

## LETTER TO THE EDITORS

# Machine-learning approaches in psychotherapy: a promising tool for advancing the understanding of the psychotherapeutic process

Braz J Psychiatry. 2019 Nov-Dec;41(6):568-569  
doi:10.1590/1516-4446-2018-0295



Mental health researchers have been using machine learning (ML) techniques to improve outcomes. For example, recent studies have reported high predictive accuracy in distinguishing patients with bipolar disorder from healthy individuals by using neuroimaging, neurocognitive data,<sup>1</sup> and biomarkers.<sup>2</sup> Another work was able to predict which disordered patients would attempt suicide.<sup>3</sup> This new approach to analyzing data is also being used in the context of psychotherapy, especially to predict responses to psychotherapeutic treatment or indications for psychotherapy.<sup>4,5</sup>

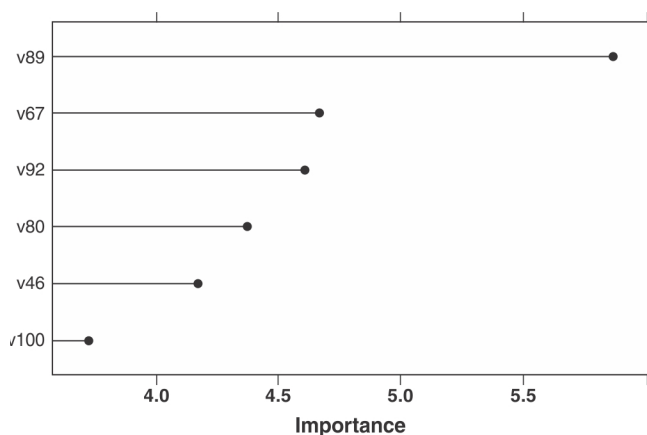
As most ML algorithms seek non-linear patterns of interrelation, they can be a useful tool in predicting the terms needed to apprehend complex phenomena such as psychotherapeutic encounters. Therefore, to evaluate their applicability, we hypothesize that ML can predict patient distress after sessions by identifying factors in the therapist that affect the patient's mental state during a session of psychotherapy.

To illustrate this, we describe the case of psychoanalytic psychotherapy of a 67-year-old woman diagnosed with somatic symptom disorder and Cluster C personality traits. The psychotherapist was a female clinical psychologist with 10 years of experience and training in psychoanalytic psychotherapy. The treatment consisted of 120 videotaped sessions and was considered successful in terms of its results. We undertook an exploratory analysis by using an ML approach. Distress following sessions was assessed by the Outcome Questionnaire (OQ-45), and the therapist's behaviors and interventions were measured by the Psychotherapy Process Q-Set (PQS). Two trained and independent judges rated randomly chose videotapes of every other session (a total of 62 sessions were therefore assessed). Inter-rater reliability showed a mean Pearson's correlation of  $r = 0.71$ . In this study, we aimed to evaluate the effect of the therapist's variables on the patient's distress and, considering that a categorical outcome variable would be more suitable for the model, we used the OQ median to classify distress after the session when high distress levels (OQ  $\geq 67$ ) or low distress levels (OQ  $< 67$ ) were present. Also, the

patient was compared to herself in terms of distress score during the process.

We assumed a classification problem with the PQS items associated with therapist effects ( $n=41$ ) as input data. Recursive feature elimination was carried out with leave-one-out cross-validation (LOOCV), using a random forest algorithm. We set the algorithm to look for a predictive model with 2 to 15 predictors out of the initial 41, with a view to obtaining a pragmatic model that could be used in clinical practice while avoiding overfitting. The best model found comprised 6 variables. Next, we used the selected variables to create predictive models with the random forest algorithm. Also known as decision tree forests, this ensemble-based method focuses only on ensembles of decision trees.<sup>6</sup> This method was developed by Leo Breiman<sup>7</sup> and combines the basic principles of "bagging" with random feature selection to add additional diversity to the decision tree models. The parameter to be adjusted was 'mtry' (an optional integer specifying the number of features to randomly select at each split) for this model. Finally, we used LOOCV to estimate model performance and plotted a receiver operating characteristic (ROC) curve, using the area under the curve (AUC) to select the best fit for each model. This process consists of training the algorithm with all subjects but one, a process that is then constantly repeated until all subjects have been used at least once. LOOCV has become the standard for estimating model performance for studies with small sample sizes.<sup>6</sup> The best model showed an AUC of 0.725, sensitivity of 79%, specificity of 62%, balanced accuracy of 70.5%, and comprised 6 variables (PQS items) as the most relevant predictors of distress after sessions: v89 (therapist helps the patient avoid or suppress disturbing content); v67 (therapist draws the patient's attention to unconscious content); v92 (patient's feelings or perceptions are linked to situations or behaviors of the past); v80 (therapist presents an experience from a different perspective); v46 (therapist communicates clearly); and v100 (therapist draws connections between the therapeutic relationship and other relationships). Figure 1 presents the variables and their importance in the model.

