Article

Development and Internal Evaluation of an Interpretable Psychometric Classification Algorithm for Mental Disorders and Suicide Risk in Psychiatric Inpatients

Diego Diaz-Milanes[1,2]
Ana Garcia-Megias[1,*]
Carmen Santin Vilariño[3]
Raquel Remesal-Cobreros[3,4,*]

[1]Department of Quantitative Methods, Universidad Loyola Andalucía, 41704 Dos Hermanas, Seville, Spain
[2]Health Research Institute, University of Canberra, 2617 Bruce, Australian Capital Territory, Australia
[3]Department of Clinical and Experimental Psychology, University of Huelva, 21007 Huelva, Huelva, Spain
[4]Mental Health Inpatient Unit, Juan Ramón Jiménez Hospital, Andalusian Health Service, 21005 Huelva, Huelva, Spain

Abstract

Background: Mental disorders (MDs) and suicide risk represent major challenges for mental healthcare. Although several psychometric instruments are available for clinical assessment, their isolated use may provide limited support for differentiating clinically relevant profiles. Interpretable machine-learning approaches may help integrate psychometric information and support preliminary classification in mental health settings. This study aimed to develop and internally evaluate a preliminary, interpretable classification algorithm based on psychometric measures to distinguish among individuals without associated psychopathology, psychiatric inpatients with MD without suicide risk, and psychiatric inpatients with MD and suicide risk.

Methods: The sample included 175 participants aged 18–76 years, distributed into three groups: control group (n = 58), MD without suicide risk group (n = 59), and MD with suicide risk group (n = 58). Clinical participants were recruited from an acute psychiatric inpatient unit. A Random Forest model was used for feature selection, and the selected predictors were entered into a Classification and Regression Trees (CART) decision tree using the Gini index as the splitting criterion. Model performance was internally evaluated using a 70/30 train-test split, with performance assessed globally and for each outcome category. Performance metrics included overall accuracy, Cohen's kappa, area under the receiver operating characteristic curve (AUC-ROC), sensitivity, specificity, balanced accuracy, precision, and F1-score.

Results: The Random Forest procedure identified depression, psychoticism, openness to experience, anxiety, and state anger as the most relevant predictors. The final CART model retained three variables: depression, openness to experience, and psychoticism. Overall accuracy was 0.68, and Cohen's kappa was 0.53, indicating moderate classification agreement. Class-specific AUC-ROC values were 0.78 for the control group, 0.65 for the MD without suicide risk group, and 0.83 for the MD with suicide risk group. Sensitivity was 0.71, 0.50, and 0.83, respectively, and specificity was 0.83, 0.92, and 0.77, respectively.

Conclusions: The proposed decision tree provides a preliminary and interpretable psychometric classification model for differentiating individuals without associated psychopathology from acute psychiatric inpatients with MD with or without suicide risk. The model showed moderate overall classification performance, with better discrimination for suicide risk than for MD without suicide risk. These findings represent preliminary evidence of classification performance within the study population and should not be interpreted as evidence of clinical utility or immedi-

Submitted: 1 July 2026 Revised: 23 August 2026 Accepted: 10 September 2026 Online First: 25 November 2026
*Corresponding author details: Ana Garcia-Megias, Department of Quantitative Methods, Universidad Loyola Andalucía, 41704 Dos Hermanas, Seville, Spain. Email: agarciamegias@al.uloyola.es; Raquel Remesal-Cobreros, Department of Clinical and Experimental Psychology, University of Huelva, 21007 Huelva, Huelva, Spain; Mental Health Inpatient Unit, Juan Ramón Jiménez Hospital, Andalusian Health Service, 21005 Huelva, Huelva, Spain. Email: mariar.remesal.sspa@juntadeandalucia.es

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

ate implementation. External evaluation in larger and more diverse samples, including primary-care and community populations, is required before considering broader clinical application.

Keywords

suicide; mental disorders; machine learning; decision trees; psychometrics

Introduction

Mental health constitutes one of the main healthcare challenges of the 21st century, owing to the high prevalence of mental disorders (MDs), their associated disability burden, and their relationship with premature mortality [1]. More than 970 million people worldwide live with a mental disorder, and mental, neurological, and substance use disorders are among the leading contributors to disability-adjusted life years (DALYs) [2–4]. Suicide represents one of the most severe outcomes associated with mental distress and psychiatric morbidity, accounting for more than 700,000 deaths each year and at least 20 suicide attempts for every completed suicide [5]. Despite this burden, mental healthcare continues to face persistent structural and clinical challenges, including shortages of specialised professionals, barriers to access, stigma, and insufficient implementation of prevention and treatment programmes [6,7]. These difficulties contribute not only to a treatment gap, but also to a quality gap, as many interventions are not delivered according to guideline-based standards or adequately targeted to those most in need [7,8]. In this context, early identification and risk stratification are central to improving detection, referral, monitoring, and suicide prevention strategies.

However, identifying MDs and suicide risk in routine practice remains complex. Clinical presentations are often heterogeneous, overlapping, and influenced by emotional, cognitive, personality-related, social, and contextual factors. Moreover, risk factors are frequently examined in isolation, limiting understanding of the interactions that may underlie psychopathology or suicidal behaviour [9,10]. Although validated psychometric instruments are available, their clinical use is not always straightforward. Some screening tools show high sensitivity but limited specificity, which may increase false positives, overburden services, and contribute to unnecessary referrals or hospitalisations [11,12]. In addition, the proliferation of partially redundant standardised measures may hinder comparability across studies and complicate the accumulation of cumulative evidence [13].

Machine learning (ML) and artificial intelligence (AI) have therefore received increasing attention as tools to support prediction, risk stratification, and clinical decision-making in mental health. These approaches are particularly relevant when clinical phenomena are multidimensional and non-linear, as they can integrate heterogeneous information and identify patterns that may not be readily apparent using conventional analytic strategies [14]. Recent reviews indicate that ML applications in psychiatry and clinical psychology have expanded rapidly, particularly in detection and diagnosis, prognosis, treatment planning, public health surveillance, and individualised care, using clinical, psychometric, biological, behavioural, and administrative data [15]. This rapid growth has also increased the need for transparent and reproducible reporting standards for clinical prediction models using machine-learning methods, as reflected in the TRIPOD+AI statement [16].

Previous studies have applied ML and related prediction methods across a broad range of mental health conditions and outcomes, with particularly frequent applications in depression, schizophrenia-spectrum disorders, anxiety-related conditions, bipolar disorder, neurodevelopmental conditions and trauma-related disorders [17–20]. Suicide-related prediction has also become a major area of development, particularly in studies examining suicidal ideation, suicide attempts, and suicide deaths in psychiatric and healthcare populations [21,22]. Fernandez *et al*. [23] developed and validated a prediction algorithm for the onset of common MDs, mainly depressive and anxiety-related conditions, in a working population, supporting the value of predictive modelling for preventive mental health strategies. Other studies have used electronic health records and routinely collected clinical data to predict suicide attempts and suicide deaths after healthcare visits [24–26]. In addition, models combining clinician assessment, patient self-report, and electronic health records suggest that patient-reported psychological information may add value beyond routine clinical judgement or administrative data alone [27].

