscellaneous	837	87.3
Presence of any sensory phenomena	589	61.4
Course of OCS (<i>n</i> = 857)		
Intermittent/episodic	417	48.7
Continuous – stable gravity	62	7.2
Chronic – worsening severity	378	44.1
Family history		
Obsessive-compulsive disorder	480	50.1
Tics	178	18.6
Age of onset (<i>n</i> = 915)		
Early (<11 years)	460	50.3
Intermediate	320	35.0
Late (≥18 years)	135	14.7

s.d., standard deviation; *n*, absolute value; %, relative value; OCS, obsessive-compulsive symptoms; OCD, obsessive-compulsive disorder; DY-BOCS, Dimensional Yale-Brown Obsessive Compulsive Scale; Y-BOCS, Yale-Brown Obsessive Compulsive Scale; BABS, Brown Assessment of Beliefs Scale.

Table 2. Descriptive results of the psychopathological variables extrinsic to the phenomenology of obsessive-compulsive disorder

	Mean (s.d.)	Median (minimum–maximum)
Severity of depressive symptoms (BDI)	16.6 (11.3)	16 (0–53)
Severity of anxiety symptoms (BAI)	16.2 (11.4)	14 (0–53)
	<i>n</i>	%
Lifetime presence of any stressful life event	580	60.5
Lifetime presence of any traumatic event	294	30.7
Lifetime psychiatric comorbidities		
Tics	267	27.8
Tourette's disorder	78	8.1
Attention deficit hyperactivity disorder	137	14.3
Separation anxiety	268	27.9
Major depression	643	67.0
Dysthymia	110	11.5
Bipolar disorder type I	37	3.90
Bipolar disorder type II	39	4.10
Alcohol dependence	75	7.80
Cannabis/cocaine addiction	24	2.50
Panic disorder/agoraphobia	191	19.9
Social phobia	328	34.2
Specific phobia	297	31.0
Generalized anxiety	329	34.3
Post-traumatic stress disorder	186	19.4
Somatization disorder	23	2.40
Body dysmorphic disorder	109	11.4
Eating disorders	107	11.2
Anorexia	23	2.40
Bulimia	22	2.30
Periodic eating disorder ('Binge')	78	8.10
Skin picking disorder	162	16.9
Intermittent explosive disorder	74	7.40
Kleptomania	27	2.80
Compulsive buying	104	10.8

s.d., standard deviation; BDI, Beck Depression Inventory; BAI, Beck Anxiety Inventory; *n*, absolute value; %, relative value.

from the systematic review by Albert et al. (2018). Two previous studies by our group on suicide in OCD found SA rates of 11% in a subsample of 582 patients ($n = 64$) (Torres et al., 2011) and 19.4% in a sub-sample of 356 ($n = 69$) (Velloso et al., 2016).

However, the sample size of these two studies was smaller than the present study.

In the WHO and World Mental Health (WMH) Survey, which 108 705 adults from 21 countries were evaluated, the prevalence of SA was 0.3% for developed countries and 0.4% for developing countries (Borges et al., 2010). In our study, the rates of SA were higher than in population-based samples, since we used a sample from tertiary psychiatric clinical services. Our findings, therefore, are in line with previous findings that suicidal behavior is more prevalent in the OCD population than in the general population (Albert et al., 2018; Angelakis et al., 2015; Fernández de la Cruz et al., 2017).

Predictive factors of suicide attempted in patients with OCD

To the best of our knowledge, this is the first study to use machine learning techniques to explore possible predictive factors for individual SA in patients with OCD. We found 18 possible predictive variables of SA, but only four were clinically and epidemiologically relevant. Regarding sociodemographic variables, our negative finding is in agreement with previous studies showing no association between suicidality and marital status, family status, level of education, employment, religion, quality of life, clinical course, age of onset, or family history of suicide or suicidality (Kamath, Reddy, & Kandavel, 2007; Torres et al., 2007, 2011).

Several studies have pointed out that suicidality in OCD may be associated with the phenomena intrinsic to the psychopathological core of the disorder, which seems to have its effect mediated by the severity of OCD symptoms (Hung et al., 2010; Torres et al., 2007): sexual/religious/moral dimensions (Dell'Osso et al., 2018; Fernández de la Cruz et al., 2017; Torres et al., 2011), aggressiveness/catastrophe (Balci & Sevincok, 2010), and symmetry/organization and arrangement (Alonso et al., 2010). Obsessions involving sexual/religious/moral and aggressive/catastrophic content are considered as 'taboo thoughts' and could present a higher degree of shame, guilt, and responsibility for the content of their obsessions, which may lead to thoughts of death, suicide ideation, SA, and suicide (Angelakis et al., 2015; Balci & Sevincok, 2010; Gupta, Avasthi, Grover, & Singh, 2014; Kamath et al., 2007). Our findings, however, show that no OCD psychopathology intrinsic variable is predictive of SA. Thus, suicidality in OCD does not appear to be related to the disorder itself. Concurrently, Sareen et al. (2005), in a population study with traditional statistical analysis, did not find an independent association between OCD and suicidality, after adjustment of sociodemographic variables and control for psychiatric comorbidities. The only intrinsic variable from the nuclear phenomenology of OCD psychopathology that remained important in the machine learning model in our study was the presence of any sensory phenomenon but did not reach the importance value of pre-stipulated clinical-epidemiological relevance.

