

Using Wearable Device and Artificial Intelligence to Predict Mood Symptoms in Bipolar Disorder

Yi-Ling Chien

Submitted to: Transfer Hub (manuscript eXchange)
on: August 28, 2024

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 4

Supplementary Files..... 14

Figures 15

Figure 1..... 16

Figure 2..... 17

Figure 3..... 18

Figure 4..... 19

Using Wearable Device and Artificial Intelligence to Predict Mood Symptoms in Bipolar Disorder

Yi-Ling Chien¹

¹Department of Psychiatry, National Taiwan University Hospital, Taipei, Taiwan Taipei TW

Corresponding Author:

Yi-Ling Chien

Department of Psychiatry, National Taiwan University Hospital, Taipei, Taiwan

Department of Psychiatry, National Taiwan University Hospital and College of Medicine, No.7, Chung-Shan South Road
Taipei
TW

Abstract

Bipolar disorder (BD) is a highly recurrent disorder. Early detection, early intervention, and prevention of recurrent bipolar mood symptoms are key for better prognosis. In this study, we build prediction models for bipolar disorder with machine learning algorithms. This study recruited 24 participants with BD. The Beck Depression Inventory (BDI) and Young Mania Rating Scale (YMRS) were used to evaluate depressive and manic episodes respectively. Using digital biomarkers collected from wearable devices as input, six machine learning algorithms (Logistic Regression, Decision Tree, K-Nearest Neighbors, Random Forest, Adaptive Boosting, and Extreme Gradient Boosting) were used to build predictive models. The prediction model for depressive symptoms achieved 83% accuracy, 0.89 Area Under the Receiver Operating Characteristic curve (AUROC), and 0.65 F1 score on testing data. The prediction model for manic symptoms achieved 91% accuracy, 0.88 AUROC, and 0.25 F1 score on testing data. With the interpretable model Shapely Additive exPlanations (SHAP), we found that relatively high resting heart rate, low activity, and lack of sleep may predict depressive symptoms. This study demonstrated that digital biomarkers could be used to predict depressive and manic symptoms. Moreover, based on the findings from the prediction model, we may provide clinical assessment and treatment earlier to prevent a recurrence.

(JMIR Preprints 28/08/2024:65857)

DOI: <https://doi.org/10.2196/preprints.65857>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in http://www.jmir.org/preprint/65857

Original Manuscript

Original Paper

Ding-Shan Liu¹, MA; Chia-Tung Wu¹, PhD; I-Ming Chen², MD, PhD; Ming-Hsien Hsieh², MD, PhD; Chen-Chun Lu³, MA; Yi-Ling Chien², MD, PhD; Feipei Lai^{1,3}, PhD

¹ Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan

² Department of Psychiatry, National Taiwan University Hospital, Taipei, Taiwan

³ Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei, Taiwan

Corresponding Author:

Yi-Ling Chien, MD, PhD

Department of Psychiatry, National Taiwan University Hospital and College of Medicine, No.7, Chung-Shan South Road, Taipei 10002, Taiwan

Phone: +886-2-2312 3456 ext. 66013

Email: ylchien71@ntu.edu.tw

Using Wearable Device and Artificial Intelligence to Predict Mood Symptoms in Bipolar Disorder

Abstract

Bipolar disorder (BD) is a highly recurrent disorder. Early detection, early intervention, and prevention of recurrent bipolar mood symptoms are key for better prognosis. In this study, we build prediction models for bipolar disorder with machine learning algorithms. This study recruited 24 participants with BD. The Beck Depression Inventory (BDI) and Young Mania Rating Scale (YMRS) were used to evaluate depressive and manic episodes respectively. Using digital biomarkers collected from wearable devices as input, six machine learning algorithms (Logistic Regression, Decision Tree, K-Nearest Neighbors, Random Forest, Adaptive Boosting, and Extreme Gradient Boosting) were used to build predictive models. The prediction model for depressive symptoms achieved 83% accuracy, 0.89 Area Under the Receiver Operating Characteristic curve (AUROC), and 0.65 F1 score on testing data. The prediction model for manic symptoms achieved 91% accuracy, 0.88 AUROC, and 0.25 F1 score on testing data. With the interpretable model Shapely Additive exPlanations (SHAP), we found that relatively high resting heart rate, low activity, and lack of sleep may predict depressive symptoms. This study demonstrated that digital biomarkers could be used to predict depressive and manic symptoms. Moreover, based on the findings from the prediction model, we may provide clinical assessment and treatment earlier to prevent a recurrence.