Despite these advances, important methodological and translational limitations remain. Reviews have highlighted heterogeneity in data sources, populations, outcomes, modelling strategies, and validation procedures, as well as concerns about reporting transparency and overfitting [17–20]. In suicide-risk prediction and precision-psychiatry research, recurrent limitations include small or highly selected samples, limited external validation, and insufficient evidence for clinical implementation [22,28–30]. Many existing models also focus on a single disorder, clinical outcome, or risk event, and often rely primarily on

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

electronic health records, administrative data, digital traces, or highly specialised data sources. Fewer studies have examined whether validated psychometric dimensions can be combined into an interpretable classification algorithm capable of distinguishing among individuals without associated psychopathology, psychiatric patients without suicide risk, and psychiatric patients with suicide risk. This distinction is clinically relevant because these profiles may require different levels of clinical attention, referral, monitoring, and intervention. In this context, transparent approaches such as decision trees may be particularly useful, as clinical value depends not only on predictive accuracy but also on interpretability, feasibility, reproducibility, and the capacity to support multidisciplinary decision-making [15,29–31].

Accordingly, the present study aimed to develop and internally evaluate a preliminary, interpretable classification algorithm based on psychometric measures and socioeconomic characteristics to distinguish among individuals without associated psychopathology, psychiatric inpatients with MD without suicide risk, and psychiatric inpatients with MD and suicide risk.

Materials and Methods

Sample and Procedure

The sample comprised a total of 175 participants, aged between 18 and 76 years (Mean (M) = 40.02; standard deviation (SD) = 12.62). The sex distribution was balanced, with 86 men (49.14%) and 89 women (50.86%).

Participants were distributed into three groups: a control group, comprising individuals without associated psychopathology and used as the comparator group (n = 58; 33.14%); a MD without suicide risk group, comprising patients with different MDs admitted to the Mental Health Inpatient Unit of Juan Ramón Jiménez Hospital in Huelva (Spain) for reasons other than a suicide attempt and without persistent suicide risk (n = 59; 33.71%); and an MD with suicide risk group, comprising psychiatric inpatients who had been admitted following a recent suicide attempt and who continued to be considered at suicide risk during the emergency assessment (n = 58; 33.14%). In this latter group, persistent suicide risk was defined by the absence of sufficient critical reappraisal or distancing from the suicidal act during the emergency assessment, which contributed to the clinical indication for psychiatric hospitalisation. A suicide attempt was required for inclusion in this group; suicidal ideation alone was insufficient. Suicide risk was assessed by at least two mental health specialists, including clinical psychologists and/or psychiatrists, through a multi-source clinical evaluation combining semi-structured interviews, clinical observation, medical-record review, and, when available, information from relatives or other key informants.

Data collection was conducted from October 2012 to March 2014. The original sample-size calculation was based on the three-group comparative design of the study. Assuming a medium effect size (f = 0.25), an alpha level of 0.05, statistical power of 0.80, and three approximately equally sized groups, a minimum total sample of approximately 159 participants (53 per group) was estimated. The final sample included 175 participants. Recruitment feasibility was additionally informed by historical activity in the acute psychiatric unit, which recorded approximately 700 admissions annually, of which approximately 6.5% were admissions following suicide attempts. Given the relative infrequency of this clinical presentation, eligible patients admitted following a suicide attempt and presenting persistent suicide risk were recruited consecutively, while recruitment of the remaining groups aimed to maintain approximately balanced group sizes. The original sample-size calculation was not specifically designed for the subsequent machine-learning analysis.

Psychometric assessment was not conducted during the acute phase immediately following admission. In both clinical groups, assessment was performed once sufficient clinical stabilisation had been achieved to allow participants to understand the assessment procedure, collaborate adequately, and provide reliable self-report data. Patients who did not meet these conditions were excluded. Among patients admitted following a suicide attempt, the interval between the attempt or initial suicide-risk assessment and psychometric testing varied according to the time required to achieve clinical stabilisation and was therefore not predefined. The same stabilisation criterion was applied to patients admitted for other psychiatric reasons. The mean length of stay in the acute psychiatric unit was approximately 12–15 days.

Among the clinical participants only (n = 117), diagnostic categories were classified according to the International Classification of Diseases, 10th Revision (ICD-10). Diagnoses were established through the same multi-source clinical assessment described above. The most frequent diagnostic categories were schizophrenia, schizotypal and delusional disorders (n = 37; 31.6%), mood disorders (n = 36; 30.8%), personality disorders (n = 25; 21.4%), neurotic, stress-related and somatoform disorders (n = 15; 12.8%), and mental and behavioural disorders due to substance use (n = 11; 9.4%). Behavioural syndromes associated with

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

physiological disturbances and physical factors were observed in one participant (n = 1; 0.9%). No cases of organic MDs, intellectual disability, or disorders of psychological development were recorded. Diagnostic categories were not mutually exclusive.

The diagnostic distribution differed between the two clinical groups. In the MD without suicide risk group, the most frequent diagnosis was schizophrenia, schizotypal and delusional disorders (n = 29; 49.2%), followed by mood disorders (n = 13; 22.0%), mental and behavioural disorders due to substance use (n = 8; 13.6%), personality disorders (n = 6; 10.2%), and neurotic, stress-related and somatoform disorders (n = 4; 6.8%). In the MD with suicide risk group, the most frequent diagnoses were mood disorders (n = 23; 39.7%), personality disorders (n = 19; 32.8%), neurotic, stress-related and somatoform disorders (n = 11; 19.0%), schizophrenia, schizotypal and delusional disorders (n = 8; 13.8%), and mental and behavioural disorders due to substance use (n = 3; 5.2%). Regarding diagnostic complexity, most clinical participants had a single recorded diagnosis (n = 109; 93.2%), whereas comorbidity was identified in eight participants (n = 8; 6.8%).

Data were collected by a trained clinical psychologist from the acute psychiatric unit. The study was conducted in accordance with the ethical principles of the Declaration of Helsinki [32] and was reviewed and approved by Research Ethics Committee for Healthcare Centres in the Province of Huelva (Comité de Ética de Investigación de centros de la provincia de Huelva), Andalusian Health Service (Servicio Andaluz de Salud), Regional Ministry of Health, Regional Government of Andalusia (reference code: CI/PI-001/12). Participation was voluntary, and written informed consent was obtained from all participants prior to their inclusion in the study. The confidentiality of personal data was guaranteed. Data were stored in an anonymous database through the assignment of a registration number, preventing individual identification. Data custody, in both physical and digital formats, was the exclusive responsibility of the research team, in accordance with current data protection regulations.

Measures

Sociodemographic variables were collected using an ad hoc questionnaire designed for the present study. These variables included age, sex, marital status, education level, employment status, and living situation.

Several validated psychometric instruments were used to assess emotional, cognitive, behavioural, symptomatic, and personality-related variables relevant to mental health. The psychometric properties of the instruments are reported in **Supplementary Table 1**, and their scoring procedures are described in **Supplementary Table 2**. The measures were selected a priori based on evidence linking affective symptoms, psychopathology, hopelessness, self-esteem, coping, anger, and personality with MD and suicidal behaviour. These domains were analysed jointly to identify the most discriminative predictors.

The State–Trait Anger Expression Inventory-2 (STAXI-2) [33], in its Spanish adaptation [34] was used to assess the intensity, frequency, and expression of anger, both as a transient emotional state and as a dispositional trait. The instrument includes measures of state anger, trait anger, internal and external anger expression, and internal and external anger control.

