UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL FACULDADE DE MEDICINA

PROGRAMA DE PÓS-GRADUAÇÃO EM PSIQUIATRIA E CIÊNCIAS DO COMPORTAMENTO

DISSERTAÇÃO DE MESTRADO

PREJUÍZO FUNCIONAL EM PACIENTES COM TRANSTORNO DE HUMOR

Kyara Rodrigues de Aguiar

Orientador: Prof. Dr. Ives Cavalcante Passos

Porto Alegre 2022

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL FACULDADE DE MEDICINA

PROGRAMA DE PÓS-GRADUAÇÃO EM PSIQUIATRIA E CIÊNCIAS DO COMPORTAMENTO

DISSERTAÇÃO DE MESTRADO

PREJUÍZO FUNCIONAL EM PACIENTES COM TRANSTORNO DE HUMOR

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

Autor: Kyara Rodrigues de Aguiar

Orientador: Prof. Dr. Ives Cavalcante Passos

Porto Alegre 2022

CIP - Catalogação na Publicação

PREJUÍZO FUNCIONAL EM PACIENTES COM TRANSTORNO DE HUMOR

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

Porto Alegre, outubro de 2022

A Comissão Examinadora, abaixo assinada, aprova a Dissertação elaborada por Kyara Rodrigues de Aguiar, como requisito parcial para a obtenção do Grau de Mestre em Psiquiatria e Ciências do Comportamento.

Banca Examinadora:

Prof.ª Dr.ª Adriane Ribeiro Rosa – Universidade Federal do Rio Grande do Sul – Relatora

Dr. Daniel Prates Baldez – Universidade Federal do Rio Grande do Sul

Dr. Francisco Diego Rabelo-da-Ponte – King's College of London

Prof. Dr. Ives Cavalcante Passos – Universidade Federal do Rio Grande do Sul – Orientador

Profª. Dra. Karen Jansen – Universidade Católica de Pelotas – Co-orientadora

AGRADECIMENTOS

Às 282 pessoas de Pelotas que se dispuseram a participar dessa pesquisa.

À Profª. Drª. Karen Jansen por transmitir e dividir comigo o ecantanto pela ciência e principalmente, pela oportunidade de poder contar com seu exemplo de trabalho em equipe e liderança.

À todos os meus colegas que dividiram os dias e o trabalho comigo no PPG Saúde e Comportamento – UCPel.

À Drª Taiane Cardoso Azevedo, por se fazer presente e sempre me ensinar com tanto carinho.

Aos meus colegas no grupo Alliance, pelos insights, pelo conhecimento compartilhado e pelos bons momentos que dividimos.

Ao meu orientador, pela parceria e pelos ensinamentos.

Por fim, a minha mãe, por desde cedo me ensinar a importância do ensino e da educação.

LISTA DE ABREVIATURAS E SIGLAS

TDM – Transtorno Depressivo Maior

TB – Transtorno Bipolar

MINI – Mini International Neuropsychiatric Interview

ABEP – Associação Brasileira das Empresas de Pesquisas

BDI – Beck Depression Inventory

ASSIST – Alcohol, Smoking and Substance Involvement Screening Test

BSI – Beck Scale for Suicide Ideation

SRQ – Self-regulation questionnaire

HCL – Hypomania Checklist

CTQ – Childhood Trauma Questionnaire

FAST – Functioning Assessment Short Test

RFE – Recursive feature elimination

ROSE – Random Over-Sampling Examples

RF – Randon Forest

LASSO – Least Absolute Shrinkage and Selection Operator

AUC – Area under the curve

RESUMO

O Transtorno Depressivo Maior e o Transtorno Bipolar são transtornos psiquiátricos associados a comprometimento psicossocial. Apesar da melhora clínica, as queixas funcionais geralmente permanecem, prejudicando principalmente o desempenho ocupacional e cognitivo. Neste estudo, objetivamos utilizar técnicas de aprendizado de máquina para prever o comprometimento funcional em pacientes eutímicos com diagnóstico de transtornos de humor. Nossos dados são provenientes de um estudo de coorte de base populacional realizado na cidade de Pelotas. Este estudo teve sua primeira etapa realizada em 2007-2009 e a segunda em 2012-2014, tendo em média 5 anos entre a avaliação de base e a avaliação de acompanhamento. Todos os participantes diagnosticados com transtorno de humor na linha de base e posteriormente reavaliados foram considerados (n= 282).

Random Forest (RF) com seleção prévia de variaveis e algoritmos LASSO foram aplicados a um conjunto de treinamento com dados imputados resultando em dois modelos principais. Após a seleção prévia de variáveis, 25 variáveis foram mantidas. O modelo de RF teve melhor desempenho em relação ao LASSO (Área sob a curva (AUC): 0,715 vs. 0,693). As variáveis mais importantes na predição do comprometimento funcional foram abuso sexual, escore total do SRQ, negligência física, abuso emocional e abuso físico.

O modelo demonstrou desempenho aceitável para predizer comprometimento funcional. No entanto, nossa amostra é composta por participantes jovens, nesse sentido, o modelo pode não ser generalizado para indivíduos mais velhos com transtornos de humor. A calculadora apresentada possui dados clínicos, sociodemográficos e ambientais, demonstrando que é possível utilizar tais informações para prever o desempenho funcional. Para estudos futuros, sugerimos a integração de dados digitais de saúde e informações biológicas.

Palavras-chave: Transtorno Depressivo Maior, Transtorno Bipolar, Transtornos do Humor, Comprometimento Funcional, Aprendizado de Máquina.

ABSTRACT

Major Depressive Disorder and Bipolar Disorder are psychiatric disorders associated with psychosocial impairment. Despite clinical improvement, functional complaints usually remain, mainly impairing occupational and cognitive performance. In this study, we aimed to use machine learning techniques to predict functional impairment in euthymic patients diagnosed with mood disorders. Our data come from a population-based cohort study carried out in the city of Pelotas. This study had its first stage carried out in 2007-2009 and the second in 2012-2014, taking an average of 5 years between the baseline assessment and the follow-up assessment. All participants diagnosed with mood disorder at baseline and later reassessed were considered (n=282).

Random forest (RF) with previous recursive feature selection and LASSO algorithms were applied to a training set with imputed data by bagged trees resulting in two main models. Following recursive feature selection, 25 variables were retained, and the RF model had the best performance compared to LASSO (Area under the curve (AUC): 0.715 vs. 0.693). The most important variables in predicting functional impairment were sexual abuse, SRQ total score, physical negligence, emotional abuse, and physical abuse.

The model demonstrated acceptable performance to predict functional impairment. However, our sample is composed of young participants and the model may not generalize to older individuals with mood disorders. Further studies are needed in this direction. The calculator presented has clinical, sociodemographic and environmental data, demonstrating that it is possible to use such information to predict functional performance. For future studies, it would be interesting to include integrating digital health data and biological information.

Keywords: Major Depressive Disorder, Bipolar Disorder, Mood Disorders, Functional Impairment, Machine Learning.

SUMÁRIO

APRESENTAÇÃO

A presente dissertação de mestrado consiste em um estudo de corteenvolvendo avaliação clínica de pacientes com TB e TDM e utilizando algoritmos de *machine learning* para analisar os dados. O artigo resultante dessa pesquisa foi aceito para publicação na *Psychiatry Research* com o título "Predicting functional impairment in patients with mood disorder: A 5-year follow-up".

O presente trabalho, apresentado ao Programa de Pós-Graduação em Psiquiatria e Ciências do Comportamento, seguira a seguinte ordem: introdução, justificativa, objetivos, métodos, aspectos éticos, artigo e considerações finais.

INTRODUÇÃO

Transtorno de Humor Bipolar e Unipolar

O Transtorno Bipolar (TB) é uma doença mental crônica e heterogênea, caracterizada por episódios depressivos e episódios maníacos (bipolar-I) e/ou hipomaníacos (bipolar-II) (DSM-5)[1], que afeta 2,4 da população mundial[2] e costuma ter seu início na juventude[3]. O TB e o Transtorno Depressivo Maior (TDM) partilham de muitas características em sua etiologia, sendo a presença de episódios maníacos e/ou hipomaníacos o único fator que separa os dois diagnósticos, de acordo com o DSM-5[1]. O TDM é caracterizado por humor negativo com distorções cognitivas e é a doença mental que afeta o maior número de indivíduos ao redor do mundo, atingindo aproximadamente 322 milhões de indivíduos[4].

Ambas patologias apresentam déficits funcionais que se perpetuam apesar da recuperação sindrômica[5,6]. A recuperação funcional nas tarefas domésticas, escolares, resgate de relações interpessoais e produtividade no trabalho envolvem processos complexos no mundo real, dessa forma, não é surpreendente que exista uma lacuna entre a remissão de sintomas clínicos e o resultado funcional, já que esses fatores vão além de estar clinicamente estável. Atualmente as evidências científicas disponíveis nos mostram que a recuperação da funcionalidade depende de variáveis como cursoda doença/gravidade, genética, comorbidades e reserva cognitiva [7,8].

No TB os prejuízos cognitivos estão associados a déficits ocupacionais. Os prejuízos cognitivos são preditores independentes mais fortes de resultados ocupacionais do que da gravidade de sintomas de humor[9]. No TDM os déficits funcionais costumam ser relatados em áreas relacionadas ao status de emprego, produtividade ocupacional, relacionamentos interpessoais e autonomia[10,11].

Portanto, as consequências do quadro clínico apresentado por indivíduos com TB e TDM acabam sendo responsáveis por uma carga econômica substancial. Ainda, os custos indiretos com essas patologias (desemprego, diminuição de produtividade) representam um ônus econômico e psicossocial ainda maior para sociedade em comparação com custos diretos (tratamento, internação, atendimento ambulatorial) [9,12,13]. Com base nessas evidências, fica manifesto o interesse vital não apenas nos

sintomas, mas no funcionamento do indivíduo de forma ampla dentro dessas patologias, demanda que confere relevância a esse estudo.

Prejuízo Funcional

O prejuízo funcional é um sintoma comum dentro de variados diagnósticos psiquiátricos. Evidências recentes demonstram que diferentes diagnósticos psiquiátricos apresentam diferentes níveis de limitações funcionais[14]. Nos transtornos de humor o prejuízo funcional parece estar associado a resistência ao tratamento, envolvimento em atividades espontâneas, estruturação de tarefas, competência e flexibilidade cognitiva[15]. Um estudo envolvendo pacientes depressivos demonstrou que o grau do funcionamento esteve relacionado à gravidade dos sintomas[16], no TB o prejuízo funcional parece estar associado a sintomas residuais de humor[17].

Um estudo que acompanhou participantes com depressão persistente, recorrente e totalmente em remissão durante 6 anos evidenciou que a melhora dos sintomas foi associada à redução do comprometimento funcional, no entanto, nenhum dos grupos diagnosticados atingiu o nível dos controles saudáveis em relação a funcionalidade[18]. Em média, houve uma correlação moderada entre a gravidade dos sintomas e as limitações no funcionamento[18]. Esse dado nos permite entender que mesmo em remissão alguns prejuízos persistem e que, apesar de haver uma correlação entre gravidade de sintomas clínicos e prejuízo funcional, estes fenômenos são aspectos distintos da doença.

No TB o comprometimento funcional social e neurocognitivo tem um impacto de moderado a grave[19]. Apesar do tratamento adequado, a disfunção social e neurocognitiva continua sendo um problema grave para esses pacientes. É relatado que 60-70% dos pacientes com TB apresentam graus variados de comprometimento dentro da doença[20] e assim como ocorre no TDM, estudos indicam que esse prejuízo persiste não apenas na fase aguda da doença, mas também em remissão[21,22]. Outro estudo, considerando 219 participantes com transtorno de humor mostrou que apenas 37,6% dos pacientes obtiveram recuperação funcional após a primeira hospitalização [23].

Dessa forma, podemos visualizar uma heterogeneidade que não ocorre somente nos fatores desencadeantes de prejuízos funcionais, mas também em um determinado subgrupo de pacientes. No entanto, o mapeamento do subgrupo depacientes que irá apresentar tais déficits ainda permanece em aberto na literatura. Aosestudos citados acima soma-se a necessidade de desenvolver trabalhos que tenham como objetivo o desenvolvimento de análises a nível individual, que possibilitem contemplar aspectos complexos dessas patologias.

Machine Learning

O grande objetivo dentro do campo do *machine learning* é utilizar algoritmos para aprender novas informações em bancos de dados que possam resultar em tomadas de decisões informadas com relação ao futuro. Com frequência nos referimosa técnicas de *machine learning* como algoritmos e apesar desse termo soar complexo, seu significado é bastante simples. Algoritmos de aprendizado de máquina podem ser entendidos como um conjunto de passos a serem seguidos durante um determinado processo, passos que um computador deve seguir[24]. Isso é importante porque parte fundamental do processo da análise é a escolha dos algoritmos a serem utilizados durante o processo. Alguns algoritmos irão performar melhor do que outros, dependendo do problema a ser resolvido, do número de observações presentes no banco e a forma com que os dados estão apresentados[25].

Machine Learning (ML) é uma subseção dentro do campo da Inteligência Artificial (IA). IA inclui qualquer tipo de técnica com o objetivo de fazer com que um sistema computacional tenha um comportamento próximo ao de um ser humano. No caso do ML, aplicamos estatísticas em conjuntos de dados na tentativa de descobrir novas informações que permitam generalizações em conjuntos de dados não vistos anteriormente[26].

O aprendizado em ML é dividido em duas categorias: aprendizado supervisionado e não supervisionado. O primeiro acontece em duas etapas, onde primeiro utilizamos uma parcela dos dados para treinar o algoritmo em identificar padrões entre as informações e em um segundo momento, aplicamos o modelo resultante dessa aprendizagem em um conjunto de dados nunca visto antes, com o objetivo de verificar o desempenho da máquina em generalizar sua aprendizagem[26].

No aprendizado não supervisionado não há diferença entre o banco de treino e teste, isso significa que enquanto os algoritmos supervisionados nos ajudam a atribuir rótulos conhecidos a novas observações (ex: sujeitos com prejuízo funcional/sem prejuízo funcional), os algoritmos não supervisionados nos ajudam a descobrir novos rótulos, ou agrupamentos das observações em nosso conjunto de dados[24].

Para realizar a aplicação de algoritmos são necessários alguns passos, começando pela coleta dos dados, preparação e exploração dos dados, divisão do banco de dados entre treino e teste, treino do modelo conforme explicado acima, e por último a avaliação do modelo. Nesta última etapa estaremos avaliando a efetividade do algoritmo empregado[27]. A avaliação do modelo resultante é feita através do número e/ou a magnitude dos erros que ele produz. Para problemas de classificação podemos observar a porcentagem de vezes que que o algoritmo realizou uma predição incorreta[24]. Da mesma forma, a porcentagem de vezes que o modelo fez predições corretas do desfecho em questão irá nos oferecer a acurácia do modelo[24].

JUSTIFICATIVA

Técnicas de aprendizado de máquina supervisionado podem fornecer uma oportunidade para desenvolver modelos lineares e não lineares multivariados com menor risco de viés e *overfitting*. Isso é possível através de métodos de seleção de variáveis que permitem identificar um subconjunto de variáveis relevantes entre um conjunto candidato de preditores potenciais e testar a generalização desses modelos em conjuntos de dados independentes. Além disso, os algoritmos de aprendizado de máquina supervisionados são adequados para modelar fenômenos não lineares complexos, já que os métodos de seleção de variaveis oferecem a oportunidade de atenuar a multicolinearidade[28], definida como variáveis independentes altamente correlacionadas, que é um problema comum na regressão linear e logística multivariada tradicional[29]. Portanto, dada a falta de homogeneidade no comprometimento funcional em pacientes com transtornos de humor e os desafios de prever quais pacientes específicos que irão apresentar comprometimento funcional, o presente estudo usa técnicas de aprendizado de máquina para prever o

comprometimento funcional em pacientes com transtornos de humor em uma amostra de base populacional, a nível individual.

OBJETIVOS

Objetivo geral

- Utilizar técnicas de aprendizado de máquina para prever o comprometimento funcional em pacientes com transtornos de humor.

Objetivos específicos

- Construir uma calculadora de risco.
- Testar o melhor modelo em três subamostras diferentes: indivíduoscom outros transtornos psiquiátricos, jovens com transtornos de humor atual e indivíduos saudáveis.