Comorbidities

The two lifetime comorbidities that remained predictive of SA were major depressive disorder and IED. In two previous systematic reviews, higher rates of suicidality were found in patients with OCD who presented concomitant comorbidities (especially higher severity of depressive and anxious symptoms) and previous history of suicidal behavior and feelings of hopelessness (Albert et al., 2018; Fernández de la Cruz et al., 2017). On the other hand, Torres et al. (2006), after controlling for variables such as the presence of

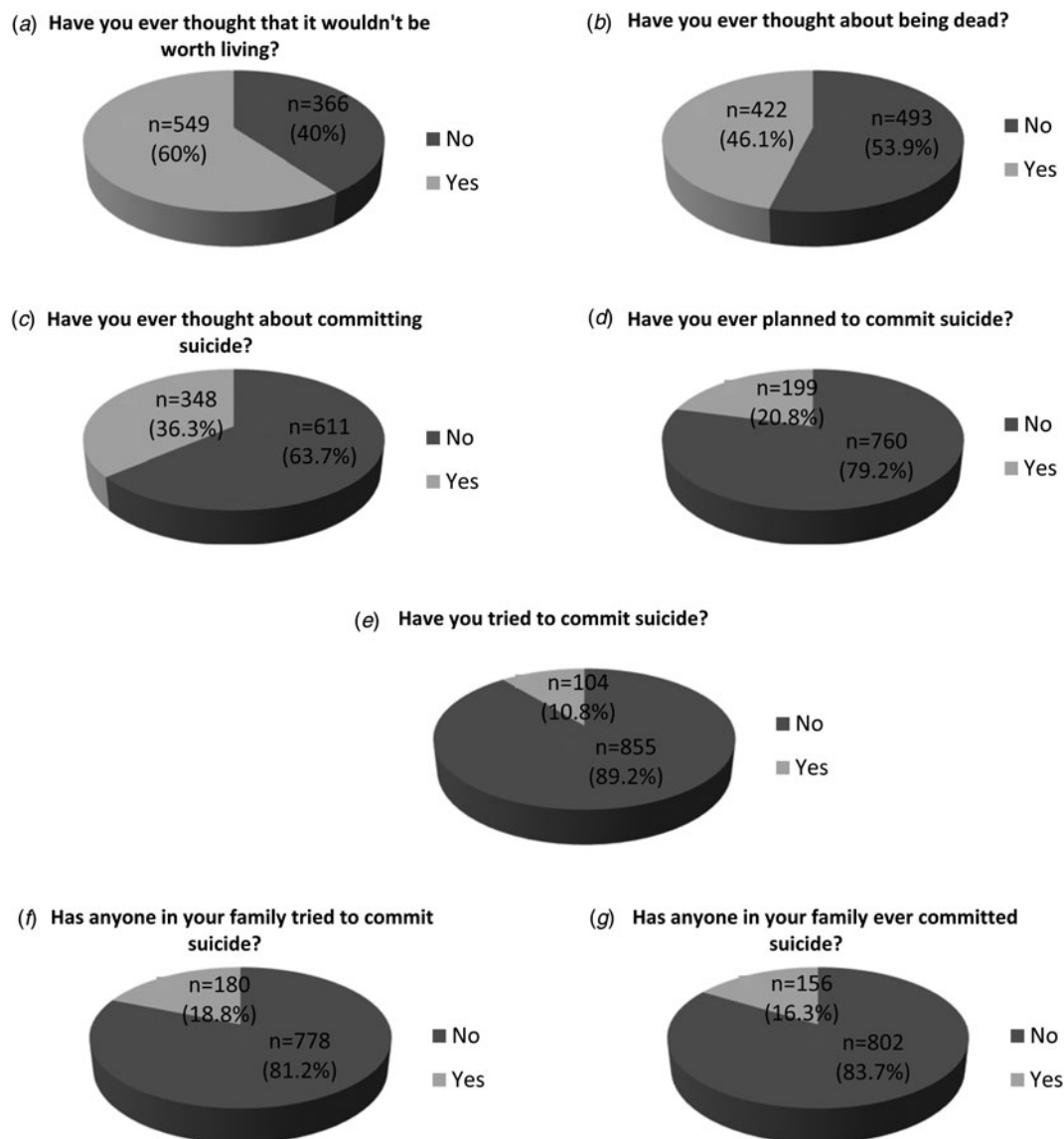


Fig. 1. Prevalence of suicidality phenomena in the sample of adults with obsessive-compulsive disorder. Graphics from 1a to 1g show the percentages of each of the aspects related to suicidality. Figure 1a, b = 915 patients with data available; Figure 1f, g = 958 patients with data available.

comorbidities, found that the risk of suicidal behavior in OCD remained significant, reaching up to 25% of prevalence, with an independent effect of OCD symptoms in suicidality, since patients with OCD without comorbidities did not differ from those with comorbidities in terms of the prevalence of suicidal behavior. Such divergence may be due to each study's samplings (clinical × epidemiological) and/or analytical approaches.

Depression as comorbidity predictive of attempted suicide in OCD

The presence of depressive symptoms has an important impact on the expression of suicidality among patients with OCD, since depression is one of the most common comorbid conditions in OCD (Kamath et al., 2007; Maina, Salvi, Tiezzi, Albert, & Bogetto, 2007; Tükel, Meteris, Koyuncu, Tecer, & Yazici, 2006). Depression alone is one of the major factors for suicidal behavior

(Bertolote et al., 2005) and is associated with all suicidal outcomes (Scocco, de Girolamo, Vilagut, & Alonso, 2008).

We can hypothesize that the existence of comorbid depressive symptoms results in more severe and incapacitating OCS, with a higher occurrence of suicidal behaviors compared to patients with OCD without depressive symptoms (Hollander et al., 1996). Since mood changes appear to be consequences of chronic stress and injury associated with the severity of OCD symptoms (Angst et al., 2004), depression diagnosis and/or severity may be an indicator of OCD severity and could 'link' OCD and suicide outcomes, even if secondary to OCD (Torres et al., 2007).