The State–Trait Depression Inventory (IDER) [35], in its Spanish adaptation [36] was used to assess depression from both state and trait perspectives. The instrument distinguishes between the presence of negative affectivity, or dysthymia, and the absence of positive affectivity, or euthymia.

The Coping Responses Inventory–Adult Form (CRI-A) [37], in its Spanish adaptation [38] was used to assess cognitive and behavioural coping responses to stress. These responses are organised into two broad coping styles: approach coping, including strategies such as logical analysis and problem-solving, and avoidance coping, including strategies such as cognitive avoidance and emotional discharge.

The Symptom Checklist-90-Revised (SCL-90-R) [39], in its Spanish adaptation [40] was used to assess psychopathological symptomatology. The instrument evaluates nine symptom dimensions: somatisation, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychoticism.

The Beck Hopelessness Scale (BHS) [41], in its Spanish adaptation [42] was used to assess negative expectations about the future and perceived inability to cope with difficulties. The scale has shown high internal consistency, as well as adequate factorial and predictive validity.

The Rosenberg Self-Esteem Scale [43], in its Spanish adaptation [44] was used to assess global self-esteem. The scale comprises 10 Likert-type items reflecting positive and negative self-evaluation.

The NEO Five-Factor Inventory (NEO-FFI) [45], in

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

its Spanish adaptation [46] was used to assess the five major domains of personality: neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness. The instrument comprises 60 items rated on a 5-point Likert-type scale and provides scores for the five broad personality factors. In the present study, the five domain scores were used as candidate predictors in the ML analyses.

Data Analysis

Baseline characteristics were compared across the three study groups to characterise between-group comparability and identify potential sources of heterogeneity. Continuous variables were compared using one-way analysis of variance when distributional assumptions were considered adequate and variances were homogeneous, Welch's analysis of variance when variances were heterogeneous, and the Kruskal-Wallis test when parametric assumptions were not met. Categorical variables were compared using Pearson's chi-square test or Fisher's exact test with 10,000 Monte-Carlo replicates when expected cell counts were insufficient. Significant omnibus comparisons were followed by Tukey, Games-Howell, or Dunn post-hoc tests, as appropriate, or by examination of adjusted standardised residuals for categorical variables. *p*-values from multiple post-hoc comparisons were adjusted using the Benjamini-Hochberg procedure. Effect sizes were reported as omega-squared or epsilon-squared for continuous variables and Cramer's V for categorical variables.

Relevant psychiatric diagnostic characteristics were additionally described for the two clinical groups. Because psychiatric diagnoses were not applicable to the control group, comparisons of diagnostic categories were restricted to participants with MD. Diagnostic categories were not mutually exclusive and were compared between the two clinical groups using Fisher's exact tests, with Benjamini-Hochberg adjustment across the reported diagnostic categories. Two-sided $p < 0.05$ was considered statistically significant.

The modelling procedure was conducted in two stages. First, a Random Forest model was applied to the 46 candidate predictors included in the original analytical framework. For this feature-selection stage, the analytical sample was restricted to participants with complete data across the candidate predictors. Random Forest was selected because it can capture non-linear relationships and interactions among predictors without requiring strong distributional assumptions and provides variable-importance estimates that allow candidate predictors to be ranked. Its role was therefore to screen and reduce the initial predictor set rather than to serve as the final classification model. The Random Forest comprised 500 trees, with predictor importance quantified using permutation-based Mean Decrease in Accuracy. No score standardisation or class weighting was applied.

The five highest-ranked predictors were retained for subsequent Classification and Regression Trees (CART) modelling. The decision to retain five predictors was based on a parsimonious heuristic considered appropriate given the modest sample size and was also intended to reduce the loss of participants associated with complete-case analysis across the original 46-variable predictor set. Following feature selection, the analytical dataset was reconstructed using only these five predictors. Because substantially fewer variables were required at this stage, a larger number of participants had complete data and could therefore be retained for the final classification analysis.

The resulting reduced-predictor dataset was then divided into a training set comprising 70% of the observations and a test set comprising the remaining 30%. Multiclass classification was subsequently performed using the CART algorithm [31]. CART was selected as the final modelling approach because it generates explicit hierarchical decision rules and data-driven cut-off points, facilitating interpretation of the resulting classification structure. The CART model was fitted using the training subset and subsequently evaluated in the test subset. Thus, the test subset was not used to fit the CART model, although Random Forest feature selection had been performed before the final train-test split.

The Gini index was used as the splitting criterion. The model used the default rpart control parameters (complexity parameter = 0.01, minimum split size = 20, minimum terminal-node size = 7, and maximum tree depth = 30), and no additional post-hoc pruning procedure was applied. The cut-off points generated by the tree represented data-driven CART split values and were not interpreted as established clinical thresholds.

Model performance in the test subset was assessed using overall accuracy and Cohen's kappa coefficient. In addition, class-specific metrics were calculated, including the area under the receiver operating characteristic curve (AUC-ROC), sensitivity, specificity, balanced accuracy, precision, and F1-score, enabling a more detailed evaluation of classification performance.

Missing observations were handled using complete-case analysis for the variables required at each modelling stage. Accordingly, complete cases across the full set of

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

candidate predictors were used during Random Forest feature selection, whereas the analytical sample for CART development and evaluation was reconstructed on the basis of complete data for the five retained predictors. This stage-specific approach allowed a larger effective sample to be used for the final classification model after dimensionality reduction.

All analyses were conducted in R (version 4.2.3; R Foundation for Statistical Computing, Vienna, Austria), using the following packages: haven (2.5.4), dplyr (1.1.4), ggplot2 (3.5.1), rpart (4.1.23), rpart.plot (3.1.1), pROC (1.18.5), caret (6.0.94), randomForest (4.7.1.1), and nnet (7.3.19).

The study was reported in accordance with the TRIPOD+AI recommendations for reporting studies that develop and evaluate clinical prediction models using regression or machine-learning methods [16]. The completed TRIPOD+AI checklist is provided in **Supplementary Table 3**.

Results

Baseline Characteristics

Baseline characteristics differed significantly across the three study groups for age ($p = 0.016$, $\omega^2 = 0.04$), sex ($p < 0.001$, Cramer's V = 0.29), marital status ($p < 0.001$, Cramer's V = 0.34), education level ($p = 0.002$, Cramer's V = 0.30), employment status ($p < 0.001$, Cramer's V = 0.40), and living situation ($p < 0.001$, Cramer's V = 0.36). Post-hoc analyses showed that participants in the MD without suicide risk group were younger than those in both the Control and MD with suicide risk groups. Additional post-hoc differences were observed across sociodemographic categories, particularly for sex, marital status, employment status, and living situation. Full descriptive statistics and between-group comparisons are provided in **Supplementary Table 4**, and multiplicity-adjusted post-hoc comparisons are reported in **Supplementary Table 5**.