The random forest algorithm achieved high accuracy and high clinical validity. It should be noted that the resultant model was predominantly composed of specific factors of the employed approach, and was congruent with psychoanalytic theory. Identifying those variables traditionally associated with psychoanalytic treatment that were active during the sessions corroborates the validity of psychoanalytic constructs in clinical settings. For example, the most important construct in the model was item 89, which postulates that the therapist intervenes to help the patient avoid or suppress disturbing ideas or feelings. This means that psychoanalytic work regarding the patient's defenses against emotional experience is fundamental within the process, and is closely associated with the variability of the patient's mental state regarding treatment with the proposed model. Next in importance, after item 89, were item 67 (which postulates that the therapist draws the patient's attention to wishes, feelings,




**Figure 1** Variable importance in the model.

or ideas that may not be in awareness) and item 92 (which denotes that several links or salient connections are made between the patient's current emotional experience and the perception of events in the past). Both items refer to central issues in psychoanalytic theory and technique - namely, unconscious content urging toward consciousness, and the importance of experiences during childhood. In brief, the proposed model clearly represents the psychoanalytic factor in action (represented by variables belonging mainly to the set called "specific factors" in psychotherapy). This means that the set of interventions proposed, when used concomitantly, had an effect on the patient's mental state, and thus proved effective.

Some limitations must be taken into account. Our findings correspond only to the case in question, and do not allow for generalizations. Also, the number of observations ( $n=62$ ) may be considered small. Although we used LOOCV, we did not have an external dataset to test our signature. Therefore, our findings may be prone to overfitting.

In conclusion, this pilot study showed that ML approaches could be a useful tool to promote advances in psychotherapeutic processes, allowing the study of numerous simultaneous variables without assuming linearity. This seems to be a promising field of study, with room for improvement. Therapeutic interaction is a complex phenomenon, and the debate as to the effective ingredients that promote therapeutic change is central.

Pricilla B. Laskoski,<sup>1</sup>  Fernanda B. Serralta,<sup>2</sup>

Ives C. Passos,<sup>1,3</sup> Simone Hauck<sup>1,3</sup> 

<sup>1</sup>Programa de Pós-Graduação em Psiquiatria e Ciências do Comportamento, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Brazil. <sup>2</sup>Programa de Pós-Graduação em Psicologia, Universidade do Vale do Rio dos Sinos (UNISINOS), São Leopoldo, RS, Brazil. <sup>3</sup>Hospital de Clínicas de Porto Alegre, Porto Alegre, RS, Brazil

Submitted Oct 02 2018, accepted Aug 08 2019.

## Acknowledgements

This research was funded by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). The funding source did not play any role in the collection, analysis, interpretation of the data, manuscript writing, or the decision to submit for publication.

## Disclosure

The authors report no conflicts of interest.

**How to cite this article:** Laskoski PB, Serralta FB, Passos IC, Hauck S. Machine-learning approaches in psychotherapy: a promising tool for advancing the understanding of the psychotherapeutic process. *Braz J Psychiatry*. 2019; 41:568-569. <http://dx.doi.org/10.1590/1516-4446-2018-0295>

## References

- 1 Wu MJ, Mwangi B, Bauer IE, Passos IC, Sanches M, Zunta-Soares GB, et al. Identification and individualized prediction of clinical phenotypes in bipolar disorders using neurocognitive data, neuroimaging scans and machine learning. *Neuroimage*. 2017;145:254-64.
- 2 Pinto JV, Passos IC, Gomes F, Reckziegel R, Kapczinski F, Mwangi B, et al. Peripheral biomarker signatures of bipolar disorder and schizophrenia: a machine learning approach. *Schizophr Res*. 2017;188:182-4.
- 3 Passos IC, Mwangi B, Cao B, Hamilton JE, Wu MJ, Zhang XY, 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;193:109-16.
- 4 Månsson KN, Frick A, Boraxbekk CJ, Marquand AF, Williams SCR, Carlbring P, et al. Predicting long-term outcome of Internet-delivered cognitive behavior therapy for social anxiety disorder using fMRI and support vector machine learning. *Transl Psychiatry*. 2015;5:e530.
- 5 Hahn T, Kircher T, Straube B, Wittchen HU, Konrad C, Ströhle A, et al. Predicting treatment response to cognitive behavioral therapy in panic disorder with agoraphobia by integrating local neural information. *JAMA Psychiatry*. 2015;72:68-74.
- 6 Lantz B. *Machine learning with R - Second Edition*. 2nd ed. Birmingham: Packt Publishing; 2015.
- 7 Breiman L. Random forests. *Mach Learn*. 2001;45:5-32.