IED as comorbidity predictive of attempted suicide in OCD

In another study of our group, with a smaller sample ($n = 582$), the authors found that impulse control disorders were independently associated with SI, suicide plans, and SA (Torres et al., 2011). Our finding that IED comorbidity in patients with OCD

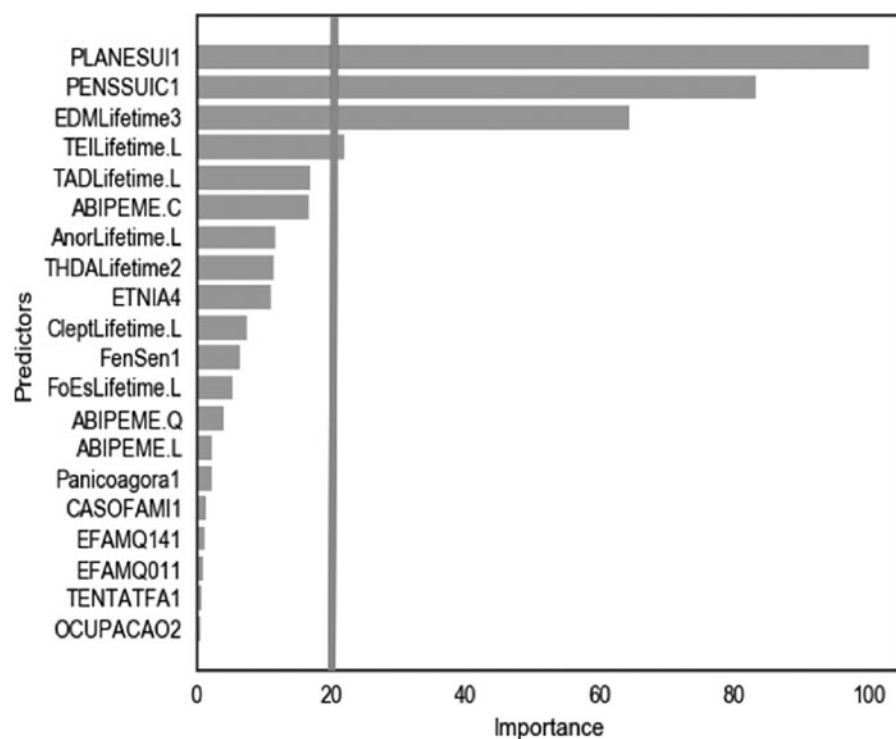


Fig. 2. Bar graph showing weighting factors assigned to each clinical variable by elastic net algorithm based on their relevance in distinguishing suicide attempters from non-attempters. The vertical line 'cuts' the importance weighting factors at the 20% level (considered clinically and epidemiologically relevant for this study).

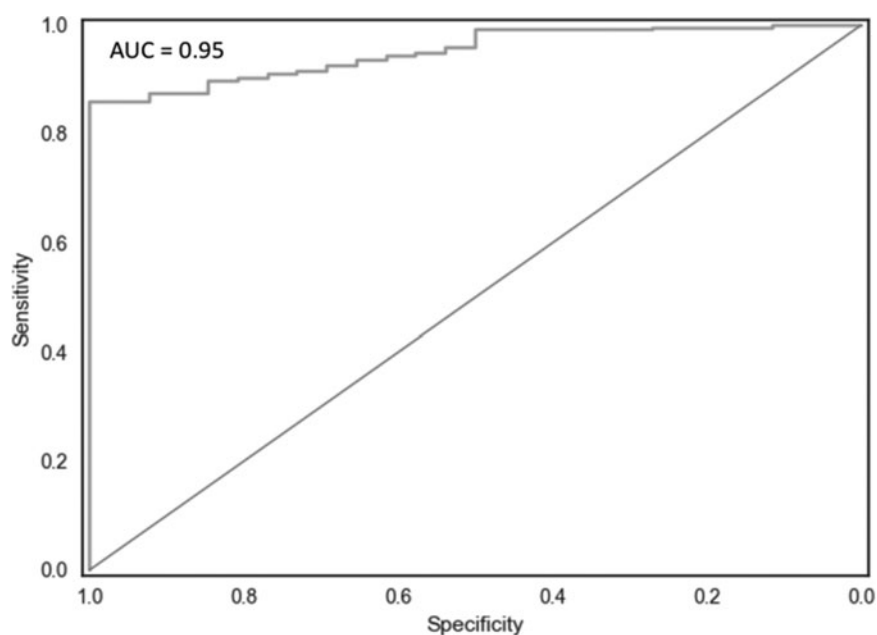


Fig. 3. Receiver operating characteristic curve of the predictive model.

may predict SA with machine learning analysis reinforces the relationship between the cited conditions. The association of IED with suicidality has been found in other studies, independently of OCD comorbidity (Gelegen & Tamam, 2018; Lee et al., 2007). A community-based survey found that impulse-control disorders were strongly associated with SA (Lee et al., 2007). Likewise, in a cross-sectional study, the lifetime prevalence of IED was 2.7 times higher among those with lifetime SA (Gelegen & Tamam, 2018). We hypothesized that impulsive and disruptive behavior can contribute to the association with suicidal phenomena in this group of patients. In our sample, the lifetime

prevalence of IED was higher (7.4%) than the 12-month prevalence in the USA (2.7%) (American Psychiatric Association, 2013), which may have impacted our results.

Psychopathological aspects of suicidality as predictive factors of SA in patients with OCD

In our study, the risk of SA in OCD can be predicted with high accuracy by SI and plans, as these two variables accounted for about 85% of the prediction of SA. As expected, the results corroborate that SI, SA, and consummated suicide represent different