Among the two clinical groups, the distribution of psychiatric diagnoses also differed for specific diagnostic categories. Schizophrenia, schizotypal, and delusional disorders were more frequent in the MD without suicide risk group than in the MD with suicide risk group (49.2% vs. 13.8%; Benjamini-Hochberg-adjusted $p < 0.001$, Cramer's V = 0.38), whereas personality disorders were more frequent in the MD with suicide risk group (32.8% vs. 10.2%; adjusted $p = 0.008$, Cramer's V = 0.28). Differences in mood disorders, neurotic, stress-related and somatoform disorders, and mental and behavioural disorders due to substance use did not remain statistically significant after correction for multiple comparisons. Detailed diagnostic distributions are presented in **Supplementary Table 4**.

Feature Selection

To optimise the model's classification capacity, a random forest-based feature selection procedure was applied. This analysis ranked the variables according to their relative importance, identifying depression (12.56), psychoticism (11.67), openness to experience (9.59), anxiety (8.83), and momentary feelings of anger (8.16) as the most relevant predictors (see Fig. 1). No socioeconomic variable was selected by the procedure. These variables were subsequently used to construct the decision tree model, providing the basis for exploring hierarchical and non-linear relationships among the groups analysed.

Decision Tree

Among the previously selected attributes, the decision tree model identified the predictors that provided the most parsimonious solution with the greatest discriminative capacity between groups, using three variables: depression, openness to experience, and psychoticism (Fig. 2). The initial splitting node corresponded to depression, with a cut-off point of 1.3. Participants with scores below this value who also showed scores equal to or greater than 33 in openness to experience were mostly classified into the control group (41% of the sample). In contrast, those with low levels of depression and low openness to experience were classified into the MD without suicide risk group (13%).

Among participants with depression scores equal to or greater than 1.3, a second split was observed at a cut-off point of 2.7. Those with depression scores equal to or greater than this value was consistently classified into the MD with suicide risk group (20%). Finally, participants with depression scores below 2.7 were classified according to their level of psychoticism: scores equal to or greater than 1.6 corresponded to the MD without suicide risk group (10%), whereas lower scores were associated with the MD with suicide risk group (16%).

Overall, these results show that depression constituted the main splitting node of the model, followed by openness to experience and psychoticism, forming a hierarchical pattern that made it possible to differentiate between individuals without associated psychopathology, patients belonging to the MD without suicide risk group, and patients in the

Actas Esp Psiquiatr 2026;1–14. https://doi.org/10.62641/aep.2447 | ISSN:1578-2735
© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

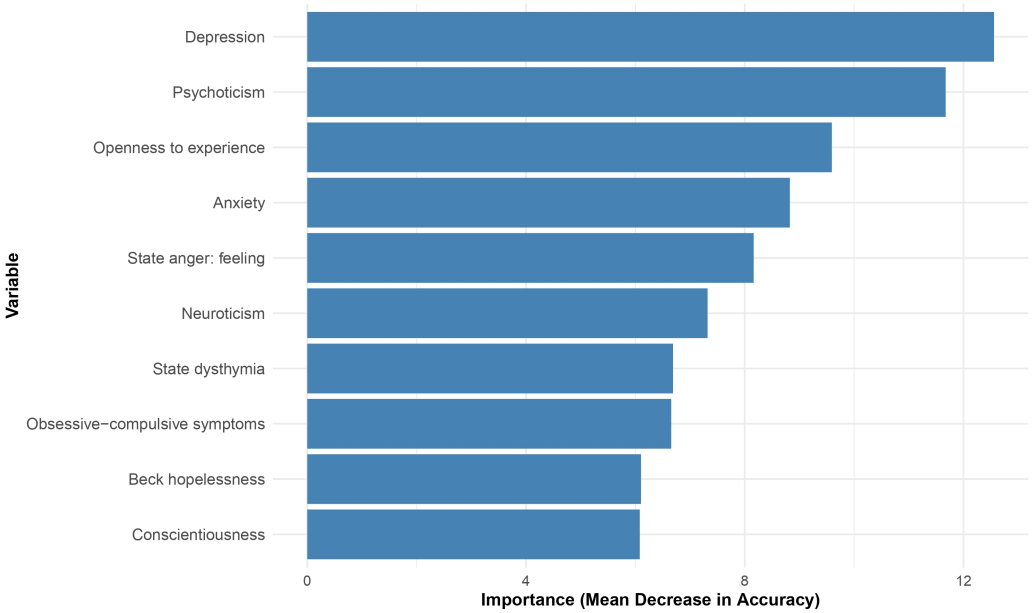


Fig. 1. Top 10 most important variables identified.

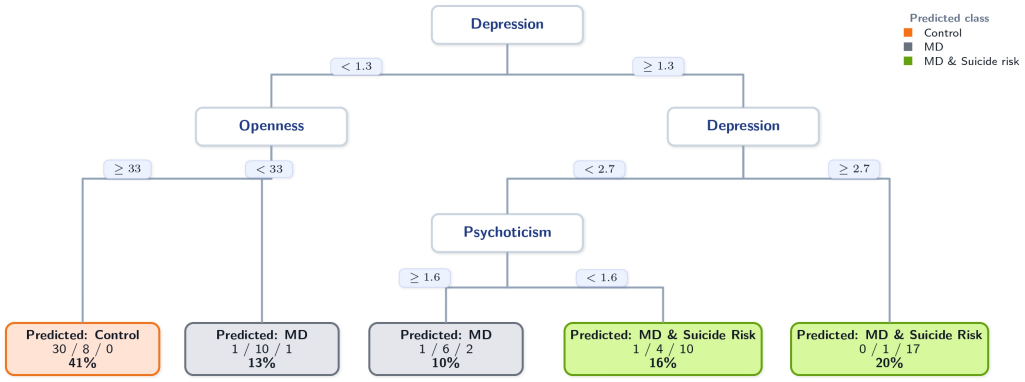


Fig. 2. Best-fitting decision tree. Note. Terminal nodes show observed counts in the order: Control/MD/MD & suicide risk. Thresholds are CART split points. MD, mental disorder.

MD with suicide risk group.

Performance Assessment

Overall, the model achieved an accuracy of 68.4% and a Kappa coefficient of 0.53, indicating moderate agreement in participant classification beyond chance.

The performance of the decision tree model by category was evaluated in the test sample using different classification metrics (see Table 1). Fig. 3 presents the multiclass ROC curve, in which the AUC values showed adequate performance in classifying the groups: control (AUC = 0.78), MD without suicide risk (AUC = 0.65), and MD with suicide risk (AUC = 0.83).

In terms of sensitivity, the model reached a value of 0.71 for the control group, 0.50 for the MD without suicide risk group, and 0.83 for the MD with suicide risk group. Specificity was high for the MD without suicide risk group (0.92) and moderate for the control (0.83) and MD with suicide risk (0.77) groups (Table 1).