MÉTODO

Critérios de inclusão

Preencher critérios diagnósticos para TDM ou TB, atual ou passado, entre 2007 e 2009 e estar eutímico na segunda avaliação, entre 2012 e 2014.

Critérios de exclusão

Incapacidade de compreensão da entrevista/avaliação clínica.

Não ter respondido ao FAST no seguimento do estudo.

Instrumentos

Informações avaliadas no baseline:

Transtornos psiquiátricos: O diagnóstico de transtornos psiquiátricos foi realizado por meio do Mini International Neuropsychiatric Interview (MINI Plus)[1]. O MINI-Plus é uma breve entrevista diagnóstica padronizada, compatível com os critérios do DSM-IV e CID-10. Destinado à avaliação aprofundada dos transtornos mentais ao

longo da vida, explora sistematicamente todos os critérios de inclusão e exclusão e a cronologia de 23 categorias diagnósticas do DSM-IV. Neste estudo, foram consideradas as seguintes categorias: (A) TDM, (B) Distimia, (D) TB I - II, (E) Pânico, (F) Agorafobia, (G) Fobia Social, (I) Transtorno Obsessivo Compulsivo (TOC), (J) Transtorno de Estresse Pós-Traumático (TEPT), (M) Transtorno Psicótico e (P) Transtorno de Ansiedade Generalizada (TAG).

Características Sociodemográficas: As informações foram coletadas através de um questionário estruturado com questões referentes a sexo, idade, estado civil, trabalho, escolaridade e nível socioeconômico por meio da Associação Brasileira das Empresas de Pesquisas (ABEP)[30].

Gravidade dos sintomas depressivos: Os sintomas depressivos foram avaliados por meio do Inventário de Depressão de Beck (BDI)[30,31]. Essa escala é composta por 21 itens que investigam a presença e intensidade de sintomas depressivos. A soma de todos os itens resulta na pontuação total, e quanto maior a pontuação, maior a gravidade dos sintomas depressivos. Este instrumento foi validado e adaptado para a população brasileira[32].

Abuso ou dependência de substâncias: Avaliado usando o Teste de Triagem de Envolvimento com Álcool, Tabagismo e Substâncias (ASSIST)[33]. É um questionário estruturado composto por oito questões sobre o uso de nove classes de substâncias psicoativas (tabaco, álcool, maconha, cocaína, estimulantes, sedativos, inalantes, alucinógenos e opiáceos). A frequência do uso da substância é verificada atualmente e nos três meses anteriores. Esse instrumento foi adaptado e validado para a população brasileira[34]. Para o álcool, um valor maior ou igual a 4 foi considerado moderado, enquanto para as demais classes de drogas o ponto de corte foi 11. A partir das classes mencionadas anteriormente, foram criadas duas variáveis e estas foram adicionadas ao conjunto de dados a ser analisado. A primeira variável assume valor 1 se o indivíduo estiver em risco moderado para álcool ou tabaco. Enquanto a segunda variável assume valor 1, se o indivíduo apresentar risco moderado em qualquer outra classe de substâncias psicoativas.

Ideação suicida: A Escala de Beck para Ideação Suicida[32,35] (BSI) foi usada para triagem de ideação suicida. A escala BSI foi traduzida e validada para uso no Brasil com confiabilidade de 0,90. O BSI avalia a presença e intensidade de pensamentos suicidas

uma semana antes da avaliação por meio de 19 itens, divididos em três partes: risco e presença de ideação suicida; gravidade da ideação, atitude e planos suicidas; gravidade da intenção de morrer e o número de tentativas de suicídio. Sua pontuação pode variar de 0 a 38, onde quanto maior a pontuação, maior o risco clínico. Para cada item, o participante tem as seguintes opções de resposta: nenhum, fraco ou moderado a forte, com pesos de 0 a 2 respectivamente. Além disso, quatro itens foram inseridos individualmente no modelo: "Desejo de viver"; "Desejo morrer"; "Quero me matar"; "Se sua vida estivesse em risco, você tentaria se salvar?".

Transtornos mentais não psicóticos: Ansiedade, depressão e sintomas somáticos foram avaliados por meio do questionário de autorregulação (SRQ-20)[36]. Esse instrumento é composto por 20 itens que buscam rastrear transtornos mentais não psicóticos por meio de opções de respostas afirmativas e negativas. As respostas afirmativas obtêm valor igual a 1, o escore total é obtido pela soma desses valores, sendo 0 = nenhuma probabilidade de presença de transtorno não psicótico e 20 = possibilidade extrema. Neste estudo, além de considerar o escore total, adicionamos separadamente ao modelo os seguintes itens: "(3) Você costuma ter noites mal dormidas?"; "(11) Você sente dificuldade em aproveitar suas atividades diárias?"; "(17) A ideia de acabar com sua vida já passou pela sua cabeça?".

Sintomas hipomaníacos: Usamos a versão brasileira do Hypomania Checklist-32 (HCL-32)[37] para investigar sintomas de hipomania. Esse instrumento foi validado após o estudo, porém, não foram feitas alterações substanciais. O HCL-32 é um instrumento autoaplicável composto por 32 questões sim/não, validado internacionalmente. Quando desenvolvido, seu objetivo era identificar características hipomaníacas em pacientes depressivos, visando facilitar o diagnóstico de TB-II. Os participantes são instruídos a se concentrarem em um período em que se sentiram com o humor "para cima" e indicar pensamentos ou emoções específicas durante esse período. Alémdisso, a escala possui 8 itens de gravidade e impacto funcional associados à duração dos episódios. Ao final do questionário, os participantes são solicitados a avaliar o grau de impacto (positivo, sem impacto, negativo ou neutro) na vida familiar, social, escolar e de lazer.

Além disso, os participantes foram questionados sobre uma série de questões clínicas e ambientais. Para melhor visualização, todas as informações de linha de base inseridas no modelo que não fazem parte dos instrumentos especificados acima serão apresentadas em formato de quadro (Quadro 1).

Quadro 1. Informações coletadas na avaliação inicial e adicionadas ao modelo.

Possui religião?	Sente que tem suporte na vida?	Seus pais já faleceram?
Participa de grupo religioso?	Possui pais divorciados?	Você conhece alguém que já cometeu suicídio?
Está estudando atualmente?	Alguma vez já realizou tratamento psicológico ou psiquiátrico?	Seus pais já receberam, alguma vez, diagnóstico psiquiátrico de TB?
Em que idade começou a trabalhar?	Alguma vez já iniciou tratamento psicológico e interrompeu abruptamente?	Está trabalhando atualmente?
Já desempenhou algum trabalho remunerado?	Com que idade, em anos, você experimentou alguma substância pela primeira vez, incluindo álcool e tabaco?	Você já foi fisicamente forçado a fazer sexo?
Você tem um parceiro fixo?	Alguma vez você já se envolveu em uma briga envolvendo agressão física?	TDM, distimia, TB-I, TB-II, pânico, agorafobia, fobia social, TOC, TEPT, TAG através do MINI[1]

Informações avaliadas no acompanhamento:

Trauma na infância: Para avaliar as experiências de trauma na infância, foi utilizado o Childhood Trauma Questionnaire (CTQ). Trata-se de um instrumento autoaplicável adaptado para o Brasil, para uso com adolescentes (a partir de 12 anos) e adultos, investigando a história de abuso e negligência na infância[38]. Por meio desse questionário, podem ser avaliados cinco componentes traumáticos: abuso físico, abuso emocional, abuso sexual, negligência física e negligência emocional. É composto por 28 afirmações relacionadas a situações da infância em uma escala Likert de cinco pontos. Esse instrumento foi incorporado ao estudo de 2012 a 2014, mas, como as questões são relacionadas à infância, optamos por incluí-las nas possíveis variáveis preditoras.

Além disso, novas variáveis foram incorporadas na entrevista de acompanhamento e adicionadas ao modelo, uma vez que são informações sobre a história e o contexto familiar dos participantes. Essas informações serão apresentadas em formato de quadro (Quadro 2).

Quadro 2. Informações coletadas no acompanhamento e adicionadas ao modelo.

Alguma vez seus pais foram internados em hospital psiquiátrico?	Alguma vez seus pais fizeram uso de medicação psiquiátrica?	Possui alguém na sua família com histórico de TB?
Você alguma vez foi internado em hospital psiquiátrico?	Seus pais já receberam diagnóstico de alguma doença psiquiátrica?	Você toma alguma medicação psiquiátrica?
Alguma vez seus pais tentaram suicídio?	Seus irmãos já receberam diagnóstico de alguma doença psiquiátrica?	Transtorno psicótico e pânico ao longo da vida através da MINI[μ]

Desfecho:

Avaliação funcional: O Functional Assessment Short Test (FAST)[39] foi utilizado para avaliar a funcionalidade dos participantes. É uma medida para avaliação objetiva de 6 áreas de funcionamento: autonomia, funcionamento ocupacional, funcionamento cognitivo, questões financeiras, relações interpessoais e tempo de lazer. Cada dimensão é composta por quatro respostas (sem dificuldade, dificuldade leve, dificuldade moderada, dificuldade severa), cujas pontuações podem variar de 0 a 3. A soma de todos os itens fornece a pontuação global, e quanto maior a pontuação, maior o dano . Foi considerado comprometimento funcional escores > 11.

4. Procedimento

2007-2009: Em um primeiro momento (tempo 0) foi realizado um estudo transversal de base populacional onde os jovens da cidade de Pelotas foram convidados a participar da pesquisa em seus domicílios e a seleção da amostra se deu por conglomerados em múltiplos estágios. Por se tratar de uma amostragem probabilística, todos os jovens residentes na zona urbana da cidade de Pelotas-RS foram considerados elegíveis. No total, 1560 adultos jovens foram avaliados por

estudantes treinados para o manuseio da *Mini International Neuropsychiatry Interview* (MINI 5.0)[1].

2012-2014: Em um segundo momento, todos os jovens avaliados no tempo 0 (n=1560) foram convidados a participar de uma reavaliação, na qual foi realizada uma avaliação de fatores ambientais e comportamentais, incluindo a avaliação do desempenho funcional e medidas clínicas de pior desfecho da doença. Tendo em sua conclusão a inclusão de 1244 adultos avaliados clinicamente.

5. *Análise estatística*

Machine Learning

O conjunto de dados inicial tinha 59 variáveis e 282 participantes. As análises foram realizadas no software R versão 4.1.127 utilizando o RStudio. Os valores omissos foram imputados usando modelos de *bagged tree* versão 0.2.0[28,29]. O *tidymodels* versão 0.2 foi usado para dividir o conjunto de dados, treinar os modelos e calcular medidas de desempenho[40]. *Recursive feature elimination* (RFE) do pacote *caret* versão 6.0.92 foi usada para seleção de variáveis 41. Para a análise principal, o conjunto de dados foi dividido em duas partes: conjunto de treinamento (75% dos participantes, n = 197) e conjunto de teste (25% dos participantes, n = 85), mantendo a proporção do desfecho em cada conjunto . Na etapa de treinamento do modelo, o método *Random Over-Sampling Examples* (ROSE)[41] foi usado para corrigir o desequilíbrio de classe.

O conjunto de treinamento foi submetido a duas estratégias diferentes e independentes. A primeira consistiu na seleção de variáveis através do RFE e partindo para o treinamento modelo de Random Forests (RF)[42] com as variáveis selecionadas e utilizando a validação cruzada de 5 vezes com 5 repetições, otimizando *mtry* e número de árvores. A segunda envolveu o treinamento do modelo *Least Absolute Shrinkage and Selection Operator* (LASSO) (validação cruzada de 5 vezes com 5 repetições, otimizando lambda)[43]. A área sob a curva (AUC) foi considerada como a principal métrica na determinação do desempenho do modelo.

O desempenho do modelo foi avaliado usando um conjunto de 85 participantes (25% da amostra). Além da curva AUC e ROC, outras medidas de desempenho foram obtidas (sensibilidade, especificidade, valor preditivo positivo, valor preditivo negativo e acurácia balanceada). A importância das variáveis foi avaliada pelo *Shappley Additive exPlanations*[44] através do pacote vip 0.3.230. Quanto maior o valor de Shapley, maior a importância da variável para o modelo.

Calculadora de risco para prejuízo funcional

Para a implementação do modelo preditivo, foi desenvolvido um aplicativo web contendo a calculadora de risco com o *framework* R denominado *shiny*[45]. Os pacotes R *fresh*[46] e *shinydashboard*[47] também foram empregados no desenvolvimento do aplicativo. O aplicativo foi publicado na página shinyapps.io de um dos autores.

ASPECTOS ÉTICOS

Todos os participantes receberam informações sobre os objetivos da pesquisa e assinaram um termo de consentimento livre e esclarecido. De acordo com a avaliação, aqueles que necessitaram de encaminhados para tratamento foram encaminhados, de acordo com a demanda, para os seguintes centros: Ambulatório de Pesquisa e Extensão em Saúde Mental da UCPel, Centro de Atenção Psicossocial – CAPS, UnidadesBásicas de Saúde – UBS ou Hospital Espírita de Pelotas – HEP. O projeto foi aprovado pelo Comitê de Ética em Pesquisa da UCPel sob o protocolo de número 2008/118.

ARTIGO

Título: Predicting functional impairment in patients with mood disorder: a 5-year follow-up
Aceito para publicação na *Psychiatry Research*
Fator de impacto 11.3

Decision on submission to Psychiatry Research   Caixa de entrada ×

Psychiatry Research <em@editorialmanager.com>
para mim ▾

sáb., 5 de ago., 03:03

Manuscript Number: PSY-D-22-02707R2

Predicting functional impairment in euthymic patients with mood disorder: A 5-year follow-up

Dear Ms. Aguiar,

Thank you for submitting your manuscript to Psychiatry Research.

I am pleased to inform you that your manuscript has been accepted for publication.

My comments, and any reviewer comments, are below.
Your accepted manuscript will now be transferred to our production department. We will create a proof which you will be asked to check, and you will also be asked to complete a number of online forms required for publication. If we need additional information from you during the production process, we will contact you directly.

We appreciate you submitting your manuscript to Psychiatry Research and hope you will consider us again for future submissions.

We encourage authors of original research papers to share the research objects – including raw data, methods, protocols, software, hardware and other outputs – associated with their paper. More information on how our open access Research Elements journals can help you do this is available at https://www.elsevier.com/authors/tools-and-resources/research-elements-journals?dgcid=ec_em_research_elements_email.

Kind regards,
Lynn E. DeLisi, MD
Editor-in-Chief

Psychiatry Research

Editor and Reviewer comments:

Predicting functional impairment in patients with mood disorder: A 5-year follow-up

Kyara Rodrigues de Aguiar[1,2,3], Bruno Braga Montezano[1,2,3], Jacson Feiten[1,2,3], Devon Watts[4,] Aline Zimerman[1,2], Thaíse Campos Mondin[5,6], Ricardo Azevedo da Silva[5], Luciano Dias de Mattos Souza[5], Flávio Kapczinski[1,2,3,4], Taiane de Azevedo Cardoso[4], Karen Jansen[5], Ives Cavalcante Passos[1,2,3].

1. Laboratory of Molecular Psychiatry, Centro de Pesquisa Experimental (CPE) and Centro de Pesquisa Clínica (CPC), Hospital de Clínicas de Porto Alegre (HCPA), Porto Alegre, RS, Brazil.

2. Universidade Federal do Rio Grande do Sul, School of Medicine, Graduate Program in Psychiatry and Behavioral Sciences, Department of Psychiatry, Porto Alegre, RS, Brazil.

3. Instituto Nacional de Ciência e Tecnologia Translacional em Medicina (INCT-TM), Porto Alegre, RS, Brazil.

4. Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, ON, Canada.

5. Programa de Pós-Graduação em Saúde e Comportamento, Universidade Católica de Pelotas (UCPel), Rua Gonçalves Chaves, 373, sala 424 C, Pelotas 96015-560 RS, Brazil.

6. Universidade Federal de Pelotas (UFPel), Pró Reitoria de Assuntos Estudantis.

Corresponding author:

Ives Cavalcante Passos, MD, PhD
Professor of Psychiatry
Department of Psychiatry - Universidade Federal do Rio Grande do Sul (UFRGS)
E-mail: ivescp1@gmail.com
R. Ramiro Barcelos, 2400 – Floresta, Porto Alegre, RS, 90035002, Brazil.