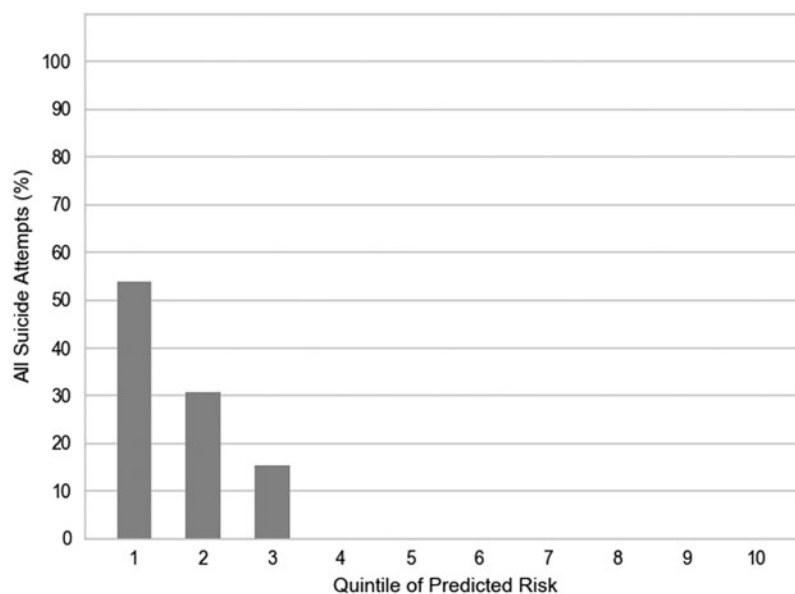


Fig. 4. Concentration of risk for suicide attempts among OCD patients.

aspects of the same psychopathological phenomenon: the spectrum of suicidality. In two other studies in our group, there was also a progressive severity continuity of the phenomenon of suicide in patients with OCD, evolving from ideation to planning and attempting suicide, culminating in '*per se*' consumption of suicide (Torres et al., 2011; Velloso et al., 2016).

According to the systematic review by Albert et al. (2018), individuals with OCD had an odds ratio ranging from 1.9 to 10.3 to present lifetime SI, when compared to the general population. The risk levels remained significant even after controlling for confounding variables (AOR between 3.8 and 5.58). In a prospective population-based study (36 000 patients with OCD), Fernández de la Cruz et al. (2016) reported that most patients with prior SA had at least one comorbid psychiatric pathology, while about 40% of those with a history of completed suicide had OCD without comorbidities. Thus, the authors concluded that OCD without comorbidity is more associated with death by suicide and OCD with comorbidities is more associated with SA. After controlling for the impact of different comorbidities, the risk remained significant, which would reinforce the idea that OCD alone poses a higher risk of suicidal behavior, contrary to our study.

In the present study, SA in patients with OCD was not predicted by characteristics of the disorder itself, but by precursor aspects of SA (thoughts and suicidal planning) and two comorbidities: major depression and IED. Such a finding seem to diverge from previous studies, where OCD was associated with suicidal behavior even after control of confounding factors; such difference may be due to the use of different analytic methodology. The use of ML algorithms is better applied in scenarios where numerous variables must be considered simultaneously to estimate the probability of an event occurring. ML techniques do not ignore small effects to perform predictions or to identify patterns (Mwangi, Ebmeier, Matthews, & Steele, 2012; Perlis, 2013).

Machine learning for prediction of suicidality and future directions

Standard investigation, including analysis in the field of suicide in OCD, has focused on traditional statistical approaches that explore a linear relationship between variables at group-level

data. In this context, machine learning approaches can be advantageous and have increasingly been used in prognostic psychiatry, as they can assume a complex relationship between variables, including non-linear patterns, and are focused at an individual patient level (Dwyer, Falkai, & Koutsouleris, 2018). In the last 6 years, therefore, several studies proposed machine learning models to predict SA in general population. A systematic review and simulation study about suicide prediction models in general population, however, showed that machine learning for suicide prediction had good classification accuracy, but with low predictive positive value (Belsher et al., 2019). It is somewhat expected that models with rare events, such as suicide in general population, have a low PPV. Accordingly, we suggest that future studies should address specific populations with higher rates for SA, such as patients with OCD or depressive disorders.

Other studies have evaluated suicide outcomes among samples from patients with specific psychiatric disorders. Passos et al. (2016) showed a clinical signature using a relevant vector machine to classify SA history throughout life in patients diagnosed with major depression or bipolar disorder. The most accurate model showed an AUC of 0.77, sensitivity of 0.72, and specificity of 0.71. The most relevant predictor variables were previous hospitalizations for depression, history of psychosis throughout life, cocaine dependence, and comorbidity with post-traumatic stress disorder (Passos et al., 2016).

Another study (Oh, Yun, Hwang, & Chae, 2017) built artificial neural network models to identify SA in 1 month, 1 year, and throughout life among individuals with depression and anxiety disorders, using self-reported psychiatric questionnaires. This report found an overall accuracy of 93.7% in 1 month, 90.8% in 1 year, and 87.4% in lifetime SA detection. Lastly, a cross-sectional study developed four machine learning models to estimate the probability of a patient with schizophrenia for attempting suicide. The model with better performance found an AUC = 0.71 (Hettige et al., 2017).

Lastly, we hypothesize that future studies integrating data from different biological levels, such as genetics, digital health, and other socio-environmental data, in specific populations could help to build more accurate models.

Limitations

The limitations of this study include the clinical-referred nature of the sample, thus possibly not representative of patients with OCD in the community. It is a cross-sectional study, subject to memory bias and therefore not ideal for the identification of predictors. With respect to the machine learning analysis, the elastic net presents some important limitations. Non-linear relationships between the outcome and the predictors are not captured, which might be the case in a complex phenomenon like SA. It also is the case that elastic net fails to include relationships between variables, only relying on the same principles that traditional regression does, although with improved capabilities for out-of-sample generalization.

Since the database was not specially constructed to study the suicidality spectrum, a specific and validated questionnaire was not applied, which impacts directly on the present report's premise. Moreover, aspects such as duration, variability, and persistence of the suicidal behaviors were not measured. Also, no control of potentially triggering variables at the time of SA was done, as well as interference of the severity of OCD symptoms and of psychiatric comorbidities to lead to SA. A smaller part of our sample had previously been evaluated in two other studies in our group, which may be argued to generate circular knowledge; however, the methodology applied differs fundamentally, confirming some results, but proposing others, diverse from the previous literature. Finally, the predictor variables included in the algorithm were defined based on previous studies and may not ideally represent all possible predictors.