Balanced accuracy ranged from 0.71 for the MD without suicide risk group to 0.80 for the MD with suicide risk group. Regarding precision and F1-score, the control group obtained values of 0.71 and 0.71, respectively; the MD without suicide risk group obtained values of 0.75 and 0.60; and the MD with suicide risk group obtained values of 0.63

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

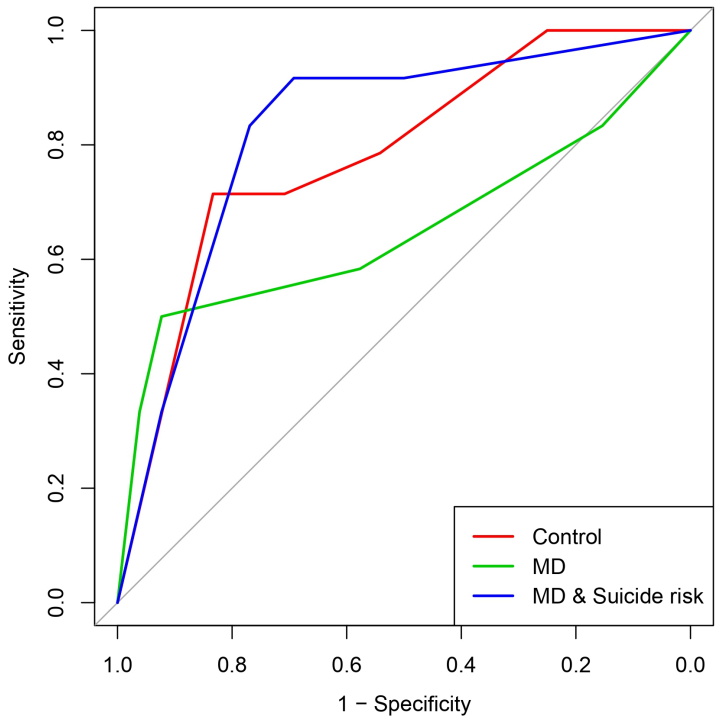


Fig. 3. Multiclass ROC curves for the decision tree model. MD, mental disorder.

and 0.71.

Discussion

The aim of this study was to develop and internally evaluate a preliminary, interpretable classification algorithm based on psychometric measures to distinguish among individuals without associated psychopathology, psychiatric inpatients with MD without suicide risk, and psychiatric inpatients with MD and suicide risk. Overall, the findings suggest that a parsimonious CART decision tree based on depressive symptomatology, openness to experience, and psychoticism provides meaningful, although moderate, discrimination between these three profiles. The model showed an overall accuracy of 0.68 and a Cohen's kappa of 0.53, with better discrimination for the suicide-risk group than for the MD without suicide risk group. Thus, the main value of the proposed model lies in its psychometric basis, parsimony, and interpretability, rather than in immediate clinical applicability.

The present algorithm should be interpreted in relation to previous machine-learning models in mental health, while recognising important differences in population, data sources, outcomes, and modelling aims. Some previous models have focused on predicting the onset of common MDs in working populations [23], whereas others have addressed disorder-specific classification, symptom severity, treatment response, or suicide-related outcomes. In the field of suicide-risk prediction, several large-scale models have used electronic health records, administrative data, routinely collected clinical information, or combinations of clinician assessment and patient self-report to predict suicide attempts, suicide deaths, or other future suicide-related outcomes [24–27]. In contrast, the present study did not predict future suicidal behaviour or incident mental disorder, but examined whether validated psychometric dimensions could classify current clinical profiles in an acute psychiatric inpatient sample. Therefore, the proposed model should not be viewed as directly competing with large-scale prognostic algorithms, but as a distinct psychometric and interpretable classification approach.

This distinction is relevant because many existing machine-learning models in mental health rely on data sources that differ substantially from the psychometric measures used in the present study, including electronic health records, neuroimaging, biological markers, social media data, smartphone-based passive sensing, wearable sensors, and other digital traces [47,48]. By contrast, the present algorithm was derived from validated symptom and per-

Actas Esp Psiquiatr 2026;1–14. https://doi.org/10.62641/aep.2447 | ISSN:1578-2735
© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

Table 1. Performance metrics in the test sample.

Group	AUC-ROC	Sensitivity	Specificity	Balanced accuracy	Precision	F1-Score
Control	0.784	0.714	0.833	0.774	0.714	0.714
MD without suicide risk	0.651	0.5	0.923	0.712	0.75	0.6
MD with suicide risk	0.825	0.833	0.769	0.801	0.625	0.714

Note. MD, mental disorder; AUC-ROC, area under the receiver operating characteristic curve. Sensitivity refers to the proportion of correctly identified cases within each group, whereas specificity refers to the proportion of correctly identified non-cases. Balanced accuracy was calculated as the average of sensitivity and specificity. Precision represents the proportion of correctly classified positive predictions, and the F1-score represents the harmonic mean of precision and sensitivity.

sonality measures routinely used in psychological and psychiatric assessment. This approach has potential conceptual and methodological value because it explores whether clinically interpretable psychometric dimensions can be reorganised into transparent decision rules. In this respect, the study extends previous work by examining a multigroup classification task that differentiates absence of associated psychopathology, MD without suicide risk, and MD with suicide risk within an inpatient context.

The use of a decision tree is also relevant from a methodological and interpretative perspective. In contrast to more complex ML approaches, classification trees provide explicit decision rules based on successive splits, making the classification process transparent and easily inspectable. Decision trees are non-parametric models that can capture non-linear relationships and clinically meaningful subgroups without requiring strong distributional assumptions [49]. Their rule-based structure may therefore be useful for understanding how specific psychometric profiles contribute to classification. However, this transparency should be understood as a strength for interpretation and hypothesis generation, rather than as evidence of clinical robustness. Simpler tree structures can be sensitive to sample-specific splits, particularly in modest datasets, and therefore require replication and external validation before applied use can be considered. This interpretation is consistent with current recommendations on transparent reporting and responsible development of machine-learning models in healthcare and mental health [16,50].

The central role of depressive symptomatology in the decision tree is clinically meaningful. Depression was the first splitting variable and therefore represented the main axis of differentiation between the three groups. However, this variable should be interpreted as the depression dimension of the SCL-90-R, rather than as a formal diagnosis of depressive disorder. This dimension captures a broad profile of depressive distress, including dysphoric mood, low motivation, reduced energy, hopelessness, and suicidal ideation [39,40]. Its relevance in the model is consistent with the broader evidence showing that depressive disorders and depressive distress are major contributors to global mental health burden and are closely linked to clinical severity and suicide-related outcomes [4,51,52]. Nevertheless, depression should not be interpreted as a sufficient marker of suicide risk. Suicide risk is a multidimensional phenomenon involving clinical, psychological, interpersonal, contextual, and temporal factors, and cannot be reduced to depressive symptoms alone [9,53,54].

Beyond depressive symptomatology, openness to experience and psychoticism contributed additional discriminative information to the model. Openness to experience, assessed as one of the broad personality dimensions of the NEO Five-Factor Inventory (NEO-FFI), may reflect differences in cognitive flexibility, tolerance of uncertainty, and engagement with internal and external experiences [45,46]. In the final tree, this dimension contributed to distinguishing individuals without associated psychopathology from clinical participants under conditions of low depressive symptomatology. Its relationship with suicide risk appears to be more complex and context-dependent than that observed for other personality traits, such as neuroticism, which has been consistently associated with increased risk. In young populations, higher scores on openness to experience have been linked to a greater likelihood of suicidal ideation and self-harm [55,56]. However, in an adult sample, Flint *et al*. [57] found that, in situations of acute crisis, higher levels of openness may act as a protective factor against the short-term transition to suicidal behaviour. This protective function may be explained by greater cognitive flexibility and a greater capacity to generate alternative responses to emotional suffering, in contrast to the rigid and constricted patterns of thinking that typically characterize suicidal crises.