Email for all authors:

Kyara Rodrigues de Aguiar: **kyarardeaguiar@gmail.com**
Bruno Braga Montezano: **brunodoyt@gmail.com**
Jacson Feiten: **jacsonfeiten95@gmail.com**
Devon Watts: **wattsd21@gmail.com**
Aline Zimerman: **zimermanaline@gmail.com**
Thaíse Campos Mondin: **thaisemondin@yahoo.com.br**
Taiane de Azevedo Cardoso: **taianeacardoso@hotmail.com**
Luciano Dias de Mattos Souza: **luciano.dms@gmail.com**
Ricardo Azevedo da Silva: **ricardo.as@uol.com.br**
Karen Jansen: **karenjansen315@hotmail.com**
Ives Cavalcante Passos: **ivescp1@gmail.com**

Abstract

Introduction: Major Depressive Disorder and Bipolar Disorder are psychiatric disorders associated with psychosocial impairment. Despite clinical improvement, functional complaints usually remain, mainly impairing occupational and cognitive performance.

Objective: To use machine learning techniques to predict functional impairment in patients with mood disorders.

Methods: Analyzes were performed using a population-based cohort study. Participants diagnosed with a mood disorder at baseline and reassessed were considered (n= 282). Random forest (RF) with previous recursive feature selection and LASSO algorithms were applied to a training set with imputed data by bagged trees resulting in two main models.

Results: Following recursive feature selection, 25 variables were retained, and the RF model had the best performance compared to LASSO (Area under the curve (AUC): 0.715 vs. 0.693). The most important variables in predicting functional impairment were sexual abuse, SRQ total score, physical negligence, emotional abuse, and physical abuse.

Conclusion: The model demonstrated acceptable performance to predict functional impairment. However, our sample is composed of young participants and the model may not generalize to older individuals with mood disorders. Further studies are needed in this direction. The calculator presented has clinical, sociodemographic and environmental data, demonstrating that it is possible to use such information to predict functional performance. For future studies, it would be interesting to include integrating digital health data and biological information.

Keywords: Major Depressive Disorder, Bipolar Disorder, Mood Disorders, Functional Impairment, Machine Learning.

1. Introduction

Major Depressive Disorder (MDD) is a highly prevalent psychiatric illness worldwide, with lifetime incidence rates of up to 21%[1] and is associated with psychosocial impairment and poor quality of life[2,3]. Similarly, Bipolar Disorder (BD) is one of the main causes of disability in the world, affecting 2.4% of the world's population[4], and is associated with functional impairment[5,6]. Despite clinical improvement (reduction of depressive/manic/hypomanic symptoms), self-reported functional complaints usually remain, mainly preventing satisfactory occupational and cognitive performance and impacting the quality of life of this population[7,8].Furthermore, there is a large degree of heterogeneity in functional deficits across patients, which complicates the ability to detect which patients are likely to show more pernicious outcomes over time[7,8].

A recent systematic review showed that residual depressive symptoms are associated with global functional impairment, assessed through Functional Assessment Short Test (FAST) in BD patients, with occupational functioning showing the largest deficits, followed by cognitive functioning[9]. Likewise, residual mood symptoms seem to be associated with functional impairments in MDD[8]. Additionally, in MDD, global functional impairment, as well as the subdomains: occupational functioning, leisure time and cognitive functioning, were associated with deficits in executive functions[10]. Regardless, these residual symptoms are usually perceived to a much greater extent by patients than by health professionals, at all stages of the disease[8].

Traditional group-based statistical analyses that assess a linear associationbetween risk factors and clinical outcomes, such as those shown above, are important to understand factors associated with functional impairment in this population, as they offer broad generalizations[8–10]. However, these approaches do not easily translate to identifying which specific patients will present with functional deficits at an individual level. In MDD and BD, predicting the outcome individually is especially important, considering that both are heterogeneous pathologies, where the course of the disease can occur in different ways, in patients with the same diagnosis.

In comparison to traditional statistical analyses, such as logistic or linear regression in the absence of resampling techniques, supervised machine learning techniques can provide an opportunity to develop multivariate linear and non-linear models with a lower risk of bias and overfitting, perform feature selection methods to identify a subset of relevant variables among a candidate set of potential predictors, and test the generalizability of these models in independent datasets. Furthermore, supervised machine learning algorithms are well-suited to model complex non-linear phenomena, and feature selection methods provide an opportunity to mitigate multicollinearity[11], defined as several highly correlated independent variables, which is a common problem in traditional multivariate linear and logistic regression[12]. Therefore, given the

lack of homogeneity in functional impairment across patients with mood disorders, and the challenges of predicting which specific patients will develop functional impairment, the current study uses machine learning techniques to predict functional impairment in patients with mood disorders in a population-based sample, at an individual level.

2. Methods

2.1 Study design

This is a population-based cohort study with baseline assessments taking place between 2007 and 2009. The sample consisted of clusters in multiple stages and all young adults aged between 18 and 24 years living in the urban area of Pelotas, Brazil, were considered eligible. Probability sampling was used to select participants, and as such, the sample was representative of the target population. More information on the selection of participants can be found in previously published articles[13,14].

In total, 1560 young adults responded to a diagnostic assessment using the Mini International Neuropsychiatric Interview (MINI 5.0)[15] in their homes. The assessment was carried out by students trained to handle the MINI[15]. Other information, such as: sociodemographic characteristics, environmental and behavioral factors, were also collected. On average 5 years after the first assessment, between 2012 and 2014, all young people were invited to participate in a reassessment. In this period, the functional performance assessment was included and used as an outcome in thisstudy. At its conclusion, 1244 adults were clinically reassessed. In both phases, participants received information about the objectives of the research and signed an informed consent form.

For the analyzes carried out in this study, all young people who met the diagnostic criteria for MDD or BD, current or past, between 2007 and 2009 and who were later reassessed, without a current episode, between 2012 and 2014, were considered. In total, 283 young people with mood disorders were included. Of the 283 subjects evaluated at follow-up, one was removed due to missing data in the study outcome. Thus, the remaining 282 individuals are distributed as follows: 226 subjects with a past depressive episode, 39 subjects with both episodes, and 17 subjects with a past episode of mania or hypomania (Figure 1).

2.2 Baseline predictores

- Psychiatric disorders: Mini International Neuropsychiatric Interview (MINI Plus)[15]

- Sociodemographic Characteristics: Associação Brasileira das Empresas de Pesquisas (ABEP)[16].

- Severity of depressive symptoms: Beck Depression Inventory (BDI)[17].

- Substance abuse or dependence: Alcohol, Smoking and Substance Involvement Screening Test (ASSIST)[19].

- Suicidal ideation: The Beck Scale for Suicide Ideation (BSI)[18,21].

- Non-psychotic mental disorders: Self-regulation questionnaire (SRQ-20)[22].

- Hypomanic symptoms: Hypomania Checklist-32 (HCL-32)[23].

Additionally, participants were asked a series of clinical and environmental questions. For better visualization, all baseline information entered in the model that are not part of the instruments specified above e a descrição detalhada de todos instrumentos listados acima will be presented in the supplementary material (Figure S1). Additionally, detailed descriptions of all instruments listed above are also availablevia supplementary material (Table S10).

2.3 Information evaluated at follow-up

- Childhood trauma: Childhood Trauma Questionnaire (CTQ)[24].

Also, new variables were incorporated in the follow-up interview and added to the model since they are information about the history and family context of the participants. This information is presented in the supplementary material (Figure S2). In addition, the detailed description of the CTQ is also available in supplementary material (Table S10).

2.4 Outcome

Functional assessment: The Functional Assessment Short Test (FAST)[26] was used to assess the participants' functionality. It is a measure for objective evaluation of6 areas of functioning: autonomy, occupational functioning, cognitive functioning, financial issues, interpersonal relationships and leisure time. Each dimension is composed of four answers (no difficulty, mild difficulty, moderate difficulty, severe difficulty), whose scores can range from 0 to 3. The sum of all items provide the global score, and the higher the score, the greater the damage. Functional impairment scores
> 11 were considered.

2.5 Machine Learning pipeline

As shown in Figure 2, the initial dataset had 59 features and 282 participants. The analyses were performed in the R software version 4.1.1[27] using RStudio. Missing values were imputed in using bagged tree models from recipes version 0.2.0 [28] [29]. The tidymodels package version 0.2 was used to partition the dataset, train models and calculate performance measures[30]. Recursive feature elimination (RFE) from caret

package version 6.0.92 was used for feature selection[31]. For the main analysis, the dataset was divided into two parts: training set (75% of participants, n = 197) and test set (25% of participants, n = 85), keeping the proportion of the functional impairment outcome in each set. In the training model step, the Random Over-Sampling Examples (ROSE) method was used to correct for class imbalance.

The training set was subjected to two different and independent strategies. The first consisted of feature selection using recursive feature elimination (RFE) and after training Random Forests (RF) with the selected variables using 5-fold cross validation with 5 repetitions optimizing mtry and number of trees. The second involved training an Least Absolute Shrinkage and Selection Operator (LASSO) (5-fold cross validation with 5 repetitions, optimizing lambda)[32]. Area under the curve (AUC) was consideredas the primary metric in determining model performance.

Model performance was assessed using a holdout set of 85 participants (25% of the sample). In addition to the AUC and ROC curve, other performance measures were obtained (sensitivity, specificity, positive predictive value, negative predictive value and balanced accuracy). The importance of variables was evaluated by Shappley Additive exPlanations[33] through the package vip 0.3.2[30]. This approach is more accurate when the predictors are correlated with each other, and provides an avenue to assess the interaction between variables[34]. The higher the Shapley value, the greater the importance of the variable for the model.

2.6 Risk calculator for functional impairment

For the implementation of the predictive model, a web application was developed containing the risk calculator with the R framework called *shiny*[35]. The*fresh*[36] and *shinydashboard*[37] R packages were also employed in web app development. The app was published on the shinyapps.io page of one of the authors.

2.7 Missing data

Variables with more than 10% missing values have been removed. Individuals with missing FAST values were excluded. Missing values were imputed using bagged tree models using the recipes package, version 0.2.0 [28] [29].

2.8 Supplementary analysis

In addition to the main analysis, seven models are available in the supplementary material (Table S5): a) RF without the previous variable selection step (recursive feature elimination), b) models without the data imputation step (only participants without any missing value were included in the analysis), c) models with different training/test division (50:50), d) model with different cross-validation (10-foldwith 10 repetitions), e) model without follow-up variables, except the outcome, f)

logistic regression. We also tested the best model in 3 different subsamples: subjects with other psychiatric disorders, young people with current mood disorders, and healthy subjects (Table S6). Descriptive statistics of subsamples and model performances on the subsamples are presented as supplementary material (Table S7, Table S8, Table S9).

3. Results

In total, 282 subjects were included in our analyses. Study participants had a mean age of 20.4 ± 2.1 years at baseline, with 69.8% (n = 197) female, 45.6% (n = 125) with incomplete high school, and 47.9 % (n = 135) in ABEP middle class. Brief description of the variables used in the analysis (Table S1), descriptive statistical analysis of demographic and clinical variables with imputation (Table S2) and without imputation (Table S3) are recounted in supplemental and the frequency of missing data for each variable is shown in Table S4.

The following variables were removed from the initial dataset because they had more than 10% missing data: *"At what age have you started working?", "Have you ever been physically forced to have sex?", "Do you participate in religious groups? ", HCL total score, "At what age did you start working?"*. Added to these, 6 variables were removed because they had low variance: *"Has your father and/or mother ever been admitted to a psychiatric hospital?" "Has your father and/or mother ever attempted suicide?", Panic disorder, Lifelong psychotic disorder, item 1 BSI, item 4 BSI*. Thus, 48 predictors were used for the RFE - RF or LASSO steps. Additionally, participants were also excluded if they had missing values in functional impairment (n = 1), leaving 282 participants for the analysis (89 with functional impairment and 193 without functionalimpairment).

RF algorithms (with a previous step of variable selection) and LASSO, were applied to the imputed training data. Following feature selection, 25 variables were retained in the RF model, and it showed better performance compared to LASSO (Area under the curve (AUC): 0.715 vs. 0.693). The test AUC of the RF model was 0.68 (confidence interval 0.56-0.80). Therefore, the model demonstrated acceptable performance to predict functional impairment at the 5-year follow-up. The model's performance measurements can be seen in detail in Table 1 (see Figure S3 for the receiver's operating characteristic curve).

A supplementary analysis was performed removing the variable selection step from the RF, which achieved a training AUC value of 0.693 (test AUC: 0.670). LASSO and RF models trained from a training dataset with only participants with no missing data showed training AUC values of 0.658 and 0.682 (test AUCs: 0.718 and 0.676), respectively. The AUCs of models trained from a different training and testing split and by different cross-validation strategies are shown in Table S5. In addition, Table S5 also shows the AUCs referring to the random forest model without the follow-up variables and logistic regression.

The five most important variables in predicting functional impairment were having been sexually abused as a child, SRQ total score, physical negligence, emotional abuse, and physical abuse through CTQ (see Figure S4 for partial dependence graphs). To visualize the contribution of the other variables in the predictions of the RF model using recursive feature elimination see Figure 3.

Table 1: Performance metrics of the random forest model after applying a variable selection algorithm in the test matrix.

Cutoff point	Balanced accuracy	Sensitivity	Specificity	PPV	NPV
0.1	0.53	1.00	0.07	0.33	1.00
0.2	0.53	0.89	0.28	0.36	0.84
0.3	**0.66**	**0.74**	**0.59**	**0.45**	**0.83**
0.4	0.63	0.63	0.64	0.45	0.79
0.5	0.56	0.37	0.76	0.42	0.72
0.6	0.57	0.30	0.84	0.47	0.72
0.7	0.55	0.19	0.91	0.50	0.71
0.8	0.50	0.04	0.97	0.33	0.68
0.9	0.49	0.00	0.98	0.00	0.68

Legend: PPV: Positive predictive value; NPV: Negative predictive value.

4. Discussion

The aim of the study was to develop a machine learning model that could predict functional impairment in patients with mood disorder in a population-based sample. The test AUC was 0.68, and, based on the metrics observed in Table 1, RF models may prove to be efficient and promising in predicting phenomena related to the global functioning of subjects with MDD and BD – considering that functional impairment is a major delimiter and indicative of the prognosis for patients with mood disorders[32]. Having suffered sexual abuse, higher SRQ scores, as well as having experienced physical neglect, emotional abuse and physical abuse were the most important variables in the predictive capacity of the final model.

The domains of sexual abuse, physical neglect, emotional abuse, physical abuse and emotional neglect of the CTQ were associated in an important way to the outcome. Robust evidence demonstrates childhood trauma as a risk factor for several lifelong mental and somatic disorders[38–40]. In MDD, childhood trauma appears to be associated with earlier onset of the disease, higher rates of comorbidities and worse treatment outcomes[41–43]. In BD patients, childhood trauma is also associated with

experiencing the first episode of the disease earlier[41], in addition to showing an increase in rapid cycling[44,4552, 53] and more severe forms of the disease[46–48]. These factors seem to contribute to a worse functional outcome in the participants. The literature shows that experiencing high and constant levels of stress in childhoodwithout having adult support triggers a continuous stress response, activated even when there is no real danger or harm[49]. This activation overwhelms developing systems, which can result in severe and lasting damage to the brain[49]. Based on our data, we can understand childhood trauma not only as a risk factor for mood disorders, but also as an enhancer of functional impairment.

Another variable with high predictive power in our model was the total score of the SRQ, an instrument used to screen non-psychotic mental disorders (depressive, anxious and somatic). This may indicate that an increase in the number of these symptoms is associated with an increase in functional impairment over time. Such impairments have been previously described in the literature by a study that assessed the presence of residual depressive symptoms in bipolar and unipolar patients inclinical remission[50]. Corroborating our data, even during the euthymic period they observed residual depressive symptoms, somatic anxiety, impact on work and activities, psychic anxiety, and somatic symptoms[50].

The literature shows that comorbidity between mood disorders and anxiety disorders is related to worse clinical outcomes[50,51]. The lifetime prevalence of some anxiety disorder in bipolar disorder is 42.7% and is associated with difficulties in diagnosing and managing the illness[52]. A study aiming to investigate clinical differences between MDD patients with and without Anxiety Disorders found a higher risk of suicide (among subjects with any anxiety disorder, panic disorder, agoraphobia, social phobia), polypsychopharmacy (among subjects with panic disorder, agoraphobia) and higher gravity of depressive symptoms (among subjects with GAD)[53]. In the direction ofthese data, the diagnoses of panic, social phobia and agoraphobia in our study also seem to contribute to a worse course and management of the disease and wereselected by the model among the most important variables in our database to predict functional impairment. Still on that line, another study showed that comorbidity with anxiety disorders seems to have a moderate impact on functional impairment[54].