Conclusions

This is the first study that uses ML algorithm to explore the predictors of individual risk of SA in patients with OCD. Our study adds to evidence that suicidality is a relevant phenomenon in OCD patients and demonstrates that an algorithm of ML can predict SA among patients with OCD. Our results also indicate that suicidality appears as a continuum and that it is imperative to analyze and investigate manifestations of suicidality actively in all patients with OCD, but more aggressively among those with comorbid depressive symptoms. The ML techniques are ideal for identification of the pattern of multiple variables with complex combinations and can be used to stratify the risk of SA for a given person by comparing the probability of SA predicted by the algorithm for their variables as compared to those calculated from a dataset of subjects with known outcomes.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/S0033291720002329>.

Acknowledgements. We thank Dr James Leckman for sharing unpublished findings with our team, which were used for to evaluate the history of stressful events in patients with OCD.

Financial support. The present work was carried out with the support of the Coordination of Improvement of Higher Education Personnel – Brazil (CAPES) – Financing Code 001 [Ygor Arzeno Ferrão and Neusa Agne were stockholders of the National Research Council (CNPq)].

Conflict of interest. The authors declare that there is no conflict of interest regarding the publication of this article.

References

Albert, U., De Ronchi, D., Maina, G., & Pompili, M. (2018). Suicide risk in obsessive-compulsive disorder and exploration of risk factors: A systematic

review. *Current Neuropharmacology*, 17(8), 681–696. <https://doi.org/10.2174/1570159X16666180620155941>.

Alonso, P., Segalàs, C., Real, E., Pertusa, A., Labad, J., Jiménez-Murcia, S., & ... Menchón, J. M. (2010). Suicide in patients treated for obsessive-compulsive disorder: A prospective follow-up study. *Journal of Affective Disorders*, 124(3), 300–308. <https://doi.org/10.1016/j.jad.2009.12.001>.

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

Angelakis, I., Gooding, P., Tarrier, N., & Panagioti, M. (2015). Suicidality in obsessive compulsive disorder (OCD): A systematic review and meta-analysis. *Clinical Psychology Review*, 39, 1–15. <https://doi.org/10.1016/j.cpr.2015.03.002>.

Angst, J., Gamma, A., Endrass, J., Goodwin, R., Ajdacic, V., Eich, D., & Rössler, W. (2004). Obsessive-compulsive severity spectrum in the community: Prevalence, comorbidity, and course. *European Archives of Psychiatry and Clinical Neuroscience*, 254(3), 156–164. <https://doi.org/10.1007/s00406-004-0459-4>.

Arsenault-Lapierre, G., Kim, C., & Turecki, G. (2004). Psychiatric diagnoses in 3275 suicides: A meta-analysis. *BMC Psychiatry*, 37(4). <https://doi.org/10.1186/1471-244X-4-37>.

Asbahr, F. R., Lotufo, N. F., Turecki, G. X., & Miguel, E. C. (1996). Escala Yale-Brown de Sintomas Obsessivo-Compulsivos, tradução brasileira., in: *Transtornos Do Espectro Obsessivo-Compulsivo*. Guanabara-Koogan, Rio de Janeiro, pp. 219–230.

Balci, V., & Sevincok, L. (2010). Suicidal ideation in patients with obsessive-compulsive disorder. *Psychiatry Research*, 175(1–2), 104–108. <https://doi.org/10.1016/j.psychres.2009.03.012>.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, 56(6), 893–897. <https://doi.org/10.1037/0022-006x.56.6.893>.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561–571. <https://doi.org/10.1001/archpsyc.1961.01710120031004>.

Belsher, B. E., Smolenski, D. J., Pruitt, L. D., Bush, N. E., Beech, E. H., Workman, D. E., ... Skopp, N. A. (2019). Prediction models for suicide attempts and deaths: A systematic review and simulation. *JAMA Psychiatry*, 76(6), 642–651. [doi:10.1001/jamapsychiatry.2019.0174](https://doi.org/10.1001/jamapsychiatry.2019.0174).

Berk, R. A. (2008). *Statistical learning from a regression perspective* (1st ed.). New York: Springer.

Bertolote, J. M., Fleischmann, A., De Leo, D., Bolhari, J., Botega, N., De Silva, D., ... Wasserman, D. (2005). Suicide attempts, plans, and ideation in culturally diverse sites: The WHO SUPRE-MISS community survey. *Psychological Medicine*, 35(10), 1457–1465. <https://doi.org/10.1017/S0033291705005404>.

Borges, G., Nock, M. K., Haro Abad, J. M., Hwang, I., Sampson, N. A., Alonso, J., ... Kessler, R. C. (2010). Twelve-month prevalence of and risk factors for suicide attempts in the World Health Organization World Mental Health Surveys. *Journal of Clinical Psychiatry*, 71(12), 1617–1628. <https://doi.org/10.4088/JCP.08m04967blu>.

Burke, T. A., Ammerman, B. A., & Jacobucci, R. (2019). The use of machine learning in the study of suicidal and non-suicidal self-injurious thoughts and behaviors: A systematic review. *Journal of Affective Disorders*, 245, 869–884.

Coryell, W. (1981). Obsessive-compulsive disorder and primary unipolar depression. Comparisons of background, family history, course, and mortality. *The Journal of Nervous Mental Disease*, 164(4), 220–224. <https://doi.org/10.1093/ajph.100.10.1511>.

de Ávila Berni, G., Rabelo-da-Ponte, F. D., Librenza-Garcia, D., Boeira, M. V., Kauer-Sant'Anna, M., Passos, I. C., ... Kapczinski, F. (2018). Potential use of text classification tools as signatures of suicidal behavior: A proof-of-concept study using Virginia Woolf's personal writings. *PLoS ONE*, 13(10), e0204820. <https://doi.org/10.1371/journal.pone.0204820>.

Dell'Osso, B., Benatti, B., Arici, C., Palazzo, C., Altamura, A.C., Hollander, E., ... Zohar, J. (2018). Prevalence of suicide attempt and clinical characteristics of suicide attempters with obsessive-compulsive disorder: A report from the International College of Obsessive-Compulsive Spectrum Disorders (ICOCs). *CNS Spectrums*, 23(1), 59–66. <https://doi.org/10.1017/S1092852917000177>.