Psychoticism, assessed through the SCL-90-R, should also be interpreted carefully. In this context, psychoticism refers to a symptomatic dimension related to interpersonal withdrawal, unusual perceptual or cognitive experiences, and psychological disconnection, rather than to a categor-

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

ical diagnosis of psychotic disorder [39,40]. Importantly, its role in the tree should not be understood as indicating a protective factor against suicide risk. Rather, psychoticism appears to have contributed to differentiating diagnostic and symptomatic profiles within the acute inpatient sample. This interpretation is especially relevant because schizophrenia-spectrum disorders were more strongly represented in the MD without suicide risk group, whereas mood disorders, personality disorders, and neurotic, stress-related, and somatoform disorders were more frequent in the suicide-risk group. Thus, psychoticism may partly reflect the diagnostic composition of the sample, rather than specifically discriminating suicide risk status.

The moderate classification performance may reflect the clinical heterogeneity and overlap between the study groups. Both clinical groups included patients with different psychiatric diagnoses and therefore potentially overlapping psychometric profiles. This may have particularly affected the MD without suicide risk group, which showed the lowest discrimination (AUC = 0.65) and sensitivity (0.50). These patients shared clinically relevant psychopathology with the suicide-risk group but lacked the suicide-related characteristic that differentiated the latter, placing them in a less clearly separable intermediate profile. By contrast, the suicide-risk group was defined by a recent suicide attempt together with persistent suicide risk at admission, which may have contributed to its greater discriminability. The modest sample size and internal train-test evaluation may also have increased variability in the estimated model performance.

The proposed algorithm also has relevant clinical implications for future research on efficient psychometric assessment. The complete battery used in the study included a broad set of psychometric instruments, whereas the final decision tree retained only three dimensions: depressive symptomatology and psychoticism from the SCL-90-R, and openness to experience from the NEO-FFI. This suggests that a limited set of validated dimensions may contain relevant classificatory information. However, this finding should not be interpreted as validation of a shortened assessment battery. The original instruments were administered in their validated format, and the present analysis only indicates that these three dimensions were retained by the classification model. Future studies should examine whether these dimensions, or selected item-level components, can be administered independently while preserving reliability, validity, and classification performance. In this sense, the potential reduction of assessment burden is a promising implication for future research rather than a conclusion established by the present study.

Another relevant finding is that the feature-selection procedure did not retain any sociodemographic variable among the most important predictors, despite statistically significant baseline differences across groups in several sociodemographic characteristics. This result should therefore not be interpreted as indicating that the groups were sociodemographically homogeneous, or that sociodemographic and contextual factors are irrelevant to mental health or suicide risk. Rather, it suggests that, within this specific classification task and dataset, symptom and personality dimensions provided greater discriminative information than the sociodemographic variables included in the model. Social determinants, age, sex, socioeconomic adversity, social exclusion, and access to care remain central to understanding mental health inequalities, suicidal behaviours, and care pathways [2,52,58].

Several limitations should be considered when interpreting these findings. First, although the achieved sample size exceeded the minimum estimated for the original three-group design, it was not specifically calculated for prediction modelling [59]. Given the modest sample size and the three-class outcome, model stability may be limited and the risk of overfitting increased. Accordingly, the findings should be considered preliminary and require evaluation in larger independent samples. Second, participants in the clinical groups were recruited from an acute psychiatric inpatient unit and therefore represented patients with established and clinically severe presentations. Although psychometric assessment was conducted only after sufficient clinical stabilisation had been achieved to allow reliable self-reporting, patients were assessed within the context of routine inpatient treatment. Consequently, potential effects of pharmacological or other therapeutic interventions received during hospitalisation on symptom scores, affective state, cognition, or questionnaire responses cannot be excluded. In addition, the time between admission, suicide attempt where applicable, and psychometric assessment was not fixed but depended on the individual time required to achieve clinical stabilisation. Third, because the clinical sample consisted of acute inpatients, the model may not adequately represent earlier, subthreshold, or less severe presentations commonly encountered in primary or community care. Finally, model performance was evaluated internally using a training–test split, and the model has not yet been evaluated in an independent external sample. Therefore, the findings should be interpreted as preliminary evidence of classification performance within the study population rather than as evidence of immediate clinical applicability.

Future research should extend and refine the present findings in several directions. First, the algorithm should be externally validated in larger, independent, and diag-

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

nostically diverse samples. Multicentre studies would allow the stability, calibration, and generalisability of the decision rules to be assessed across healthcare contexts and diagnostic profiles. Second, future work should examine whether the model performs adequately in outpatient, primary care, and community samples, but only as part of formal validation studies. Third, prospective designs are needed to determine whether psychometric decision-tree models can predict future clinical outcomes, including changes in symptom severity, service use, suicide attempts, or other clinically relevant endpoints. Future studies should also include benchmark comparisons, calibration analyses, decision-curve analysis, and subgroup or fairness assessments. Finally, item-level analyses could explore whether shorter assessment formats can be developed without compromising predictive accuracy. In this regard, network analysis may be useful for examining interactions between symptoms, items, and instruments, providing a more fine-grained understanding of the variables involved in MD and suicide-risk profiles.

Conclusions

This study developed and internally evaluated a preliminary, interpretable psychometric classification model for distinguishing individuals without associated psychopathology from acute psychiatric inpatients with MD with or without suicide risk. The final decision tree achieved moderate overall classification performance, with stronger discrimination for the suicide-risk group than for the MD without suicide risk group. Given the modest sample size, the highly selected inpatient clinical population, and the absence of external evaluation, these findings should be interpreted as preliminary evidence of classification performance within the study population rather than as evidence of clinical utility. External evaluation in larger and more diverse samples, including primary-care and community populations, is required before considering its use in broader healthcare settings.

Availability of Data and Materials

The data that support the findings of this study are not publicly available due to ethical and privacy restrictions. An anonymised dataset may be made available from the corresponding authors upon reasonable request, subject to approval by the research team and compliance with applicable data protection regulations.

Author Contributions

DDM and RRC contributed to the conception and design of the study. Conceptualisation and methodology were led by DDM. Data acquisition and clinical supervision were conducted by RRC. Data curation, formal analysis, and visualisation, including the preparation of figures and tables, were performed by DDM and AGM. The original draft of the manuscript was prepared by DDM, AGM, and RRC. RRC and CSV contributed to clinical interpretation and manuscript review. All authors reviewed and edited the manuscript, approved the final version, and agreed to its submission. All authors agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

Ethics Approval and Consent to Participate

The study was reviewed and approved by the Research Ethics Committee for Healthcare Centres in the Province of Huelva (Comité de Ética de Investigación de centros de la provincia de Huelva), Andalusian Health Service (Servicio Andaluz de Salud), Regional Ministry of Health, Regional Government of Andalusia (reference code: CI/PI-001/12). This Committee was the competent ethics review body for research conducted at Juan Ramón Jiménez Hospital, where the clinical study was carried out. The study was conducted in accordance with the ethical principles of the Declaration of Helsinki. Written informed consent was obtained from all participants involved in the study.

Acknowledgment

The authors thank the acute psychiatric unit of Juan Ramón Jiménez Hospital in Huelva and the Andalusian Health Service for its support during data collection, as well as all participants who took part in the study. Artificial intelligence-assisted tools were used only to support language editing, translation, and stylistic refinement of the manuscript. The authors reviewed and verified all AI-assisted content and take full responsibility for the final version of the manuscript, including the accuracy, integrity, and interpretation of the scientific content.