The severity of depressive symptoms also plays an important role in the outcome. In the literature, for both MDD and BD, residual mood symptoms seem to be associated with functional impairment[9, 8]. Among BD patients, the occupational domain seems to be the most compromised[9], while in patients with MDD, cognitive symptoms seem to have the greatest impact on functionality[8]. Furthermore, cognitive functioning seems to have a strong impact on occupational productivity, in addition to occupational status[55]. Even employed BD patients may have difficulties in emotionally bonding with work tasks and present a habitual pattern of absence from work[9], which

would characterize the impact on occupational productivity, derived from cognitive deficits.

Another point that we would like to highlight is the selection of socioeconomic status and educational level variables. This information can be understood through previously published studies demonstrating that even BD patients with a high level of education, when compared to controls, are at a salary disadvantage and in worse jobs[56,57]. However, employed BD patients have lower scores in the FAST total score and in the autonomy, occupational and interpersonal domains compared to unemployed BD patients[58]. These data may be reflecting the importance of social involvement and continued cognitive stimulation in this population. Since quality of life includes subjective perceptions regarding social position, cultural context, value system, goals and personal expectations[59], it does not depend only on clinical remission, but also on satisfactory functional performance[60]. However, the job market can be a challenge for these subjects and such complications may be contributing to a worse functional outcome. In this sense, the variable referring to currently working on the baseline was also selected by our model to compose the variables of importance, although it was not among the most important in the predictive capacity of the final model.

It is important to note that this study has some limitations. Initially, the use of machine learning techniques requires a larger sample size to improve the model and in order to maintain the greatest sampling power possible, we chose to develop our analyses considering both MDD and BD patients, despite the existence of characteristics that are particular to each psychiatric condition. Third, we chose to consider some variables that were collected in the follow-up, since they are information regarding the life context of the participants that can be easily collected by clinicians and especially useful if included in the development of the risk calculator. Last but not least, functional impairment was not assessed at baseline, meaning that some individuals may already have impaired functionality from baseline. Therefore, the data presented should be interpreted with caution, given the potential confounding bias.

Furthermore, there are a number of remaining challenges within clinical prediction models in psychiatry. For instance, in a recent systematic review mapping the process of clinical application models in psychiatry, a substantial risk of bias was observed across 94.5% of studies, and 68.6% of those which incorporated external validation. Moreover, as the authors mention, methods such as random split-sampling in regression models, which involves random splitting a sample into training and testing sets, can occasionally provide overoptimistic or biased model accuracy[61], and as discussed elsewhere, approaches such as bootstrap-based optimism correction may be a useful alternative[62]. Additionally, methods such as stratified train-test splits can be used to partly circumvent this issue[63]. Similarly, autocorrelation, defined as the

relationship between observations at different points of time, can present challenges in classification and regression tasks, irrespective of the presence of multiple timepoints.

Finally, we developed an algorithm capable of predicting functional impairment in patients with mood disorder in up to 5 years, through data collected in a population sample. The model presented has clinical, sociodemographic and environmental data, demonstrating that it is possible to use such information to predict functional performance. Still, the model presented may be improved in the future, through other studies that can use a larger sample size, as well as integrate digital health data and biological information. Future studies may take into account the screening objective of predictive tools and collect features that are easily accessible to health professionals. Furthermore, the advancement of dimensional approaches in the understanding of severe mental disorders may be a paradigm on the rise, given the possibility of predicting risk through computational models, and considering the complexity of outcomes that are not always binary. The results of the study provide insights into some key features for the appearance of functional impairment in subjects with mood disorders in euthymia. This finding is exploratory, but it is a first step towards finding clues to sociodemographic and clinical predictors for functional impairment, given the high cost of collecting biological or genetic data in various settings.The important variables presented in this study alert health professionals about the need to be analyzed from the first evaluations. In addition, it warns about the gains that can be obtained if included in the goals and treatment strategies from the first contact with the clinician.

5. Risk calculator for functional impairment

The calculator can be accessed through the link: https://brunomontezano.shinyapps.io/functional_impairment_risk_calculator/.

6. Acknowledgements/Funding

7. Interest conflicts

ICP receives authorship royalties from Springer Nature and ArtMed. ICP has been served as a consultant, advisor, or CME speaker for the following entities: Janssen, LundBeck, Libbs, Daiichi Sankyo, EMS and Pfizer.

7. References

1. Gutiérrez-Rojas L, Porras-Segovia A, Dunne H, Andrade-González N, Cervilla JA. Prevalence and correlates of major depressive disorder: a systematic review. Braz J Psychiatry [Internet]. novembro de 2020;42(6):657–72. Disponível em: http://dx.doi.org/10.1590/1516-4446-2020-0650

2. Greer TL, Kurian BT, Trivedi MH. Defining and Measuring Functional. CNS Drugs [Internet]. 1° de abril de 2010;24(4):267–84. Disponível em: https://doi.org/10.2165/11530230-000000000-00000

3. Saragoussi D, Christensen MC, Hammer-Helmich L, Rive B, Touya M, Haro JM. Long-term follow-up on health-related quality of life in major depressive disorder: A 2-year European cohort study. Neuropsychiatr Dis Treat [Internet]. 22 de maio de 2018;14:1339–50. Disponível em: https://psycnet.apa.org/fulltext/2018-29738-001.pdf

4. Karam, Ladea, Medina-Mora. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. Arch Gen Psychiatry [Internet]. 2011; Disponível em: https://jamanetwork.com/journals/jamapsychiatry/article-abstract/211192

5. Catalá-López F, Gènova-Maleras R, Vieta E, Tabarés-Seisdedos R. The increasing burden of mental and neurological disorders. Eur Neuropsychopharmacol [Internet]. novembro de 2013;23(11):1337–9. Disponível em: http://dx.doi.org/10.1016/j.euroneuro.2013.04.001

6. Hay, Abajobir, Abate, Abbafati, Abbas. … national disability-adjusted life-years (DALYs) for 333 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990– 2016 …. Lancet [Internet]. 2017; Disponível em: https://www.sciencedirect.com/science/article/pii/S014067361732130X

7. Solé B, Bonnin CM, Jiménez E, Torrent C, Torres I, Varo C, et al. Heterogeneity of functionaloutcomes in patients with bipolar disorder: a cluster-analytic approach. Acta Psychiatr Scand [Internet]. junho de 2018;137(6):516–27. Disponível em: https://onlinelibrary.wiley.com/doi/abs/10.1111/acps.12871

8. Christensen MC, Wong CMJ, Baune BT. Symptoms of major depressive disorder and their impact on psychosocial functioning in the different phases of the disease: Do the perspectives of patients and healthcare providers differ? Front Psychiatry [Internet]. 24 de abril de 2020;11:280. Disponível em: https://www.frontiersin.org/articles/10.3389/fpsyt.2020.00280/full

9. Léda-Rêgo G, Bezerra-Filho S, Miranda-Scippa Â. Functioning in euthymic patients with bipolar disorder: A systematic review and meta-analysis using the Functioning Assessment Short Test. Bipolar Disord [Internet]. setembro de 2020;22(6):569–81. Disponível em: https://onlinelibrary.wiley.com/doi/10.1111/bdi.12904

10. Knight MJ, Baune BT. Social cognitive abilities predict psychosocial dysfunction in major depressive disorder. Depress Anxiety [Internet]. janeiro de 2019;36(1):54–62. Disponível em: http://dx.doi.org/10.1002/da.22844

11. Chan JYL, Leow SMH, Bea KT, Cheng WK, Phoong SW, Hong ZW, et al. Mitigating the Multicollinearity Problem and Its Machine Learning Approach: A Review. Sci China Ser A Math [Internet]. 12 de abril de 2022 [citado 25 de agosto de 2022];10(8):1283. Disponível em: https://www.mdpi.com/2227-7390/10/8/1283

12. Shrestha. Detecting multicollinearity in regression analysis. Am J Appl Math Stat [Internet]. Disponível em: https://www.researchgate.net/profile/Noora-Shrestha/publication/342413955_Detecting_Multicollinearity_in_Regression_Analysis/links/5eff2033458515505087a949/Detecting-Multicollinearity-in-Regression-Analysis.pdf

13. Vieira DC, de Azevedo Cardoso T, Mondin TC, Jansen K, da Silva RA, de Mattos Souza LD, et al. Mood disorders and prospective suicidality in young adults: a population-based cohort study. Acta Psychiatr Scand [Internet]. fevereiro de 2018;137(2):109–15. Disponível em:

https://onlinelibrary.wiley.com/doi/abs/10.1111/acps.12846

14. Ballester PL, Cardoso T de A, Moreira FP, da Silva RA, Mondin TC, Araujo RM, et al. 5-year incidence of suicide-risk in youth: A gradient tree boosting and SHAP study. J Affect Disord [Internet]. 1º de dezembro de 2021;295:1049–56. Disponível em: http://dx.doi.org/10.1016/j.jad.2021.08.033

15. Amorim. Mini International Neuropsychiatric Interview (MINI): validação de entrevista breve para diagnóstico de transtornos mentais. Brazilian Journal of Psychiatry [Internet]. 2000; Disponível em: https://www.scielo.br/pdf/rbp/v22n3/v22n3a01.pdf

16. Abep A. Associação Brasileira de Empresas de Pesquisa. Classificação Econômica Brasil Critério Padrão de; 2009.

17. Cunha JA. Escalas beck. São Paulo: Casa do Psicólogo; 2001.

18. Cunha. [CITAÇÃO] Manual da versão em português das Escalas Beck. São Paulo: casa do psicólogo. 2001;

19. WHO ASSIST Working Group. The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): development, reliability and feasibility. Addiction [Internet]. setembro de 2002;97(9):1183–94. Disponível em: http://dx.doi.org/10.1046/j.1360-0443.2002.00185.x

20. Henrique, De Micheli, Lacerda. Validação da versão brasileira do teste de triagem do envolvimento com álcool, cigarro e outras substâncias (ASSIST). Rev Assoc Bras Pesqui s Negr s - ABPN [Internet]. 2004; Disponível em: https://www.scielo.br/pdf/ramb/v50n2/20784.pdf

21. Beck AT, Steer RA. Beck Scale for Suscide Ideation: Manual. Psychological Corporation; 1993.

22. Gonçalves, Stein. Avaliação de desempenho do Self-Reporting Questionnaire como instrumento de rastreamento psiquiátrico: um estudo comparativo com o Structured Clinical …. Cad Saude Publica [Internet]. 2008; Disponível em: https://www.scielosp.org/article/ssm/content/raw/?resource_ssm_path=/media/assets/csp/v24n2/16.pdf

23. Soares OT, Moreno DH, Moura EC de, Angst J, Moreno RA. Reliability and validity of a Brazilian version of the Hypomania Checklist (HCL-32) compared to the Mood Disorder Questionnaire (MDQ). Braz J Psychiatry [Internet]. dezembro de 2010;32(4):416–23. Disponível em: http://dx.doi.org/10.1590/s1516-44462010000400015

24. Bernstein DP, Stein JA, Newcomb MD, Walker E, Pogge D, Ahluvalia T, et al. Development and validation of a brief screening version of the Childhood Trauma Questionnaire. Child Abuse Negl [Internet]. fevereiro de 2003;27(2):169–90. Disponível em: http://dx.doi.org/10.1016/s0145-2134(02)00541-0

25. Grassi-Oliveira, Stein, Pezzi. Tradução e validação de conteúdo da versão em português do Childhood Trauma Questionnaire. Rev Saude Publica [Internet]. 2006; Disponível em: https://www.scielo.br/pdf/rsp/v40n2/28529.pdf

26. Rosa AR, Sánchez-Moreno J, Martínez-Aran A, Salamero M, Torrent C, Reinares M, et al. Validity and reliability of the Functioning Assessment Short Test (FAST) in bipolar disorder. Clin PractEpidemiol Ment Health [Internet]. 7 de junho de 2007;3:5. Disponível em: http://dx.doi.org/10.1186/1745-0179-3-5

27. Bunn A, Korpela M. Crossdating in dplR [Internet]. cran.microsoft.com; 2014 [citado 28 de maio de 2022]. Disponível em: https://cran.microsoft.com/snapshot/2014-09-08/web/packages/dplR/vignettes/xdate-dplR.pdf

28. Kuhn M, Wickham H. recipes: Preprocessing and Feature Engineering Steps for Modeling [Internet]. 2022. Disponível em: https://github.com/tidymodels/recipes

29. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. J Stat Softw [Internet]. 12 de dezembro de 2011 [citado 28 de maio de 2022];45:1–67. Disponível em: https://www.jstatsoft.org/index.php/jss/article/view/v045i03

30. Kuhn, Wickham. [CITAÇÃO] Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles. Boston, MA, USA[(accessed on 10 December 2020)]. 2020;

31. Kuhn M. Classification and Regression Training [R package caret version 6.0-92]. 19 de abril de 2022 [citado 29 de maio de 2022]; Disponível em: https://CRAN.R-project.org/package=caret

32. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw [Internet]. 2010;33(1):1–22. Disponível em: https://www.ncbi.nlm.nih.gov/pubmed/20808728

33. Lundberg, Lee. A unified approach to interpreting model predictions. information processing systems [Internet]. 2017; Disponível em: https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html

34. Molnar C. Interpretable Machine Learning [Internet]. Lulu.com; 2020. 320 p. Disponível em: https://play.google.com/store/books/details?id=jBm3DwAAQBAJ

35. Web Application Framework for R [R package shiny version 1.7.1]. 2 de outubro de 2021 [citado 28 de maio de 2022]; Disponível em: https://CRAN.R-project.org/package=shiny

36. Create Custom "Bootstrap" Themes to Use in "Shiny" [R package fresh version 0.2.0]. 29 de maio de 2020 [citado 8 de agosto de 2022]; Disponível em: https://cran.r-project.org/package=fresh

37. Create Dashboards with "Shiny" [R package shinydashboard version 0.7.2]. 30 de setembro de 2021 [citado 8 de agosto de 2022]; Disponível em: https://cran.r-project.org/package=shinydashboard

38. Selous C, Kelly-Irving M, Maughan B, Eyre O, Rice F, Collishaw S. Adverse childhood experiences and adult mood problems: evidence from a five-decade prospective birth cohort. Psychol Med [Internet]. outubro de 2020;50(14):2444–51. Disponível em: http://dx.doi.org/10.1017/S003329171900271X

39. Hughes K, Bellis MA, Hardcastle KA, Sethi D, Butchart A, Mikton C, et al. The effect of multiple adverse childhood experiences on health: a systematic review and meta-analysis. Lancet Public Health [Internet]. agosto de 2017;2(8):e356–66. Disponível em: http://dx.doi.org/10.1016/S2468-2667(17)30118-4

40. Norman RE, Byambaa M, De R, Butchart A, Scott J, Vos T. The long-term health consequences of child physical abuse, emotional abuse, and neglect: a systematic review and meta-analysis. PLoS Med [Internet]. 27 de novembro de 2012;9(11):e1001349. Disponível em: http://dx.doi.org/10.1371/journal.pmed.1001349

41. Agnew-Blais J, Danese A. Childhood maltreatment and unfavourable clinical outcomes in bipolar disorder: a systematic review and meta-analysis. Lancet Psychiatry [Internet]. abril de 2016;3(4):342–9. Disponível em: http://dx.doi.org/10.1016/S2215-0366(15)00544-1

42. Hovens JGFM, Giltay EJ, Wiersma JE, Spinhoven P, Penninx BWJH, Zitman FG. Impact of childhood life events and trauma on the course of depressive and anxiety disorders. Acta Psychiatr Scand [Internet]. setembro de 2012;126(3):198–207. Disponível em: http://dx.doi.org/10.1111/j.1600-0447.2011.01828.x

43. Wiersma JE, Hovens JGFM, van Oppen P, Giltay EJ, van Schaik DJF, Beekman ATF, et al. The importance of childhood trauma and childhood life events for chronicity of depression in adults. J Clin Psychiatry [Internet]. julho de 2009;70(7):983–9. Disponível em: http://dx.doi.org/10.4088/jcp.08m04521