- DiGangi, J. A., Gomez, D., Mendoza, L., Jason, L. A., Keys, C. B., & Koenen, K. C. (2013). Pretrauma risk factors for posttraumatic stress disorder: A systematic review of the literature. *Clinical Psychology Review*, 33(6), 728–744. <https://doi.org/10.1016/j.cpr.2013.05.002>.
- Dwyer, D. B., Falkai, P., & Koutsouleris, N. (2018). Machine learning approaches for clinical psychology and psychiatry. *Annual Review of Clinical Psychology*, 14(1), 91–118. <https://doi.org/10.1146/annurev-clinpsy-032816-045037>.
- Eisen, J. L., Mancebo, M. A., Pinto, A., Coles, M. E., Pagano, M. E., Stout, R., ... Rasmussen, S. A. (2006). Impact of obsessive-compulsive disorder on quality of life. *Comprehensive Psychiatry*, 47(4), 270–275. <https://doi.org/10.1016/j.comppsy.2005.11.006>.
- Fernández de la Cruz, L., Rydell, M., Runeson, B., D'Onofrio, B. M., Brander, G., Rück, C., ... Mataix-Cols, D. (2017). Suicide in obsessive-compulsive disorder: A population-based study of 36 788 Swedish patients. *Molecular Psychiatry*, 22(11), 1626–1632. <https://doi.org/10.1038/mp.2016.115>.
- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. (1997). *Structured clinical interview for DSM-IV axis I disorders: clinical version (SCID-CV)*. Washington, DC: American Psychiatric Press.
- Franklin, J. C., Ribeiro, J. D., Fox, K. R., Bentley, K. H., Kleiman, E. M., Huang, X., ... Nock, M. K. (2017). Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychological Bulletin*, 143(2), 187–232. <https://doi.org/10.1037/bul0000084>.
- Gelegen, V., & Tamam, L. (2018). Prevalence and clinical correlates of intermittent explosive disorder in Turkish psychiatric outpatients. *Comprehensive Psychiatry*, 83, 64–70. <https://doi.org/10.1016/j.comppsy.2018.03.003>.
- Goodwin, D. W., Guze, S. B., & Robins, E. (1969). Follow-up studies in obsessional neurosis. *Archives of General Psychiatry*, 20(2), 182–187. <https://doi.org/10.1001/archpsyc.1969.01740140054006>.
- Gorenstein, C., & Andrade, L. (1996). Validation of a Portuguese version of the Beck Depression Inventory and the State-Trait Anxiety Inventory in Brazilian subjects. *Brazilian Journal of Medical and Biological Research*, 29(4), 453–457.
- Greenhalgh, T., Howick, J., & Maskrey, N. (2014). Evidence based medicine: A movement in crisis? *British Medical Journal*, 348, g3725. <https://doi.org/10.1136/bmj.g3725>.
- Gupta, G., Avasthi, A., Grover, S., & Singh, S. M. (2014). Factors associated with suicidal ideations and suicidal attempts in patients with obsessive compulsive disorder. *Asian Journal of Psychiatry*, 12, 140–146. <https://doi.org/10.1016/j.ajp.2014.09.004>.
- Hawton, K., & van Heeringen, K. (2009). Suicide. *Lancet (London, England)*, 373(9672), 1372–1381. [https://doi.org/10.1016/S0140-6736\(09\)60372-X](https://doi.org/10.1016/S0140-6736(09)60372-X).
- Hettige, N. C., Nguyen, T. B., Yuan, C., Rajakulendran, T., Baddour, J., Bhagwat, N., ... De Luca, V. (2017). Classification of suicide attempters in schizophrenia using sociocultural and clinical features: A machine learning approach. *General Hospital Psychiatry*, 47, 20–28. <https://doi.org/10.1016/j.genhosppsych.2017.03.001>.
- Hollander, E., Greenwald, S., Neville, D., Johnson, J., Hornig, C. D., & Weissman, M. M. (1996). Uncomplicated and comorbid obsessive-compulsive disorder in an epidemiologic sample. *Depression and Anxiety*, 4(3), 111–119. [https://doi.org/10.1002/\(SICI\)1520-6394\(1996\)4:3<111::AID-DA3>3.0.CO;2-J](https://doi.org/10.1002/(SICI)1520-6394(1996)4:3<111::AID-DA3>3.0.CO;2-J).
- Hung, T.-C., Tang, H.-S., Chiu, C.-H., Chen, Y.-Y., Chou, K.-R., Chiou, H.-C., ... Chang, H. J. (2010). Anxiety, depressive symptom and suicidal ideation of outpatients with obsessive compulsive disorders in Taiwan. *Journal of Clinical Nursing*, 19(2122), 3092–3101. <https://doi.org/10.1111/j.1365-2702.2010.03378.x>.
- Huys, Q. J. M., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience*, 19, 404–413. <https://doi.org/10.1038/nn.4238>.
- Jacoby, R. J., Leonard, R. C., Riemann, B. C., & Abramowitz, J. S. (2014). Predictors of quality of life and functional impairment in obsessive-compulsive disorder. *Comprehensive Psychiatry*, 55(5), 1195–1202. <https://doi.org/10.1016/j.comppsy.2014.03.011>.