Funding

This research received no external funding.

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.62641/aep.2447.

References

[1] World Health Organization. World mental health report: Transforming mental health for all. 2022. Available at: https://www.who.int/publications/i/item/9789240049338 (Accessed: 21 May 2026).

[2] Patel V, Chisholm D, Parikh R, Charlson FJ, Degenhardt L, Dua T, *et al*. Addressing the burden of mental, neurological, and substance use disorders: key messages from Disease Control Priorities, 3rd edition. Lancet. 2016; 387: 1672–1685. https://doi.org/10.1016/S0140-6736(15)00390-6.

[3] Prince M, Patel V, Saxena S, Maj M, Maselko J, Phillips MR, *et al*. No health without mental health. Lancet. 2007; 370: 859–877. https://doi.org/10.1016/S0140-6736(07)61238-0.

[4] COVID-19 Mental Disorders Collaborators. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. Lancet. 2021; 398: 1700–1712. https://doi.org/10.1016/S0140-6736(21)02143-7.

[5] World Health Organization. Suicide worldwide in 2019. 2021. Available at: http://who.int/publications/i/item/9789240026643 (Accessed: 21 May 2026).

[6] Kohn R, Saxena S, Levav I, Saraceno B. The treatment gap in mental health care. Bulletin of the World Health Organization. 2004; 82: 858–866.

[7] Thornicroft G. Most people with mental illness are not treated. Lancet. 2007; 370: 807–808. https://doi.org/10.1016/S0140-6736(07)61392-0.

[8] Lancet Global Mental Health Group. Scale up services for mental disorders: a call for action. Lancet. 2007; 370: 1241–1252. https://doi.org/10.1016/S0140-6736(07)61242-2.

[9] Franklin JC, Ribeiro JD, Fox KR, Bentley KH, Kleiman EM, Huang X, *et al*. Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. Psychological bulletin. 2017; 143: 187–232. https://doi.org/10.1037/bul0000084.

[10] Ribeiro JD, Huang X, Fox KR, Franklin JC. Depression and hopelessness as risk factors for suicide ideation, attempts and death: Meta-analysis of longitudinal studies. The British journal of psychiatry: the journal of mental science. 2018; 212: 279–286. https://doi.org/10.1192/bjp.2018.27.

[11] Ngin NLX, Hassan NB, Koh SLS. Predicting suicide and suicide attempts in adults in acute hospitals: A systematic review of diagnostic accuracy evaluating risk scales. International journal of nursing studies. 2022; 136: 104361. https://doi.org/10.1016/j.ijnurstu.2022.104361.

[12] Pollak OH, Sheehan AE, Walsh RFL, Stephenson AR, Zell H, Mayes J, *et al*. Assessment of suicidal thoughts and behaviors in adults: A systematic review of measure psychometric properties and implications for clinical and research utility. Clinical psychology review. 2024; 112: 102464. https://doi.org/10.1016/j.cpr.2024.102464.

[13] Elson M, Hussey I, Alsalti T, Arslan RC. Psychological measures aren't toothbrushes. Communications psychology. 2023; 1: 25. https://doi.org/10.1038/s44271-023-00026-9.

[14] Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. The New England journal of medicine. 2019; 380: 1347–1358. https://doi.org/10.1056/NEJMra1814259.

[15] Dwyer DB, Falkai P, Koutsouleris N. Machine Learning Approaches for Clinical Psychology and Psychiatry. Annual review of clinical psychology. 2018; 14: 91–118. https://doi.org/10.1146/annurev-clinpsy-032816-045037.

[16] Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, *et al*. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024; 385: e078378. https://doi.org/10.1136/bmj-2023-078378.

[17] Iyortsuun NK, Kim SH, Jhon M, Yang HJ, Pant S. A Review of Machine Learning and Deep Learning Approaches on Mental Health Diagnosis. Healthcare. 2023; 11: 285. https://doi.org/10.3390/healthcare11030285.

[18] Kannan KD, Jagatheesaperumal SK, Kandala RNVPS, Lotfaliany M, Alizadehsanid R, Mohebbi M. Advancements in Machine Learning and Deep Learning for Early Detection and Management of Mental Health Disorder. 2024. https://doi.org/10.48550/arXiv.2412.06147. (preprint)

[19] Sharma CM, Chariar VM. Diagnosis of mental disorders using machine learning: Literature review and bibliometric mapping from 2012 to 2023. Heliyon. 2024; 10: e32548. https://doi.org/10.1016/j.heliyon.2024.e32548.

[20] Shatte ABR, Hutchinson DM, Teague SJ. Machine learning in mental health: a scoping review of methods and applications. Psychological medicine. 2019; 49: 1426–1448. https://doi.org/10.1017/S0033291719000151.

[21] Atmakuru A, Shahini A, Chakraborty S, Seoni S, Salvi M, Hafeez-Baig A, *et al*. Artificial intelligence-based suicide prevention and prediction: A systematic review (2019–2023). Information Fusion. 2025; 114: 102673. https://doi.org/10.1016/j.inffus.2024.102673.

[22] Pigoni A, Delvecchio G, Turtulici N, Madonna D, Pietrini P, Cecchetti L, *et al*. Machine learning and the prediction of suicide in psychiatric populations: a systematic review. Translational psychiatry. 2024; 14: 140. https://doi.org/10.1038/s41398-024-02852-9.

[23] Fernandez A, Salvador-Carulla L, Choi I, Calvo R, Harvey SB, Glozier N. Development and validation of a prediction algorithm for the onset of common mental disorders in a working population. The Australian and New Zealand journal of psychiatry. 2018; 52: 47–58. https://doi.org/10.1177/0004867417704506.

[24] Barak-Corren Y, Castro VM, Nock MK, Mandl KD, Madsen EM, Seiger A, *et al*. Validation of an Electronic Health Record-Based Suicide Risk Prediction Modeling Approach across Multiple Health Care Systems. JAMA network open. 2020; 3: e201262. https://doi.org/10.1001/jamanetworkopen.2020.1262.

[25] Simon GE, Shortreed SM, Johnson E, Rossom RC, Lynch FL, Ziebell R, *et al*. What health records data are required for accu-

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

rate prediction of suicidal behavior? Journal of the American Medical Informatics Association: JAMIA. 2019; 26: 1458–1465. https://doi.org/10.1093/jamia/ocz136.

[26] Shortreed SM, Walker RL, Johnson E, Wellman R, Cruz M, Ziebell R, *et al*. Complex modeling with detailed temporal predictors does not improve health records-based suicide risk prediction. NPJ digital medicine. 2023; 6: 47. https://doi.org/10.1038/s41746-023-00772-4.

[27] Nock MK, Millner AJ, Ross EL, Kennedy CJ, Al-Suwaidi M, Barak-Corren Y, *et al*. Prediction of Suicide Attempts Using Clinician Assessment, Patient Self-report, and Electronic Health Records. JAMA Network Open. 2022; 5: e2144373. https://doi.org/10.1001/jamanetworkopen.2021.44373.

[28] Belsher BE, Smolenski DJ, Pruitt LD, Bush NE, Beech EH, Workman DE, *et al*. Prediction Models for Suicide Attempts and Deaths: A Systematic Review and Simulation. JAMA Psychiatry. 2019; 76: 642–651. https://doi.org/10.1001/jamapsychiatry.2019.0174.