44. Suppes T, Leverich GS, Keck PE, Nolen WA, Denicoff KD, Altshuler LL, et al. The Stanley Foundation Bipolar Treatment Outcome Network. II. Demographics and illness characteristics of the first 261 patients. J Affect Disord [Internet]. dezembro de 2001;67(1-3):45–59. Disponível em: http://dx.doi.org/10.1016/s0165-0327(01)00432-3

45. Kupka RW, Luckenbaugh DA, Post RM, Suppes T, Altshuler LL, Keck PE Jr, et al. Comparison of rapid-cycling and non-rapid-cycling bipolar disorder based on prospective mood ratings in 539 outpatients. Am J Psychiatry [Internet]. julho de 2005;162(7):1273–80. Disponível em: http://dx.doi.org/10.1176/appi.ajp.162.7.1273

46. Varese F, Smeets F, Drukker M, Lieverse R, Lataster T, Viechtbauer W, et al. Childhood adversities increase the risk of psychosis: a meta-analysis of patient-control, prospective- and cross-sectional cohort studies. Schizophr Bull [Internet]. junho de 2012;38(4):661–71. Disponível em: http://dx.doi.org/10.1093/schbul/sbs050

47. Janssen I, Krabbendam L, Bak M, Hanssen M, Vollebergh W, de Graaf R, et al. Childhood abuse as a risk factor for psychotic experiences. Acta Psychiatr Scand [Internet]. janeiro de 2004;109(1):38–45. Disponível em: http://dx.doi.org/10.1046/j.0001-690x.2003.00217.x

48. Schäfer I, Fisher HL. Childhood trauma and psychosis - what is the evidence? Dialogues Clin Neurosci [Internet]. 2011;13(3):360–5. Disponível em: https://www.tandfonline.com/doi/full/10.31887/DCNS.2011.13.2/ischaefer

49. Scientific Council National. Excessive stress disrupts the development of brain architecture. J Child Serv [Internet]. 1º de janeiro de 2014;9(2):143–53. Disponível em: https://doi.org/10.1108/JCS-01-2014-0006

50. Vieta E, Sánchez-Moreno J, Lahuerta J, Zaragoza S, EDHIPO Group (Hypomania Detection Study Group). Subsyndromal depressive symptoms in patients with bipolar and unipolar disorder during clinical remission. J Affect Disord [Internet]. abril de 2008;107(1-3):169–74. Disponível em: http://dx.doi.org/10.1016/j.jad.2007.08.007

51. Cavicchioli FL, Maes M, Roomruangwong C, Bonifacio KL, Barbosa DS, Anderson G, et al. Associations between severity of anxiety and clinical and biological features of major affective disorders. Psychiatry Res [Internet]. fevereiro de 2018;260:17–23. Disponível em: http://dx.doi.org/10.1016/j.psychres.2017.11.024

52. Nabavi B, Mitchell AJ, Nutt D. A Lifetime Prevalence of Comorbidity Between Bipolar Affective Disorder and Anxiety Disorders: A Meta-analysis of 52 Interview-based Studies of Psychiatric Population. EBioMedicine [Internet]. outubro de 2015;2(10):1405–19. Disponível em: http://dx.doi.org/10.1016/j.ebiom.2015.09.006

53. Dold M, Bartova L, Souery D, Mendlewicz J, Serretti A, Porcelli S, et al. Clinical characteristics and treatment outcomes of patients with major depressive disorder and comorbid anxiety disorders - results from a European multicenter study. J Psychiatr Res [Internet]. agosto de 2017;91:1–13. Disponível em: http://dx.doi.org/10.1016/j.jpsychires.2017.02.020

54. Gili M, García Toro M, Armengol S, García-Campayo J, Castro A, Roca M. Functional impairment in patients with major depressive disorder and comorbid anxiety disorder. Can J Psychiatry [Internet]. dezembro de 2013;58(12):679–86. Disponível em: http://dx.doi.org/10.1177/070674371305801205

55. Knight MJ, Air T, Baune BT. The role of cognitive impairment in psychosocial functioning in remitted depression. J Affect Disord [Internet]. 1º de agosto de 2018;235:129–34. Disponível em: http://dx.doi.org/10.1016/j.jad.2018.04.051

56. Morselli PL, Elgie R, Cesana BM. GAMIAN-Europe/BEAM survey II: cross-national analysis of unemployment, family history, treatment satisfaction and impact of the bipolar disorder on life style. Bipolar Disord [Internet]. dezembro de 2004;6(6):487–97. Disponível em:

https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1399-5618.2004.00160.x

57. Chang HC, Huang KC, Chiu WC, Huang KC, Tang CH, Su KP. Change in employment status in bipolar disorder: a longitudinal study using national claims data. J Clin Psychiatry [Internet]. abril de 2016;77(4):e429–35. Disponível em: http://dx.doi.org/10.4088/JCP.14m09576

58. Liu X, Ma X, Wang W, Zhang J, Sun X, Luo X, et al. The functional impairment of different subtypes and occupational states in euthymic patients with bipolar disorder. BMC Psychiatry [Internet]. 7 de maio de 2021;21(1):240. Disponível em: http://dx.doi.org/10.1186/s12888-021-03242-x

59. Group W. The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. Soc Sci Med [Internet]. novembro de 1995;41(10):1403–9. Disponível em: http://dx.doi.org/10.1016/0277-9536(95)00112-k

60. Vieta E, Torrent C. Functional remediation: the pathway from remission to recovery in bipolar disorder. World Psychiatry [Internet]. outubro de 2016;15(3):288–9. Disponível em: http://dx.doi.org/10.1002/wps.20351

61. Meehan AJ, Lewis SJ, Fazel S, Fusar-Poli P, Steyerberg EW, Stahl D, et al. Clinical prediction models in psychiatry: a systematic review of two decades of progress and challenges. Mol Psychiatry [Internet]. junho de 2022;27(6):2700–8. Disponível em: http://dx.doi.org/10.1038/s41380-022-01528-4

62. Austin PC, Steyerberg EW. Events per variable (EPV) and the relative performance of different strategies for estimating the out-of-sample validity of logistic regression models. Stat Methods Med Res [Internet]. abril de 2017;26(2):796–808. Disponível em: http://dx.doi.org/10.1177/0962280214558972

63. Farias F, Ludermir T, Bastos-Filho C. Similarity Based Stratified Splitting: an approach to train better classifiers [Internet]. arXiv [cs.LG]. 2020. Disponível em: http://arxiv.org/abs/2010.06099

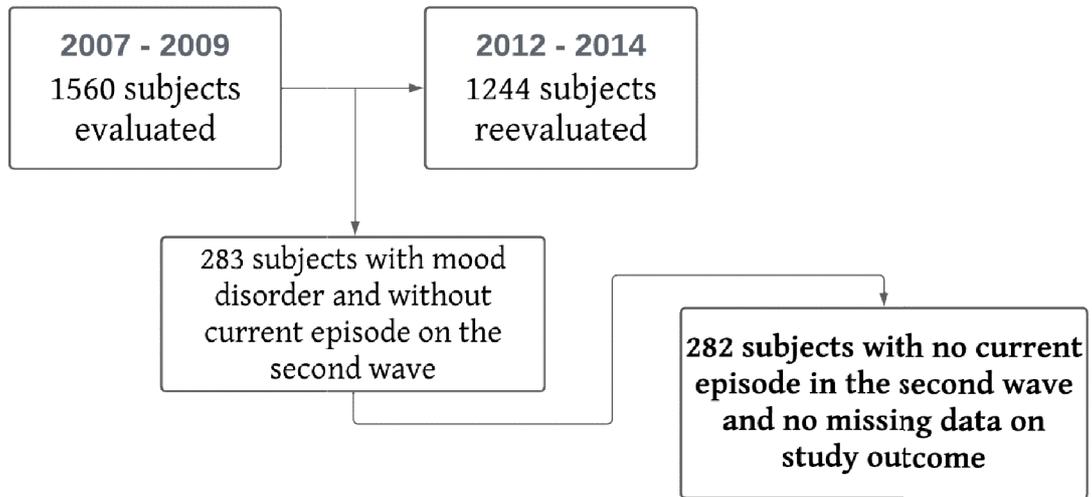**Figure 1.** Flowchart of sample selection in the first and second wave of the study.

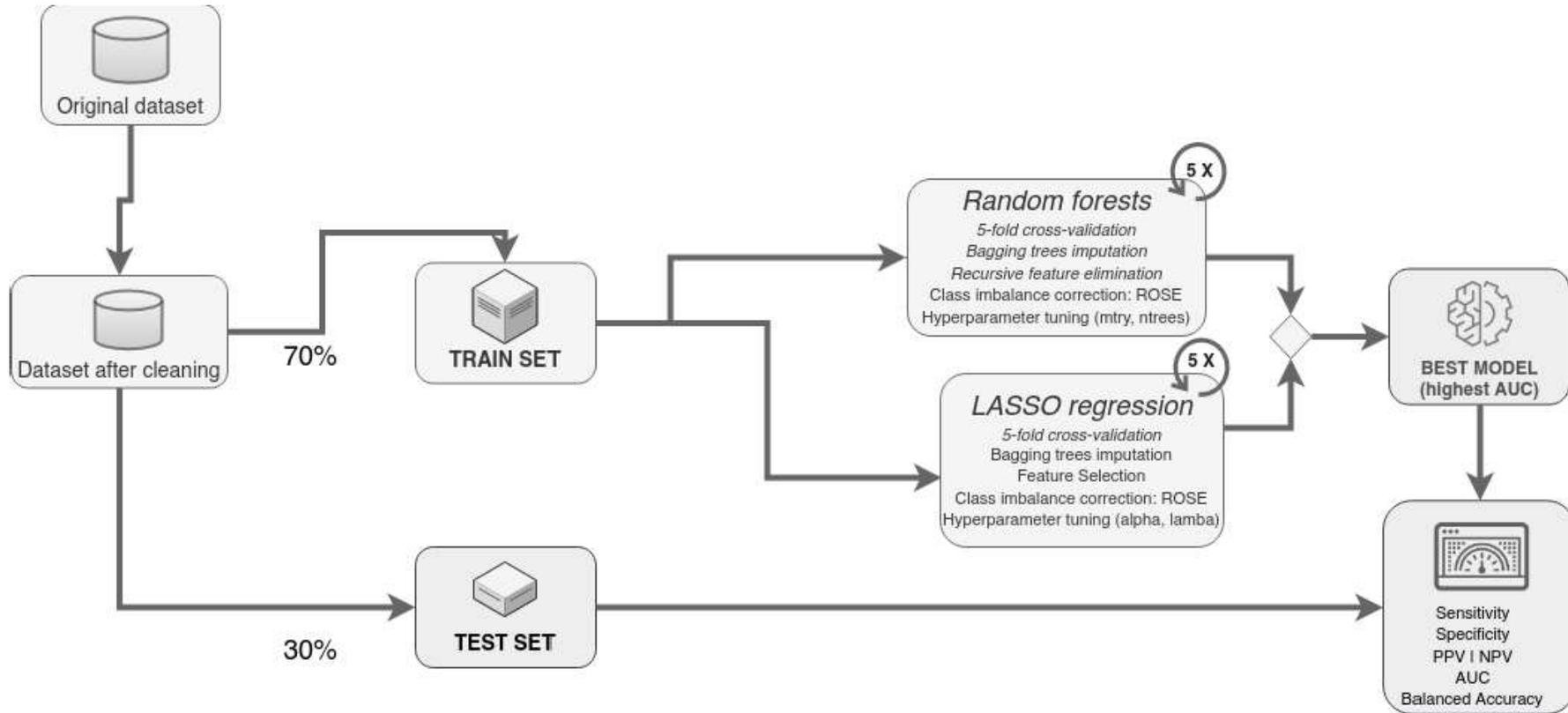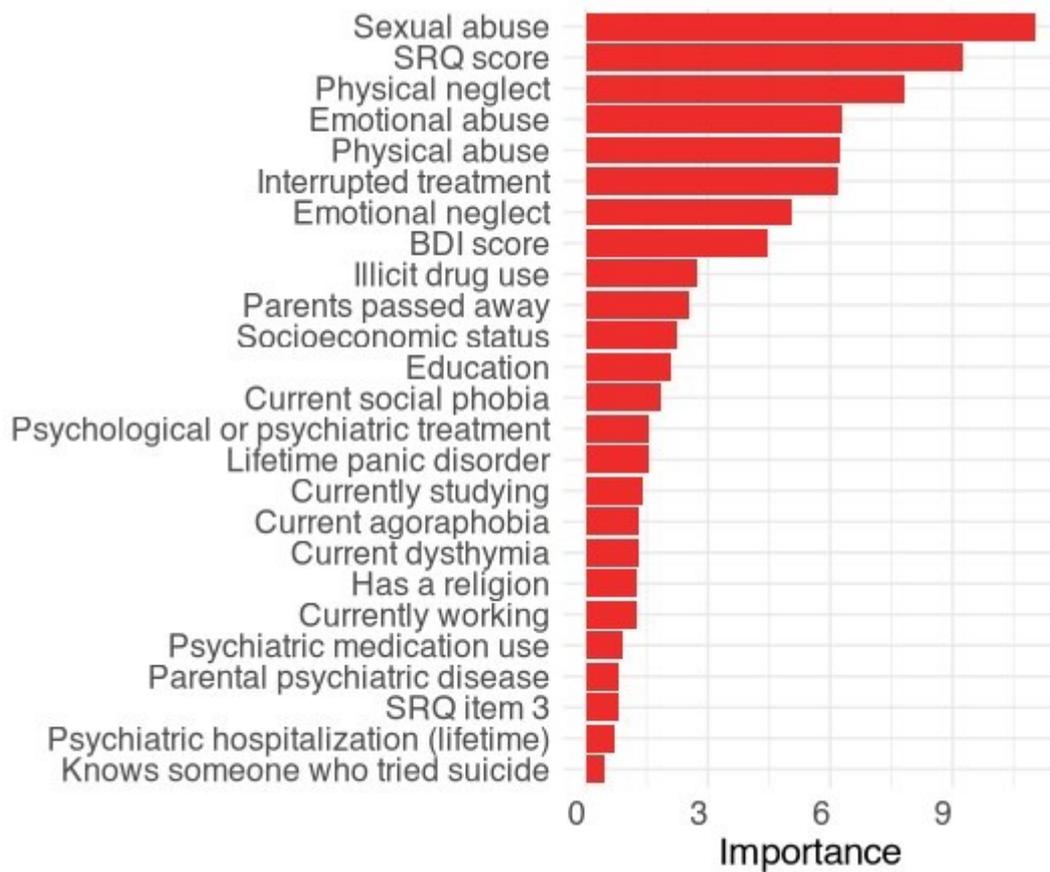**Figure 2.** Machine learning protocol.

Figure 3: Importance of variables based on Shapley values of the 25 variables of the random forest model with recursive feature elimination fitted in the test matrix.

Legend: SRQ (Self-Reporting Questionnaire); BDI (Beck's Depression Inventory).

Supplementary material list.

Name	Description
Table S1	Brief description of the variables used in the analysis.
Table S2	Descriptive data with Multivariate Imputation by Chained Equations (MICE).
Table S3	Descriptive data without Imputation.
Table S4	Absolute and relative number of missing instances.
Table S5	Supplementary analyses for the predictive model.
Table S6	Areas under the receiver operating characteristic curve for the model tested on additional subsamples and their respective outcome frequencies.
Table S7	Model performance metrics predicting in a sample of patients in current mood episode at follow-up.
Table S8	Model performance metrics predicting in a sample of patients with a psychiatric condition other than a mood disorder.
Table S9	Model performance metrics predicting in a sample of healthy participants.
Table S10	Detailed description of the instruments used in the study.
Figure S1	Baseline information added to the model.
Figure S2	Follow-up information added to the model.
Figure S3	Test set ROC of the main model.
Figure S4	Partial dependence plot of the most important variables of the main model.

Table S1: Brief description of the variables used in the analysis.