- Kamath, P., Reddy, Y. C. J., & Kandavel, T. (2007). Suicidal behavior in obsessive-compulsive disorder. *Journal of Clinical Psychiatry*, 68(11), 1741–1750. <https://doi.org/10.4088/jcp.v68n1114>.
- Kanwar, A., Malik, S., Prokop, L. J., Sim, L. A., Feldstein, D., Wang, Z., ... Murad, M. H. (2013). The association between anxiety disorders and suicidal behaviors: A systematic review and meta-analysis. *Depression and Anxiety*, 30(10), 917–929. <https://doi.org/10.1002/da.22074>.
- Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., ... Ryan, N. (1997). Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): Initial reliability and validity data. *Journal of the American Academy Child Adolescent Psychiatry*, 36(10), 980–988. <https://doi.org/10.1097/00004583-199707000-00021>.
- Kessler, R. C., Chiu, W. T., Demler, O., Merikangas, K. R., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62(6), 617–627. <https://doi.org/10.1001/archpsyc.62.6.617>.
- Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, 28(5), 1–26. <https://doi.org/10.18637/jss.v028.i05>.
- Larsen, M. E., Cummins, N., Boonstra, T. W., O'Dea, B., Tighe, J., Nicholas, J., ... Christensen, H. (2015). The use of technology in Suicide Prevention. Paper presented at the 37th annual international conference of the IEEE Engineering in Medicine and Biology Society, Milano, Italy, pp. 7316–7319. <https://doi.org/10.1109/EMBC.2015.7320081>.
- Lee, S., Fung, S. C., Tsang, A., Liu, Z. R., Huang, Y. Q., He, Y. L., ... Kessler, R. C. (2007). Lifetime prevalence of suicide ideation, plan, and attempt in metropolitan China. *Acta Psychiatrica Scandinavica*, 116(6), 429–437. <https://doi.org/10.1111/j.1600-0447.2007.01064.x>.
- Leonard Westgate, C., Shiner, B., Thompson, P., & Watts, B. V. (2015). Evaluation of veterans' suicide risk with the use of linguistic detection methods. *Psychiatric Services*, 66(10), 1051–1056. <https://doi.org/10.1176/appi.ps.201400283>.
- Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., ... Memish, Z. A. (2012). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the Global Burden of Disease Study 2010. *Lancet (London, England)*, 380(9859), 2095–2128. [https://doi.org/10.1016/S0140-6736\(12\)61728-0](https://doi.org/10.1016/S0140-6736(12)61728-0).
- Maina, G., Salvi, V., Tiezzi, M. N., Albert, U., & Bogetto, F. (2007). Is OCD at risk for suicide? A case-control study. *Clinical Neuropsychiatry: Journal of Treatment Evaluation*, 4(3), 117–121.
- Maziade, M., Roy, M. A., Fournier, J. P., Cliche, D., Mérette, C., Caron, C., ... Dion, C. (1992). Reliability of best-estimate diagnosis in genetic linkage studies of major psychoses: Results from the Quebec pedigree studies. *American Journal of Psychiatry*, 149(12), 1674–1686. <https://doi.org/10.1176/ajp.149.12.1674>.
- Miguel, E. C., Ferrão, Y. A., do Rosário, M. C. de Mathis, M. A., Torres, A. R., Fontenelle, L. F., ... da Silva, E. D. (2008). The Brazilian Research Consortium on Obsessive-Compulsive Spectrum Disorders: Recruitment, assessment instruments, methods for the development of multicenter collaborative studies and preliminary results. *Revista Brasileira de Psiquiatria*, 30(3), 185–196. <http://dx.doi.org/10.1590/S1516-44462008000300003>.
- Murray, C. J. L., Lopez, A. D., & World Health Organization, World Bank & Harvard School of Public Health. (1996). The Global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020. In Christopher J. L. Murray, Alan D. Lopez (Eds.), Boston: Harvard School of Public Health. Retrieved from <https://apps.who.int/iris/handle/10665/41842>.
- Mwangi, B., Ebmeier, K. P., Matthews, K., & Steele, J. D. (2012). Multi-centre diagnostic classification of individual structural neuroimaging scans from patients with major depressive disorder. *Brain*, 135, 1508–1521.
- Niculescu, A. B., Levey, D. F., Phalen, P. L., Le-Niculescu, H., Dainton, H. D., Jain, N., ... Salomon, D. R. (2015). Understanding and predicting suicidality using a combined genomic and clinical risk assessment approach. *Molecular Psychiatry*, 20(11), 1266–1285. <https://doi.org/10.1038/mp.2015.112>.
- Nock, M. K., Hwang, I., Sampson, N. A., & Kessler, R. C. (2010). Mental disorders, comorbidity and suicidal behavior: Results from the National Comorbidity Survey Replication. *Molecular Psychiatry*, 15(8), 868–876. <https://doi.org/10.1038/mp.2009.29>.