[29] Koutsouleris N, Hauser TU, Skvortsova V, De Choudhury M. From promise to practice: towards the realisation of AI-informed mental health care. The Lancet. Digital health. 2022; 4: e829–e840. https://doi.org/10.1016/S2589-7500(22)00153-4.

[30] Salazar de Pablo G, Studerus E, Vaquerizo-Serrano J, Irving J, Catalan A, Oliver D, *et al*. Implementing Precision Psychiatry: A Systematic Review of Individualized Prediction Models for Clinical Practice. Schizophrenia bulletin. 2021; 47: 284–297. https://doi.org/10.1093/schbul/sbaa120.

[31] Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and Regression Trees. Chapman and Hall/CRC: New York, NY. 2017.

[32] World Medical Association. World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Participants. JAMA. 2025; 333: 71–74. https://doi.org/10.1001/jama.2024.21972.

[33] Spielberger CD. State–Trait Anger Expression Inventory–2: STAXI-2: Professional Manual. Psychological Assessment Resources: Odessa, FL. 1999.

[34] Spielberger CD, Miguel-Tobal JJ, Casado Morales MI, Cano-Vindel A. STAXI-2: Inventario de expresión de ira estado-rasgo: Manual. TEA Ediciones: Madrid, Spain. 2001.

[35] Spielberger CD, Jacobs GA, Russell SF, Crane RS. State–Trait Depression Inventory: ST–DEP Form Y. Psychological Assessment Resources: Odessa, FL. 1997.

[36] Spielberger CD, Buela-Casal G, Agudelo Vélez DM. IDER: Inventario de Depresión Estado-Rasgo: Manual. TEA Ediciones: Madrid, Spain. 2008.

[37] Moos RH. Coping Responses Inventory: CRI-Adult Form: Professional Manual. Psychological Assessment Resources: Odessa, FL. 1993.

[38] Moos RH, Kirchner Nebot T, Forns i Santacana M. CRI-A: Inventario de Respuestas de Afrontamiento-Adultos: Manual. TEA Ediciones: Madrid, Spain. 2010.

[39] Derogatis LR. SCL-90-R: Administration, Scoring and Procedures Manual. 3rd ed. National Computer Systems: Minneapolis. 1994.

[40] González de Rivera JL, De Las Cuevas C, Rodríguez-Abuín M, Rodríguez-Pulido F. El cuestionario de 90 síntomas: Adaptación española del SCL-90-R de L. R. Derogatis. TEA Ediciones: Madrid, Spain. 2002.

[41] Beck AT, Weissman A, Lester D, Trexler L. The measurement of pessimism: The Hopelessness Scale. Journal of consulting and clinical psychology. 1974; 42: 861–865. https://doi.org/10.1037/h0037562.

[42] Sánchez-Teruel D, Robles-Bello MA, Camacho-Conde JA. Validity of the spanish version of the Herth Hope Index and The Beck Hopelessness Scale in people who have attempted suicide. Actas españolas de psiquiatría. 2020; 48: 163–168.

[43] Rosenberg M. Society and the Adolescent Self-Image. Princeton University Press: Princeton, NJ. 1965.

[44] Martín-Albo J, Núñez JL, Navarro JG, Grijalvo F. The Rosenberg Self-Esteem Scale: translation and validation in university students. The Spanish journal of psychology. 2007; 10: 458–467. https://doi.org/10.1017/s1138741600006727.

[45] Costa PT Jr, McCrae RR. Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional Manual. Psychological Assessment Resources: Odessa, FL. 1992.

[46] Cordero Pando A, Pamos A, Seisdedos Cubero N, Costa PT Jr. Inventario de Personalidad NEO Revisado (NEO PI-R), Inventario NEO Reducido de Cinco Factores (NEO-FFI): Manual profesional. TEA Ediciones: Madrid, Spain. 1999.

[47] Chekroud AM, Bondar J, Delgadillo J, Doherty G, Wasil A, Fokkema M, *et al*. The promise of machine learning in predicting treatment outcomes in psychiatry. World psychiatry: official journal of the World Psychiatric Association (WPA). 2021; 20: 154–170. https://doi.org/10.1002/wps.20882.

[48] Khoo LS, Lim MK, Chong CY, McNaney R. Machine Learning for Multimodal Mental Health Detection: A Systematic Review of Passive Sensing Approaches. Sensors. 2024; 24: 348. https://doi.org/10.3390/s24020348.

[49] Song YY, Lu Y. Decision tree methods: applications for classification and prediction. Shanghai archives of psychiatry. 2015; 27: 130–135. https://doi.org/10.11919/j.issn.1002-0829.215044.

[50] Perlis RH. Artificial Intelligence and the Potential Transformation of Mental Health. JAMA Psychiatry. 2026; 83: 409–413. https://doi.org/10.1001/jamapsychiatry.2025.4116.

[51] GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The lancet. Psychiatry. 2022; 9: 137–150. https://doi.org/10.1016/S2215-0366(21)00395-3.

[52] Turecki G, Brent DA, Gunnell D, O'Connor RC, Oquendo MA, Pirkis J, *et al*. Suicide and suicide risk. Nature reviews. Disease primers. 2019; 5: 74. https://doi.org/10.1038/s41572-019-0121-0.

[53] Fazel S, Runeson B. Suicide. The New England journal of medicine. 2020; 382: 1078. https://doi.org/10.1056/NEJMx200005.

[54] O'Connor RC, Nock MK. The psychology of suicidal behaviour. The lancet. Psychiatry. 2014; 1: 73–85. https://doi.org/10.1016/S2215-0366(14)70222-6.

[55] Takahashi M, Imahara K, Miyamoto Y, Myojo K, Yasuda M. Association between the Big Five personality traits and suicide-related behaviors in Japanese institutionalized youths. PCN reports: psychiatry and clinical neurosciences. 2024; 3: e186. https://doi.org/10.1002/pcn5.186.

[56] Mota MSSD, Ulguim HB, Jansen K, Cardoso TDA, Souza LDDM.

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.

Are big five personality traits associated to suicidal behaviour in adolescents? A systematic review and meta-analysis. Journal of affective disorders. 2024; 347: 115–123. https://doi.org/10.1016/j.jad.2023.11.002.

[57] Flint J, Cohen L, Nath D, Habib Z, Guo X, Galynker I, *et al*. The Association Between the Suicide Crisis Syndrome and Suicidal Behaviors: The Moderating Role of Personality Traits. European psychiatry: the journal of the Association of European Psychiatrists. 2021; 64: e63. https://doi.org/10.1192/j.eurpsy.2021.2235.

[58] Lund C, Brooke-Sumner C, Baingana F, Baron EC, Breuer E, Chandra P, *et al*. Social determinants of mental disorders and the Sustainable Development Goals: a systematic review of reviews. The lancet. Psychiatry. 2018; 5: 357–369. https://doi.org/10.1016/S2215-0366(18)30060-9.

[59] Riley RD, Ensor J, Snell KIE, Harrell FE, Martin GP, Reitsma JB, *et al*. Calculating the sample size required for developing a clinical prediction model. BMJ. 2020; 368: m441. https://doi.org/10.1136/bmj.m441.

© 2026 The Author(s). This is an open access article under the CC BY 4.0 license.