Variable	Type	Selected by RFE	Selected on follow-up
(Outcome) FAST score > 11	Categorical	Not applicable	
Skin color (dichotomy)	Categorical		
Sex	Categorical		
Age (in years)	Quantitative		
Socioeconomic status	Categorical	✓	
Years of study	Categorical		
Religion (dichotomy)	Categorical	✓	
Education (in years)	Categorical	✓	
Parent medication	Categorical		✓
Psychiatric illness of parents	Categorical	✓	✓
Siblings with psychiatric illness	Categorical		✓
Psychiatric medication	Categorical	✓	✓
Psychiatric hospitalization (lifelong)	Categorical	✓	✓
Worked for money (lifelong)	Categorical		
Currently works	Categorical	✓	
Currently studying	Categorical	✓	
Feel that you have support in life	Categorical		
Has divorced parents	Categorical		
Have ever had psychiatric or psychological treatment	Categorical	✓	
Interrupted treatment before completion	Categorical	✓	
Has a partner	Categorical		
Have you ever been involved in a	Categorical		

physical fight?

Variable	Type		
Have deceased parents	Categorical	✓	
Know someone who has attempted suicide	Categorical	✓	
Know someone who committed suicide	Categorical		
Family member with BD	Categorical		✓
Current depressive episode	Categorical		
Current melancholic depressive episode	Categorical		
Current dysthymia	Categorical	✓	
Current manic or hypomanic episode	Categorical		
Current agoraphobia	Categorical	✓	
Current social phobia	Categorical	✓	
Current obsessive compulsive disorder	Categorical		
Current post-traumatic stress disorder	Categorical		
Current generalized anxiety disorder	Categorical		
Panic disorder (lifelong)	Categorical	✓	✓
Item 3 SRQ	Categorical	✓	
Item 11 SRQ	Categorical		
Item 17 SRQ	Categorical		
BDI total score	Quantitative	✓	
SRQ total score	Quantitative	✓	
Emotional abuse (CTQ domain)	Quantitative	✓	
Physical abuse (CTQ domain)	Quantitative	✓	
Sexual abuse (CTQ domain)	Quantitative	✓	
Emotional neglect (CTQ domain)	Quantitative	✓	
Physical neglect (CTQ domain)	Quantitative	✓	

BSI Item 2	Categorical
BSI Item 5	Categorical
Moderate risk of some illicit drug	Categorical
	√
Moderate or high risk of alcohol or tobacco	Categorical

Legend: Beck Scale for Suicide Ideation (BSI); Functioning Assessment Short Test (FAST); Self Regulation Questionaire (SRQ); Childhood Trauma Questionaire (CTQ).

Table S2: Descriptive data with imputed values by bagged tree models.

Variables	Participants with mood disorder (n = 282)	Functional impairment (n = 89)	No functional impairment (n = 193)	p
Skin color				
Non-white (%)	83 (29.4)	26 (29.2)	57 (29.5)	1
Sex				
Male (%)	85 (30.1)	23 (25.8)	62 (32.1)	0.345
Age (mean (SD))	20.4 (2.06)	20.3 (2.13)	20.5 (2.03)	0.601
Socioeconomic status (%)				0.009*
Upper	99 (35.1)	21 (23.6)	78 (40.4)	
Middle	135 (47.9)	47 (52.8)	88 (45.6)	
Lower	48 (17.0)	21 (23.6)	27 (14.0)	
Studying in the current year				
Yes (%)	127 (45.0)	31 (34.8)	96 (49.7)	0.026
Religion				
Yes (%)	170 (60.3)	50 (56.2)	120 (62.2)	0.362
Education (%)				0.020*
Incomplete high school	130 (46.1)	46 (51.7)	84 (43.5)	
Complete high school	108 (38.3)	37 (41.6)	71 (36.8)	
Complete higher education	44 (15.6)	6 (6.7)	38 (19.7)	
Parents have used psychiatric medication				
Yes (%)	72 (25.5)	26 (29.2)	46 (23.8)	0.37
Parents had psychiatric disease				
Yes (%)	81 (28.7)	31 (34.8)	50 (25.9)	0.152
Siblings had psychiatric disease	50 (17.7)	17 (19.1)	33 (17.1)	0.748

	Yes (%)			
Takes psychiatric medication				
Yes (%)	110 (39.0)	38 (42.7)	72 (37.3)	0.425
Psychiatric hospitalization (lifetime)				
Yes (%)	12 (4.3)	3 (3.4)	9 (4.7)	0.767
Has worked for money				
Yes (%)	235 (83.3)	71 (79.8)	164 (85.0)	0.308
Currently working				
Yes (%)	93 (33.0)	23 (25.8)	70 (36.3)	0.098
Feel that have support in life				
Yes (%)	261 (92.6)	79 (88.8)	182 (94.3)	0.148
Has divorced parents				
Yes (%)	106 (37.6)	31 (34.8)	75 (38.9)	0.593
Had psychological or psychiatric treatment				
Yes (%)	55 (19.5)	24 (27.0)	31 (16.1)	0.037*
Interrupted treatment before completion				
Yes (%)	23 (8.2)	16 (18.0)	7 (3.6)	<0.001*
Has a partner				
Yes (%)	206 (73.0)	64 (71.9)	142 (73.6)	0.769
Engaged in a physical fight				
Yes (%)	32 (11.3)	14 (15.7)	18 (9.3)	0.153
Parents passed away				
Yes (%)	48 (17.0)	20 (22.5)	28 (14.5)	0.126
Knows someone who tried suicide				
Yes (%)	127 (45.0)	40 (44.9)	87 (45.1)	1
Knows someone who killed themselves	107 (37.9)	40 (44.9)	67 (34.7)	0.115

	Yes (%)		
Family bipolar disorder history			
Yes (%)	20 (7.1)	8 (9.0)	12 (6.2)	0.437
Major depressive episode			
Yes (%)	66 (23.4)	25 (28.1)	41 (21.2)	0.231
Melancholic depressive episode			
Yes (%)	42 (14.9)	21 (23.6)	21 (10.9)	0.008*
Dysthymia			
Yes (%)	17 (6.0)	9 (10.1)	8 (4.1)	0.068
Manic or hypomanic episode			
Yes (%)	72 (25.5)	22 (24.7)	50 (25.9)	0.882
Agoraphobia			
Yes (%)	57 (20.2)	21 (23.6)	36 (18.7)	0.354
Social phobia			
Yes (%)	18 (6.4)	9 (10.1)	9 (4.7)	0.108
Obsessive compulsive disorder			
Yes (%)	17 (6.0)	6 (6.7)	11 (5.7)	0.778
Post-traumatic stress disorder			
Yes (%)	13 (4.6)	10 (11.2)	3 (1.6)	<0.001*
Generalized anxiety disorder			
Yes (%)	37 (13.1)	10 (11.2)	27 (14.0)	0.571
Panic disorder (lifetime)			
Yes (%)	14 (5.0)	10 (11.2)	4 (2.1)	0.002*
Item 3 (SRQ)			
Yes (%)	109 (38.7)	39 (43.8)	70 (36.3)	0.024*
Item 11 (SRQ)	53 (18.8)	21 (23.6)	32 (16.6)	0.192

Yes (%)				
Item 17 (SRQ)				
Yes (%)	34 (12.1)	15 (16.9)	19 (9.8)	0.126
BDI score (mean (SD))	8.99 (8.66)	12.0 (10.4)	7.59 (7.33)	<0.001*
SRQ score (mean (SD))	5.76 (4.12)	6.91 (4.28)	5.22 (3.93)	0.001*
Emotional abuse (mean (SD))	8.55 (3.73)	10.0 (4.48)	7.88 (3.12)	<0.001*
Physical abuse (mean (SD))	7.10 (3.08)	8.02 (3.93)	6.67 (2.49)	0.003*
Sexual abuse (mean (SD))	5.78 (2.51)	6.06 (2.72)	5.65 (2.41)	0.233
Emotional neglect (mean (SD))	9.67 (4.7)	10.9 (5.39)	9.09 (4.23)	0.004*
Physical neglect (mean (SD))	6.96 (2.72)	7.66 (3.04)	6.63 (2.51)	0.005*
Item 2 (BSI)				
Yes (%)	26 (9.2)	9 (10.1)	17 (8.8)	0.846
Item 5 (BSI)				
Yes (%)	19 (6.7)	7 (7.9)	12 (6.2)	0.616
Any illicit drug				
Yes (%)	39 (13.8)	21 (23.6)	18 (9.3)	0.002*
Alcohol or tobacco use				
Yes (%)	85 (30.1)	34 (38.2)	51 (26.4)	0.052

Legend: Beck Scale for Suicide Ideation (BSI); Self-Reporting Questionnaire (SRQ); Beck Depression Inventory (BDI). Statistically significant difference (p<0.05)*

Table S3: Descriptive data without Imputation.

Variables	Participants with mood disorder (n = 282)	Functional impairment (n = 89)	No functional impairment (n = 193)	p	% of missing values
Skin color					
Non-white (%)	83 (29.4)	26 (29.2)	57 (29.5)	1	0.4
Sex					
Male (%)	85 (30.1)	23 (25.8)	62 (32.1)	0.321	0
Age (mean (SD))	20.4 (2.06)	20.3 (2.13)	20.5 (2.04)	0.594	0.7
Socioeconomic status (%)				0.016*	0
Upper	99 (35.1)	21 (23.6)	78 (40.4)		
Middle	135 (47.9)	47 (52.8)	88 (45.6)		
Lower	48 (17.0)	21 (23.6)	27 (14.0)		
Studying in the current year					
Yes (%)	127 (45.0)	31 (34.8)	96 (49.7)	0.017*	0
Religion					
Yes (%)	170 (60.3)	50 (56.2)	120 (62.2)	0.358	0
Education (%)				0.019*	2.8
Incomplete high school	125 (44.3)	44 (49.4)	81 (42.0)		
Complete high school	105 (37.2)	36 (40.4)	69 (35.8)		
Complete higher education	44 (15.6)	6 (6.7)	38 (19.7)		
Parents have used psychiatric medication					
Yes (%)	72 (25.5)	26 (29.2)	46 (23.8)	0.38	0
Parents had psychiatric disease	81 (28.7)	31 (34.8)	50 (25.9)	0.174	0

	Yes (%)				
Siblings had psychiatric disease					
Yes (%)	50 (17.7)	17 (19.1)	33 (17.1)	0.73	0
Takes psychiatric medication					
Yes (%)	110 (39.0)	38 (42.7)	72 (37.3)	0.428	0
Psychiatric hospitalization (lifetime)					
Yes (%)	12 (4.3)	3 (3.4)	9 (4.7)	0.761	0
Has worked for money					
Yes (%)	234 (83.0)	71 (79.8)	163 (84.5)	0.282	0.4
Currently working					
Yes (%)	93 (33.0)	23 (25.8)	70 (36.3)	0.094	0
Feel that have support in life					
Yes (%)	261 (92.6)	79 (88.8)	182 (94.3)	0.151	0
Has divorced parents					
Yes (%)	106 (37.6)	31 (34.8)	75 (38.9)	0.603	0
Had psychological or psychiatric treatment					
Yes (%)	53 (18.8)	24 (27.0)	29 (15.0)	0.048*	7.8
Interrupted treatment before completion					
Yes (%)	23 (8.2)	16 (18.0)	7 (3.6)	<0.001*	0
Has a partner					
Yes (%)	188 (66.7)	56 (62.9)	132 (68.4)	0.544	7.1
Engaged in a physical fight					
Yes (%)	32 (11.3)	14 (15.7)	18 (9.3)	0.137	0
Parents passed away					
Yes (%)	48 (17.0)	20 (22.5)	28 (14.5)	0.121	0

Knows someone who tried suicide				
Yes (%)	127 (45.0)	40 (44.9)	87 (45.1)	1	0
Knows someone who killed themselves				
Yes (%)	107 (37.9)	40 (44.9)	67 (34.7)	0.124	0.4
Family bipolar disorder history				
Yes (%)	20 (7.1)	8 (9.0)	12 (6.2)	0.455	0
Major depressive episode				
Yes (%)	66 (23.4)	25 (28.1)	41 (21.2)	0.225	0
Melancholic depressive episode				
Yes (%)	42 (14.9)	21 (23.6)	21 (10.9)	0.004*	0
Dysthymia				
Yes (%)	17 (6.0)	9 (10.1)	8 (4.1)	0.055	0
Manic or hypomanic episode				
Yes (%)	72 (25.5)	22 (24.7)	50 (25.9)	0.882	0
Agoraphobia				
Yes (%)	57 (20.2)	21 (23.6)	36 (18.7)	0.329	0
Social phobia				
Yes (%)	18 (6.4)	9 (10.1)	9 (4.7)	0.12	0
Obsessive compulsive disorder				
Yes (%)	17 (6.0)	6 (6.7)	11 (5.7)	0.789	0
Post-traumatic stress disorder				
Yes (%)	13 (4.6)	10 (11.2)	3 (1.6)	0.001*	0
Generalized anxiety disorder				
Yes (%)	37 (13.1)	10 (11.2)	27 (14.0)	0.564	0
Panic disorder (lifetime)				
Yes (%)	14 (5.0)	10 (11.2)	4 (2.1)	0.001*	0

Item 3 (SRQ)					
Yes (%)	109 (38.7)	39 (43.8)	70 (36.3)	0.242	0
Item 11 (SRQ)					
Yes (%)	53 (18.8)	21 (23.6)	32 (16.6)	0.176	0
Item 17 (SRQ)					
Yes (%)	34 (12.1)	15 (16.9)	19 (9.8)	0.123	0
BDI score (mean (SD))	8.99 (8.66)	12.0 (10.4)	7.59 (7.33)	<0.001*	0
SRQ score (mean (SD))	5.76 (4.12)	6.91 (4.28)	5.22 (3.93)	0.001*	0
Emotional abuse (mean (SD))	8.55 (3.73)	10.0 (4.48)	7.88 (3.12)	<0.001*	0
Physical abuse (mean (SD))	7.10 (3.08)	8.02 (3.93)	6.67 (2.49)	0.003*	0
Sexual abuse (mean (SD))	5.78 (2.51)	6.06 (2.72)	5.65 (2.41)	0.233	0
Emotional neglect (mean (SD))	9.67 (4.70)	10.9 (5.39)	9.09 (4.23)	0.004*	0
Physical neglect (mean (SD))	7.60 (3.12)	8.39 (3.51)	6.85 (2.49)	0.005*	0
Item 2 (BSI)					
Yes (%)	26 (9.2)	9 (10.1)	17 (8.8)	0.844	0
Item 5 (BSI)					
Yes (%)	19 (6.7)	7 (7.9)	12 (6.2)	0.622	0
Any ilicit drug					
Yes (%)	39 (13.8)	21 (23.6)	18 (9.3)	0.004*	0
Alcohol or tobacco use					
Yes (%)	85 (30.1)	34 (38.2)	51 (26.4)	0.055	0

Legend: Beck Depression Inventory (BDI); Beck Scale for Suicide Ideation (BSI); Self-Reporting Questionnaire (SRQ). Statistically significant difference (p<0.05)*

Table S4: Absolute and relative number of missing instances.

Variable	Absolute number of missing instances	Relative number of missing instances (%)
Suffered forced sex attempt	117	41.48
Participates of religious group	113	40.07
HCL-32 total score	77	27.30
Age in which worked for first time	55	19.50
Age in which used drugs for first time	26	9.21
Have ever had psychiatric or psychological treatment	22	7.80
Has a partner	20	7.09
Education (in years)	8	2.83
Age (in years)	2	0.70
Worked for money (lifelong)	1	0.35
Skin color (dichotomy)	1	0.35
BSI Item 1	1	0.35
Know someone who committed suicide	1	0.35
Sex	0	0.00
Socioeconomic status	0	0.00
Currently studying	0	0.00
Religion (dichotomy)	0	0.00
Parent medication	0	0.00
Psychiatric illness of parents	0	0.00
Siblings with psychiatric illness	0	0.00
Psychiatric medication	0	0.00
Psychiatric hospitalization	0	0.00

(lifelong)

Currently works	0	0.00
Feel that have support in life	0	0.00
Has separate parents	0	0.00
Interrupted treatment before completion	0	0.00
Have ever been involved in a physical fight	0	0.00
Have deceased parents	0	0.00
Know someone who has attempted suicide	0	0.00
Family member with BD	0	0.00
Current depressive episode	0	0.00
Current melancholic depressive episode	0	0.00
Current dysthymia	0	0.00
Current manic or hypomanic episode	0	0.00
Current agoraphobia	0	0.00
Current social phobia	0	0.00
Current obsessive compulsive disorder	0	0.00
Current post-traumatic stress disorder	0	0.00
Current generalized anxiety disorder	0	0.00
Panic disorder (lifelong)	0	0.00
Psychotic disorder (collected at follow-up)	0	0.00

Item 3 (SRQ)	0	0.00
Item 11 (SRQ)	0	0.00
Item 17 (SRQ)	0	0.00
BDI-II total score	0	0.00
SRQ total score	0	0.00
Emotional abuse (CTQ domain)	0	0.00
Physical abuse (CTQ domain)	0	0.00
Sexual abuse (CTQ domain)	0	0.00
Emotional neglect (CTQ domain)	0	0.00
Physical neglect (CTQ domain)	0	0.00
BSI Item 2	0	0.00
BSI Item 5	0	0.00
Moderate risk of some illicit drug	0	0.00
Moderate or high risk of alcohol or tobacco	0	0.00

Legend: Beck Scale for Suicide Ideation (BSI); Functioning Assessment Short Test (FAST); Self Regulation Questionaire (SRQ); Childhood Trauma Questionaire (CTQ).