- O'Dea, B., Wan, S., Batterham, P. J., Calear, A. L., Paris, C., & Christensen, H. (2015). Detecting suicidality on Twitter. *Internet Interventions*, 2(2), 183–188. <https://doi.org/10.1016/j.invent.2015.03.005>.
- Oh, J., Yun, K., Hwang, J.-H., & Chae, J. H. (2017). Classification of suicide attempts through a machine learning algorithm based on multiple systemic psychiatric scales. *Frontiers in Psychiatry*, 8, 192. <https://doi.org/10.3389/fpsy.2017.00192>.
- Ozer, E. J., Best, S. R., Lipsey, T. L., & Weiss, D. S. (2003). Predictors of post-traumatic stress disorder and symptoms in adults: A meta-analysis. *Psychological Bulletin*, 129, 52–73. <https://doi.org/10.1037/0033-2909.129.1.52>.
- Passos, I. C., Ballester, P. L., Barros, R. C., Librenza-Garcia, D., Mwangi, B., Birmaher, B., ... Kapczynski, F. (2019a). Machine learning and big data analytics in bipolar disorder: A position paper from the International Society for Bipolar Disorders Big Data Task Force. *Bipolar Disorders*, 21(7), 582–594. <https://doi.org/10.1111/bdi.12828>.
- Passos, I. C., & Mwangi, B. (2018). Machine learning-guided intervention trials to predict treatment response at an individual patient level: An important second step following randomized clinical trials. *Molecular Psychiatry*, 25(4), 701–702. <https://doi.org/10.1038/s41380-018-0250-y>.
- Passos, I. C., Mwangi, B., Cao, B., Hamilton, J. E., Wu, M. J., Zhang, X. Y., ... Soares, J. C. (2016). Identifying a clinical signature of suicidality among patients with mood disorders: A pilot study using a machine learning approach. *Journal of Affective Disorders*, 193, 109–116. <https://doi.org/10.1016/j.jad.2015.12.066>.
- Passos, I. C., Mwangi, B., & Kapczynski, F. (2019b). *Personalized psychiatry: Big data analytics in mental Health* (1st ed.). New York, United States: Springer International Publishing Company, Incorporated.
- Perlis, R. H. (2013). A clinical risk stratification tool for predicting treatment resistance in major depressive disorder. *Biological Psychiatry*, 74(1), 7–14. <https://doi.org/10.1016/j.biopsych.2012.12.007>.
- Rajkumar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *The New England Journal of Medicine*, 380, 1347–1358. <https://doi.org/10.1056/NEJMr1814259>.
- R Core Team (2013) A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from: <http://www.R-project.org>.
- Rosario-Campos, M. C., Miguel, E. C., Quatrano, S., Chacon, P., Ferrao, Y., Findley, D., ... Leckman, J. F. (2006). The Dimensional Yale-Brown Obsessive-Compulsive Scale (DY-BOCS): An instrument for assessing obsessive-compulsive symptom dimensions. *Molecular Psychiatry*, 11(5), 495–504. <https://doi.org/10.1038/sj.mp.4001798>.
- Ruscio, A. M., Stein, D. J., Chiu, W. T., & Kessler, R. C. (2010). The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. *Molecular Psychiatry*, 15(1), 53–63. <https://doi.org/10.1038/mp.2008.94>.
- Sareen, J., Cox, B. J., Afifi, T. O., de Graaf, R., Asmundson, G. J. G., & ten Have, M. (2005). Anxiety disorders and risk for suicidal ideation and suicide attempts: A population-based longitudinal study of adults. *Archives of General Psychiatry*, 62(11), 1249–1257. <https://doi.org/10.1001/archpsyc.62.11.1249>.
- Scocco, P., de Girolamo, G., Vilagut, G., & Alonso, J. (2008). Prevalence of suicide ideation, plans, and attempts and related risk factors in Italy: Results from the European Study on the Epidemiology of Mental Disorders – World Mental Health study. *Comprehensive Psychiatry*, 49(1), 13–21. <https://doi.org/10.1016/j.comppsy.2007.08.004>.
- Simon, G. E., Johnson, E., Lawrence, J. M., Rossom, R. C., Ahmedani, B., Lynch, F. L., ... Shortreed, S. M. (2018). Predicting suicide attempts and suicide deaths following outpatient visits using electronic health records. *American Journal of Psychiatry*, 175(10), 951–960. <https://doi.org/10.1176/appi.ajp.2018.17101167>.
- SPSS Inc. Released (2009). PASW Statistics for Windows, Version 18.0. Chicago: SPSS Inc.
- Tolin, D. F., & Foa, E. B. (2006). Sex differences in trauma and posttraumatic stress disorder: A quantitative review of 25 years of research. *Psychological Bulletin*, 132(6), 959–992. <https://doi.org/10.1037/0033-2909.132.6.959>.
- Torous, J., Larsen, M. E., Depp, C., Cosco, T. D., Barnett, I., Nock, M. K., ... Firth, J. (2018). Smartphones, sensors, and machine learning to advance real-time prediction and interventions for suicide prevention: A review of current progress and next steps. *Current Psychiatry Reports*, 20(7), 51. <https://doi.org/10.1007/s11920-018-0914-y>.
- Torres, A. R., de Abreu Ramos-Cerqueira, A. T., Torresan, R. C., de Souza Domingues, M., Hercos, A. C. R., & Guimarães, A. B. C. (2007). Prevalence and associated factors for suicidal ideation and behaviors in obsessive-compulsive disorder. *CNS Spectrums*, 12, 771–778. <https://doi.org/10.1017/s1092852900015467>.
- Torres, A. R., Prince, M. J., Bebbington, P. E., Bhugra, D., Brugha, T. S., Farrell, M., ... Singleton, N. (2006). Obsessive-compulsive disorder: Prevalence, comorbidity, impact and help-seeking in the British National Psychiatric Morbidity Survey of 2000. *American Journal of Psychiatry*, 163(11), 1978–1985. <https://doi.org/10.1176/ajp.2006.163.11.1978>.
- Torres, A. R., Ramos-Cerqueira, A. T. A., Ferrão, Y. A., Fontenelle, L. F., do Rosário, M. C., & Miguel, E. C. (2011). Suicidality in obsessive-compulsive disorder: Prevalence and relation to symptom dimensions and comorbid conditions. *The Journal of Clinical Psychiatry*, 72(1), 17–26. <https://doi.org/10.4088/JCP.09m05651blu>.
- Tükel, R., Meteris, H., Koyuncu, A., Tecer, A., & Yazici, O. (2006). The clinical impact of mood disorder comorbidity on obsessive-compulsive disorder. *European Archives of Psychiatry and Clinical Neuroscience*, 256(4), 240–245. <https://doi.org/10.1007/s00406-006-0632-z>.
- Velloso, P., Piccinato, C., Ferrão, Y., Aliende Perin, E., Cesar, R., Fontenelle, L., ... do Rosário, M. C. (2016). The suicidality continuum in a large sample of obsessive-compulsive disorder (OCD) patients. *European Psychiatry: The Journal of the Association of European Psychiatrists*, 38, 1–7. <https://doi.org/10.1016/j.eurpsy.2016.05.003>.
- World Health Organization. (2014). Preventing suicide: A global imperative. Retrieved from: https://apps.who.int/iris/bitstream/handle/10665/131056/9789241564779_eng.pdf?sequence=1.
- World Health Organization. (2017). Depression and other common mental disorders: Global health estimates. Retrieved from: <https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf>.
- Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2), 301–320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>.