Table S5: Supplementary analyses for the predictive model.

Analysis	Method	Cross-validation	Split (train:test)	RFE before model training	Data imputation	Training AUC
M1 - Model described in the main manuscript	Random Forests	5-fold with 5 repetitions	75:25	√	√	0.715
M2 - Model described in the main manuscript	LASSO	5-fold with 5 repetitions	75:25	-	√	0.693
M3 - Model 1 without RFE	Random Forests	5-fold with 5 repetitions	75:25	-	√	0.693
M4 - Model 1 without data imputation	Random Forests	5-fold with 5 repetitions	75:25	√	-	0.682
M5 - Model 2 without data imputation	LASSO	5-fold with 5 repetitions	75:25	-	-	0.658
M6 - Model 1 with different train/test set split	Random Forests	5-fold with 5 repetitions	50:50	√	√	0.654
M7 - Model 1 with different train/test set split and cross-validation	Random Forests	10-fold with 10 repetitions	50:50	√	√	0.677
M8 - Model 1 without follow-up variables	Random Forests	5-fold with 5 repetitions	75:25	√	√	0.706
M9 - Logistic regression	Logistic regression	5-fold with 5 repetitions	75:25	-	√	-

Table S6: Areas under the receiver operating characteristic curve for the model tested on additional subsamples and their respective outcome frequencies.

Subsample	Functional Impairment		AUC (confidence interval)
	Yes (%)	No (%)	
Current mood episode at follow-up (n=143)	94 (65.7)	49 (34.3)	0.67 (058 - 0.77)
Psychiatric condition other than a mood disorder (n=92)	33 (35.9)	59 (64.1)	0.76 (0.65 - 0.86)
Healthy participants (n=718)	117 (16.3)	602 (83.8)	0.63 (0.58 - 0.68)

Table S7. Model performance metrics predicting in a sample of patients in current mood episode at follow-up.

Cutoff point	Balanced accuracy	Sensitivity	Specificity	PPV	NPV
0.1	0.53	1.00	0.06	0.67	1.00
0.2	0.60	0.97	0.22	0.71	0.79
0.3	**0.61**	**0.89**	**0.33**	**0.72**	**0.62**
0.4	0.61	0.84	0.39	0.72	0.56
0.5	0.59	0.70	0.47	0.72	0.45
0.6	0.61	0.52	0.69	0.77	0.43
0.7	0.58	0.33	0.84	0.79	0.39
0.8	0.57	0.18	0.96	0.89	0.38
0.9	0.50	0.00	1.00	NA	0.34

Legend: PPV: Positive predictive value; NPV: Negative predictive value.

Table S8. Model performance metrics predicting in a sample of patients with a psychiatric condition other than a mood disorder.

Cutoff point	Balanced accuracy	Sensitivity	Specificity	PPV	NPV
0.1	0.54	1.00	0.08	0.38	1.00
0.2	0.61	0.94	0.29	0.42	0.89
0.3	0.70	0.88	0.53	0.51	0.89
0.4	**0.71**	**0.76**	**0.66**	**0.56**	**0.83**
0.5	0.67	0.55	0.80	0.60	0.76
0.6	0.60	0.33	0.86	0.58	0.70
0.7	0.55	0.18	0.92	0.55	0.67
0.8	0.51	0.03	0.98	0.50	0.64
0.9	0.50	0.00	1.00	NA	0.64

Legend: PPV: Positive predictive value; NPV: Negative predictive value.

Table S9. Model performance metrics predicting in a sample of healthy participants.

Cutoff point	Balanced accuracy	Sensitivity	Specificity	PPV	NPV
0.1	0.55	0.89	0.20	0.18	0.90
0.2	**0.60**	**0.58**	**0.62**	**0.23**	**0.88**
0.3	0.57	0.37	0.77	0.24	0.86
0.4	0.56	0.23	0.88	0.27	0.85
0.5	0.53	0.13	0.94	0.28	0.85
0.6	0.52	0.07	0.97	0.28	0.84
0.7	0.51	0.03	0.99	0.33	0.84
0.8	0.50	0.01	1.00	1.00	0.84
0.9	0.50	0.00	1.00	NA	0.84

Legend: PPV: Positive predictive value; NPV: Negative predictive value.

Table S10. Detailed description of the instruments used in the study.

Instruments	Detailed Description
Mini International Neuropsychiatric Interview (MINI Plus)	The MINI-Plus is a brief standardized diagnostic interview, compatible with the DSM-IV and ICD-10 criteria. Intended for the in-depth assessment of mental disorders throughoutlife, it systematically explores all inclusion and exclusion criteria and the chronology of 23 diagnostic categories from the DSM-IV. In this study, the following categories were considered: (A) MDD, (B) Dysthymia, (D) BD I - II, (E) Panic, (F) Agoraphobia, (G) Social Phobia, (I) Obsessive Compulsive Disorder (OCD), (J) Post Traumatic Stress Disorder (PTSD), (M) Psychotic Disorder and (P) Generalized Anxiety Disorder (GAD).
Associação Brasileira das Empresas de Pesquisas (ABEP)	Structured questionnaire with questions regarding gender, age, marital status, work, education and socioeconomic level.
Beck Depression Inventory (BDI)	This scale is composed of 21 items that investigate the presence and intensity of depressive symptoms. The sum of all items gives the total score, and the higher the score, the greater the severity of depressive symptoms. This instrument was validated and adapted for the Brazilian population.
Alcohol, Smoking and Substance Involvement Screening Test (ASSIST)	It is a structured questionnaire composed of eight questions about the use of nine classes of psychoactive substances (tobacco, alcohol, marijuana, cocaine, stimulants, sedatives, inhalants, hallucinogens and opiates). Frequency of substance use is verified currently and in the previous three months. In addition, problems related to substance use, concerns about use by people close to the user, functional impairment, frustrated attempts to stop or reduce use, the feeling of binge and use via injection were also evaluated. This instrument was adapted and validated for the Brazilian population. For alcohol, a value greater than or equal to 4 was considered moderate, while for the other drug classes the cut-off point was 11. From the previously mentioned classes, two variables were created and these were added to the data set to be analyzed. The first variable assumes a value of 1 if the individual is at moderate risk for alcohol or tobacco. While the second variable assumes a value of 1, if the individual is at moderate risk in any other class of psychoactive substances.
Beck Scale for Suicide Ideation (BSI)	The BSI scale was translated and validated for use in Brazil with a reliability of 0.90. The BSI assesses the presence and

intensity of suicidal thoughts one week before the assessment through 19 items, divided into three parts: risk and presence of suicidal ideation; severity of suicidal ideation, attitude and plans; severity of the intention to die and the number of suicide attempts. Its score can range from 0 to 38, where the higher the score, the higher the clinical risk. For each item, the participant has the following response options: none, weak or moderate to strong, with weights from 0 to 2 respectively. Additionally, four items were individually inserted into the model: "Desire to live";"I wish to die"; "I want to kill myself"; "If your life were at risk, would you try to save yourself?".

Self-regulation questionnaire (SRQ-20)	This instrument consists of 20 items that seek to screen for non-psychotic mental disorders through affirmative and negative response options. Affirmative answers obtain a value equal to 1, the total score is obtained through the sum of these values, with 0 = no probability of the presence of non-psychotic disorder and 20 = extreme possibility. In this study, in addition to considering the total score, we added the following items separately to the model: "(3) Do you regularly have a bad night's sleep?"; "(11) Do you find it difficult to enjoy your daily activities?"; "(17) Has the thought of ending your life ever crossed your mind?".
Hypomania Checklist-32 (HCL-32)	This instrument was validated after the study, however, no substantial changes were made. The HCL-32 is a self-administered instrument composed of 32 yes/no questions, internationally validated. When developed, its objective was to identify hypomanic features in depressive patients, aiming to facilitate the diagnosis of BD-II. Participants are instructed to focus on a period when they felt "up" and indicate specific thoughts or emotions during that period. In addition, the scale has 8 severity and functional impact items associated with the duration of the episodes. At the end of the questionnaire, participants are asked to rate the degree of impact (positive, no impact, negative or neutral) on family, social, school and leisure life.
Childhood Trauma Questionnaire (CTQ)	This is a self-administered instrument adapted for Brazil, for use with adolescents (from 12 years old) and adults, investigating the history of abuse and neglect during childhood. Through this questionnaire, five traumatic components can be evaluated: physical abuse, emotional abuse, sexual abuse, physical neglect and emotional neglect. It consists of 28 statements related to childhood situations on a five-point Likert scale. This instrument was

incorporated in the study from 2012 to 2014, but, since the questions are childhood related, we chose to include them in the possible predictor variables.

Figure S1. Baseline information added to the model.

Do you have any religion?	Do you feel like you have support in life?	Has your father or mother passed away?
Do you participate in religious groups?	Do you have separate parents?	Do you now someone who has committed suicide?
Are you currently studying?	Have you ever had any psychological or psychiatric treatment?	Does your father or mother have/had a psychiatric diagnosis of BD?
At what age have you started working?	Have you ever started psychological treatment, but stopped abruptly?	Are you currently working?
Have you ever done paid work?	At what age, in years, did you try any substance for the first time, including alcohol and tobacco?	Have you ever been physically forced to have sex?
Do you have a steady partner?	Have you ever gotten into a fight involving physical aggression?	MDD, Dysthymia, BD I - II, Panic, Agoraphobia, Social Phobia, OCD, PTSD, GAD through MINI[7]

Figure S2. Follow-up information added to the model.

Has your father and/or mother ever been admitted to a psychiatric hospital?	Does your father and/or mother take psychiatric medication?	Does anyone in your family have a psychiatric diagnosis of BD?
Have you ever been admitted to a psychiatric hospital?	Does your father and/or mother have a diagnosis of a psychiatric illness?	Do you take psychiatric medication?
Has your father and/or mother ever attempted suicide?	Do you have siblings with psychiatric illnesses?	Lifelong psychotic disorder and panic through MINI[7]

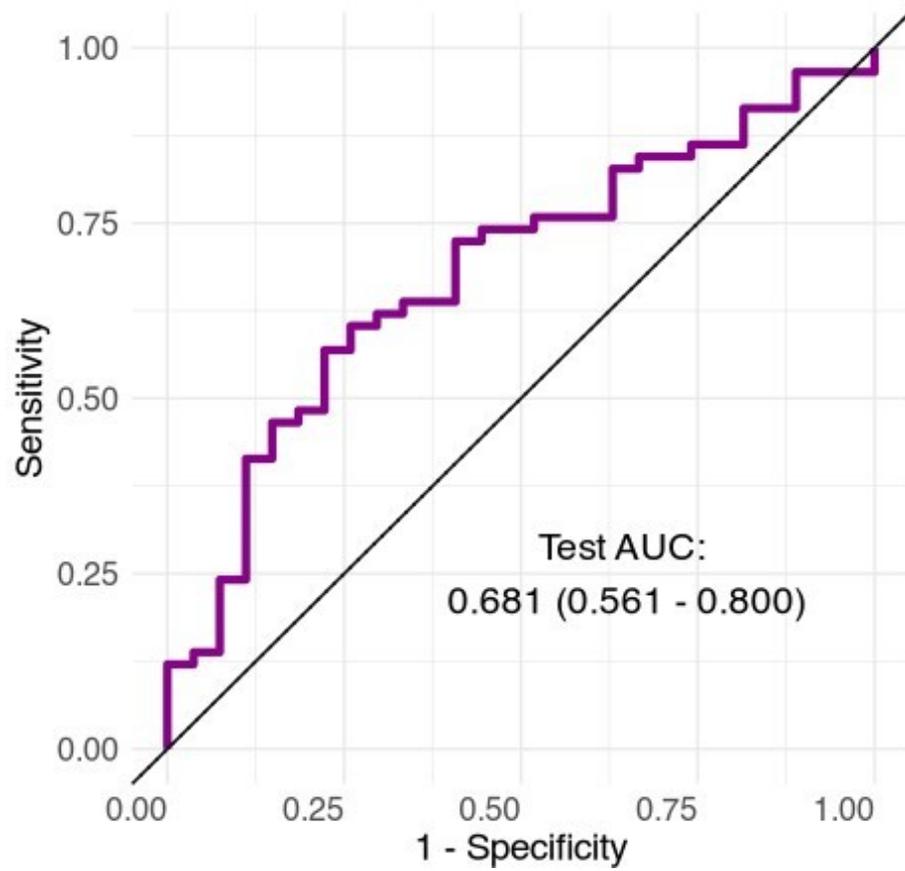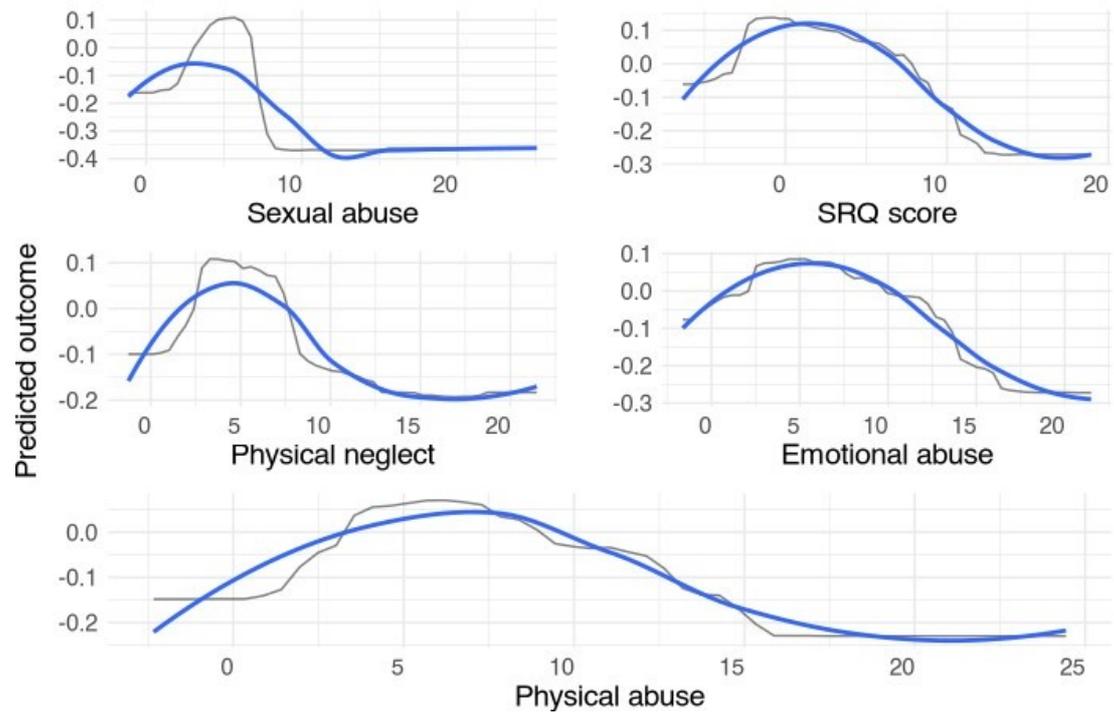**Figure S3.** The test AUC of the Random Forest model.

Figure S4: Partial dependence graphs of the five most important variables in predicting functional impairment.

CONSIDERAÇÕES FINAIS

A presente dissertação tem como resultado final o desenvolvimento de um algoritmo capaz de prever o comprometimento funcional em pacientes com transtorno de humor em até 5 anos, por meio de dados coletados em uma amostra populacional. O modelo apresentado possui dados clínicos, sociodemográficos e ambientais, demonstrando que é possível utilizar tais informações para predizer o desempenho funcional. Os resultados apresentados fornecem insights sobre algumas características fundamentais para o aparecimento de comprometimento funcional em indivíduos com transtornos de humor em eutimia, entre elas: abuso sexual na infância, maiores escores no SRQ, negligência física, abuso emocional e abuso físico na infância.

Os dados apresentados devem ser interpretados levando em consideração alguns aspectos importantes do estudo. Primeiro, análises de aprendizado de máquina requerem grandes números de observações. Apesar do nosso estudo contar com um tamanho amostral superior ao de outros trabalhos que buscaram identificar fatores associados ao prejuízo funcional em pacientes com transtorno de humor[48,49], estudos futuros que possam utilizar um tamanho amostral superior ainda são necessários, por oferecerem a possibilidade de desenvolver modelos mais acurados. Outro ponto importante é que não tínhamos a avaliação da funcionalidade dos participantes no *baseline* porque o FAST foi introduzido ao estudo apenas no acompanhamento. No entanto, em nossa amostra os sujeitos com e sem prejuízo funcional não diferiam estatisticamente com relação a estar trabalhando atualmente e já ter realizado trabalho remunerado no passado. Essas informações podem refletir aspectos da funcionalidade dos participantes na avaliação de linha de base.

O modelo apresentado deve ser entendido como um primeiro passo dentro da temática de predição da funcionalidade. Os dados apresentados devem ser sofisticados no futuro por meio de trabalhos que além de utilizar um tamanho amostral superior possam integrar dados digitais de saúde, informações biológicas, genéticas e metabólicas. As 25 variáveis de importância apresentadas no artigo oriundo dessa dissertação alertam os profissionais de saúde sobre a necessidade de serem analisadas desde as primeiras avaliações, mas de forma alguma buscam

substituir a avaliação clínica. Ainda, a calculadora divulgada através desse trabalho tem como objetivo a divulgação acadêmica e não deve ser generalizada para prática clínica. Ainda, nosso modelo foi desenvolvido a partir de dados coletados em adultos jovens, o que limita a generalização dos resultados para indivíduos de outras faixas etárias com diagnóstico de transtornos de humor.

Por fim, o modelo foi aplicado a diferentes subamostras e obteve uma AUC acima de 0.70 apenas no grupo de pacientes com outros transtornos psiquiátricos (AUC 0.76). Ainda que os prejuízos funcionais difiram em relação a níveis e áreas mais afetadas dentro de diferentes diagnósticos e que esse seja um dado exploratório,nosso modelo parece ser promissor em identificar variáveis associadas ao comprometimento funcional, mesmo em diferentes diagnósticos psiquiátricos. Nasamostras de sujeitos saudáveis e em episódio atual de transtorno de humor o modelo apresentou AUC de 0.63 e 0.67 respectivamente.

REFERÊNCIAS

1. Amorim P. Mini International Neuropsychiatric Interview (MINI): validation of a short structured diagnostic psychiatric interview. Braz J Psychiatry [Internet]. 2000 Sep [cited 2022 Jul 8];22(3):106–15. Available from: https://www.scielo.br/j/rbp/a/8BcLbLX3QHtPMsxjtFCHqKb/abstract/?format=html&lang=en&stop=previous

2. Merikangas KR, Jin R, He JP, Kessler RC, Lee S, Sampson NA, et al. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. Arch Gen Psychiatry [Internet]. 2011;68(3):241–51. Available from: https://jamanetwork.com/journals/jamapsychiatry/article-abstract/211192

3. Vieta E, Salagre E, Grande I, Carvalho AF, Fernandes BS, Berk M, et al. Early Intervention in Bipolar Disorder. Am J Psychiatry [Internet]. 2018 May 1;175(5):411–26. Available from: http://dx.doi.org/10.1176/appi.ajp.2017.17090972

4. Buist-Bouwman MA, Ormel J, de Graaf R, de Jonge P, van Sonderen E, Alonso J, et al. Mediators of the association between depression and role functioning. Acta Psychiatr Scand [Internet]. 2008 Dec;118(6):451–8. Available from: http://dx.doi.org/10.1111/j.1600-0447.2008.01285.x

5. Grande I, Berk M, Birmaher B, Vieta E. Bipolar disorder. Lancet [Internet]. 2016 Apr 9;387(10027):1561–72. Available from: http://dx.doi.org/10.1016/S0140-6736(15)00241-X

6. Christensen MC, Wong CMJ, Baune BT. Symptoms of major depressive disorder and their impact on psychosocial functioning in the different phases of the disease: Do the perspectives of patients and healthcare providers differ? Front Psychiatry [Internet]. 2020 Apr 24;11:280. Available from: https://www.frontiersin.org/articles/10.3389/fpsyt.2020.00280/full

7. Levy B, Manove E, Weiss RD. Recovery of cognitive functioning in patients with co-occurring bipolar disorder and alcohol dependence during early remission from an acute mood episode. Ann Clin Psychiatry [Internet]. 2012 May;24(2):143–54. Available from: https://www.ncbi.nlm.nih.gov/pubmed/22563570

8. Sanchez-Moreno J, Martinez-Aran A, Tabarés-Seisdedos R, Torrent C, Vieta E, Ayuso-Mateos JL. Functioning and disability in bipolar disorder: an extensive review. Psychother Psychosom [Internet]. 2009 Jul 11;78(5):285–97. Available from: http://dx.doi.org/10.1159/000228249

9. McIntyre RS, Soczynska JZ, Woldeyohannes HO, Alsuwaidan MT, Cha DS, Carvalho AF, et al. The impact of cognitive impairment on perceived workforce performance: results from the International Mood Disorders Collaborative Project. Compr Psychiatry [Internet]. 2015 Jan;56:279–82. Available from: http://dx.doi.org/10.1016/j.comppsych.2014.08.051

10. Bortolato B, Carvalho AF, McIntyre RS. Cognitive dysfunction in major depressive disorder: a state-of- the-art clinical review. CNS Neurol Disord Drug Targets [Internet]. 2014;13(10):1804–18. Available from: http://dx.doi.org/10.2174/1871527313666141130203823

11. Cambridge OR, Knight MJ, Mills N, Baune BT. The clinical relationship between cognitive impairment and psychosocial functioning in major depressive disorder: A systematic review. Psychiatry Res [Internet]. 2018 Nov;269:157–71. Available from: http://dx.doi.org/10.1016/j.psychres.2018.08.033

12. Olesen J, Gustavsson A, Svensson M, Wittchen HU, Jönsson B, CDBE2010 study group, et al. The economic cost of brain disorders in Europe. Eur J Neurol [Internet]. 2012 Jan;19(1):155–62. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-1331.2011.03590.x

13. Mintz J, Mintz LI, Arruda MJ, Hwang SS. Treatments of depression and the functional capacity to work. Arch Gen Psychiatry [Internet]. 1992 Oct;49(10):761–8. Available from: http://dx.doi.org/10.1001/archpsyc.1992.01820100005001

14. Balestrieri M, Isola M, Bonn R, Tam T, Vio A, Linden M, et al. Validation of the Italian version of Mini-ICF-APP, a short instrument for rating activity and participation restrictions in psychiatric disorders. Epidemiol Psychiatr Sci [Internet]. 2013 Mar;22(1):81–91. Available from: http://dx.doi.org/10.1017/S2045796012000480

15. Rosburg T, Kunz R, Trezzini B, Schwegler U, Jeger J. The assessment of capacity limitations in psychiatric work disability evaluations by the social functioning scale Mini-ICF-APP. BMC Psychiatry [Internet]. 2021 Sep 30;21(1):480. Available from: http://dx.doi.org/10.1186/s12888-021-03467-w

16. Collard RM, Wassink-Vossen S, Schene AH, Naarding P, Verhaak P, Oude Voshaar RC, et al. Symptomatic and functional recovery in depression in later life. Soc Psychiatry Psychiatr Epidemiol [Internet]. 2018 Oct;53(10):1071–9. Available from: http://dx.doi.org/10.1007/s00127-018-1540-z

17. Léda-Rêgo G, Bezerra-Filho S, Miranda-Scippa Â. Functioning in euthymic patients with bipolar disorder: A systematic review and meta-analysis using the Functioning Assessment Short Test. Bipolar Disord [Internet]. 2020 Sep;22(6):569–81. Available from: http://dx.doi.org/10.1111/bdi.12904

18. Iancu SC, Wong YM, Rhebergen D, van Balkom AJLM, Batelaan NM. Long-term disability in major depressive disorder: a 6-year follow-up study. Psychol Med [Internet]. 2020 Jul;50(10):1644–52. Available from: http://dx.doi.org/10.1017/S0033291719001612

19. Levy B, Manove E. Functional outcome in bipolar disorder: the big picture. Depress Res Treat [Internet]. 2012;2012:949248. Available from: http://dx.doi.org/10.1155/2012/949248

20. Henry C, Etain B, Godin O, Dargel AA, Azorin JM, Gard S, et al. Bipolar patients referred to specialized services of care: Not resistant but impaired by sub-syndromal symptoms. Results from the FACE-BD cohort. Aust N Z J Psychiatry [Internet]. 2015 Oct;49(10):898–905. Available from: http://dx.doi.org/10.1177/0004867415585582

21. Martinez-Aran A, Vieta E, Torrent C, Sanchez-Moreno J, Goikolea JM, Salamero M, et al. Functional outcome in bipolar disorder: the role of clinical and cognitive factors. Bipolar Disord [Internet]. 2007 Feb;9(1-2):103–13. Available from: http://doi.wiley.com/10.1111/j.1399-5618.2007.00327.x

22. Suppes T, Leverich GS, Keck PE, Nolen WA, Denicoff KD, Altshuler LL, et al. The Stanley Foundation Bipolar Treatment Outcome Network. II. Demographics and illness characteristics of the first 261 patients. J Affect Disord [Internet]. 2001 Dec;67(1-3):45–59. Available from: http://dx.doi.org/10.1016/s0165-0327(01)00432-3

23. Tohen M, Hennen J, Zarate CM Jr, Baldessarini RJ, Strakowski SM, Stoll AL, et al. Two-year syndromal and functional recovery in 219 cases of first-episode major affective disorder with psychotic features. Am J Psychiatry [Internet]. 2000 Feb;157(2):220–8. Available from: http://dx.doi.org/10.1176/appi.ajp.157.2.220

24. Nwanganga F, Chapple M. Practical Machine Learning in R [Internet]. John Wiley & Sons; 2020. 464 p. Available from: https://play.google.com/store/books/details?id=g1DYDwAAQBAJ

25. Conway D, White JM. Machine Learning for Hackers: Case Studies and Algorithms to Get You Started [Internet]. "O'Reilly Media, Inc."; 2012. 324 p. Available from: https://play.google.com/store/books/details?id=2VY6Bbqd0F8C

26. Lantz B. Machine Learning with R: Expert techniques for predictive modeling, 3rd Edition [Internet]. Packt Publishing Ltd; 2019. 458 p. Available from: https://play.google.com/store/books/details?id=iNuSDwAAQBAJ

27. Shalev-Shwartz S, Ben-David S. Understanding machine learning: From theory to algorithms [Internet]. Cambridge university press; 2014. Available from: http://103.47.12.35/bitstream/handle/1/1069/understanding-machine-learning-theory-algorithms.pdf?sequence=1&isAllowed=y

28. Chan JYL, Leow SMH, Bea KT, Cheng WK, Phoong SW, Hong ZW, et al. Mitigating the Multicollinearity Problem and Its Machine Learning Approach: A Review. Sci China Ser A Math [Internet]. 2022 Apr 12 [cited 2022 Aug 25];10(8):1283. Available from: https://www.mdpi.com/2227-7390/10/8/1283

29. [No title] [Internet]. [cited 2022 Sep 23]. Available from: https://www.researchgate.net/profile/Noora-Shrestha/publication/342413955_Detecting_Multicollinearity_in_Regression_Analysis/links/5eff203345 8515505087a949/Detecting-Multicollinearity-in-Regression-Analysis.pdf

30. Brasil, Filiadas. ABEP: Associação Brasileira de Empresas de Pesquisa. http://www abep org Acesso em.

31. Cunha JA. Psicodiagnóstico - V [Internet]. Artmed Editora; 2009. Available from: https://play.google.com/store/books/details?id=KwcrFmZ4mvgC

32. Cunha. [CITAÇÃO] Manual da versão em português das Escalas Beck. São Paulo: casa do psicólogo. 2001;

33. WHO ASSIST Working Group. The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): development, reliability and feasibility. Addiction [Internet]. 2002 Sep;97(9):1183–94. Available from: http://dx.doi.org/10.1046/j.1360-0443.2002.00185.x

34. Henrique, De Micheli, Lacerda. Validação da versão brasileira do teste de triagem do envolvimento com álcool, cigarro e outras substâncias (ASSIST). Rev Assoc Bras Pesqui s Negr s - ABPN [Internet]. 2004; Available from: https://www.scielo.br/pdf/ramb/v50n2/20784.pdf

35. Beck AT, Steer RA. Beck Scale for Suscide Ideation: Manual. Psychological Corporation; 1993.

36. Gonçalves, Stein. Avaliação de desempenho do Self-Reporting Questionnaire como instrumento de rastreamento psiquiátrico: um estudo comparativo com o Structured Clinical …. Cad Saude Publica [Internet]. 2008; Available from: https://www.scielosp.org/article/ssm/content/raw/?resource_ssm_path=/media/assets/csp/v24n2/16.pdf

37. Soares OT, Moreno DH, Moura EC de, Angst J, Moreno RA. Reliability and validity of a Brazilian version of the Hypomania Checklist (HCL-32) compared to the Mood Disorder Questionnaire (MDQ). Braz J Psychiatry [Internet]. 2010 Dec;32(4):416–23. Available from: http://dx.doi.org/10.1590/s1516-44462010000400015

38. Grassi-Oliveira, Stein, Pezzi. Tradução e validação de conteúdo da versão em português do Childhood Trauma Questionnaire. Rev Saude Publica [Internet]. 2006; Available from: https://www.scielo.br/pdf/rsp/v40n2/28529.pdf

39. Rosa AR, Sánchez-Moreno J, Martínez-Aran A, Salamero M, Torrent C, Reinares M, et al. Validity and reliability of the Functioning Assessment Short Test (FAST) in bipolar disorder. Clin Pract Epidemiol Ment Health [Internet]. 2007 Jun 7;3:5. Available from: http://dx.doi.org/10.1186/1745-0179-3-5

40. Kuhn, Wickham. [CITAÇÃO] Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles. Boston, MA, USA[(accessed on 10 December 2020)]. 2020;

41. Menardi G, Torelli N. Training and assessing classification rules with imbalanced data. Data Min Knowl Discov [Internet]. 2014 Jan 1;28(1):92–122. Available from: https://doi.org/10.1007/s10618-012-0295-5

42. Breiman L. Random Forests. Mach Learn [Internet]. 2001 Oct 1;45(1):5–32. Available from: https://doi.org/10.1023/A:1010933404324

43. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw [Internet]. 2010;33(1):1–22. Available from: https://www.ncbi.nlm.nih.gov/pubmed/20808728

44. Lundberg, Lee. A unified approach to interpreting model predictions. information processing systems [Internet]. 2017; Available from: https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html

45. Web Application Framework for R [R package shiny version 1.7.2]. 2022 Jul 19 [cited 2022 Sep 24]; Available from: https://CRAN.R-project.org/package=shiny

46. Create Custom "Bootstrap" Themes to Use in "Shiny" [R package fresh version 0.2.0]. 2020 May 29 [cited 2022 Sep 24]; Available from: https://cran.r-project.org/package=fresh

47. Create Dashboards with "Shiny" [R package shinydashboard version 0.7.2]. 2021 Sep 30 [cited 2022 Sep 24]; Available from: https://cran.r-project.org/package=shinydashboard

48. Knight MJ, Baune BT. Social cognitive abilities predict psychosocial dysfunction in major depressive disorder. Depress Anxiety [Internet]. 2019 Jan;36(1):54–62. Available from: http://dx.doi.org/10.1002/da.22844

49. Sanchez-Moreno J, Bonnin CM, González-Pinto A, Amann BL, Solé B, Balanzá-Martinez V, et al. Factors associated with poor functional outcome in bipolar disorder: sociodemographic, clinical, and neurocognitive variables. Acta Psychiatr Scand [Internet]. 2018 Aug;138(2):145–54. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/acps.12894