Keywords: bipolar disorder; wearable device; machine learning; prediction, explainable model

Introduction

Bipolar disorder (BD) is a recurrent disorder characterized by fluctuations in mood and energy from depression to mania, that often results in enormous functional impairment, and high disease burden [21]. The 5-year recurrence rate of BD was as high as 73% [18]. A meta-analysis estimated the recurrence rates under treatment were 55.2% (naturalistic studies) and 39.3% (randomized controlled trials) versus 60.6% under placebo [23]. Although recurrence rate can be reduced under appropriate

treatment [8], irregular compliance usually precipitates recurrence of mood episodes and compromises the outcome. Despite high recurrence and low compliance, the frequency of routine follow-up in current practice is usually insufficient to detect early signs of relapse, causing delayed treatment until an acute episode has fully developed [22]. It is important to detect early signs of relapse so that an upcoming mood episode can be aborted by appropriate intervention. Furthermore, clinical practice mostly relies on patients' self-report of symptom changes, which depends on illness insight [6] and is subject to recall bias [7].

Recent studies have developed machine learning algorithms to predict depression [15] based on socio-demographic features, personal or family health history, and symptoms coded by psychiatrists [19]. In BD, lower heart rate variability was associated with manic and depressive states [13]. Activity level may be a promising digital phenotype for mood episode, with decreased activity for depressive and increased activity for manic episode [12]. Also, sleep problems may indicate relapse in BD [5]. However, most current studies targeted full-blown mood episodes and followed patients for shorter time periods. Besides, information about feature importance was lacking.

This study established machine learning algorithms to predict early signs of upcoming depressive or manic symptoms based on digital biomarkers collected by actigraphy and mobile devices, including activity, sleep hours, and heart rate. We hypothesized that greater activity level and shorter sleep hours may predict manic symptoms. We also hypothesized that high resting heart rate and lower activity level and sleep hour changes may predict depressive symptoms. Moreover, considering that the change of a feature (e.g., sleep or activity) from the individual's baseline may be important when predicting mood symptoms, we also integrated individualized parameters for model prediction.

Methods

Participants

This study recruited 24 BD patients (aged 20-65, mean 38 ± 9 ; male $n=9$) from the Psychiatric Department of National Taiwan University Hospital from October 2020 to July 2022, and prospectively followed for 6.39 ± 4.85 months. The diagnosis was made by board-certified senior psychiatrists based on the Diagnostic and Statistical Manual of Mental Disorders 5th edition [1]. Participants' educational levels ranged from middle school to graduate school.

This study was approved by the Research Ethics Committee at National Taiwan University Hospital (202002006RINA) before its implementation. The investigation was carried out in accordance with the latest version of the Declaration of Helsinki. Informed consent of the participants was obtained after the nature of the procedures had been fully explained.

Clinical Measures

The Beck Depression Inventory (BDI) and Young Mania Rating Scale (YMRS) were self-rated weekly on a mobile app on the personal smartphone to evaluate depressive and manic symptoms, respectively.

The BDI is a 21-item self-reported inventory measuring the severity of depression in adolescents and adults [3]. The BDI-II was revised in 1996 [2] to be more consistent with the DSM-IV criteria of depression. The inventory consists of 21 items, in which four response options are presented on a Likert scale from 0 to 3. The BDI was translated into Chinese version with good internal consistency (Cronbach alpha = 0.85) and concurrent validity [26]. Higher total scores indicate more severe depressive symptoms. As for the standard cut-off scores, 30–63 may indicate severe depression, 19–29 indicates moderate depression, 10–18 indicates mild depression, 0–9 indicates minimal

depression [26]. The distribution of BDI scores in our sample were summarized in Supplementary Figure S1.

The Young Mania Rating Scale (YMRS), an 11-item interviewer rated scale, is designed for assessing the severity of manic symptoms [24]. The items are rated on five grades of severity. Four items among them are double weighted, including irritability, speech, thought content and disruptive/aggressive behavior. The YMRS is by far the most commonly used standardized measure of bipolar manic symptoms for clinical trials in acute mania. As for psychometric properties, the inter-rater reliability was adequate for total score (0.93) and for individual items ranged from 0.67 to 0.95. There are no firmly established scoring criteria that relate to diagnostic classification [24]. The distribution of YMRS scores in our sample were summarized in Supplementary Figure S2.

To detect early signs of relapse, a state of mild depression ($BDI > 13$) or mania ($YMRS > 13$) was labeled true [10].

Digital Biomarkers

The wristwatch-like actimetry sensor was worn by each participant, and continually measuring as well as recording the motor activity, sleep length, and heart rate (HR).

The motor activities which were monitored included steps, distance traveled, floors climbed. The sleep length was determined with stages of sleep: (1) deep, (2) light, (3) rapid eye movement, and (4) awake stages. The heart rate was divided into different status: (1) the minimum HR values, (2) the maximum HR values, (3) the average HR during the past 7 days, and (4) the average HR at rest.

Machine Learning Algorithms

Fig. 1A shows workflow architecture including interpretable model Shapely Additive exPlanations (SHAP). To achieve more balanced data, under-sampling technique was applied for the training set. For individualized features, we transformed the raw feature into 'difference from the personal mean' by subtracting the raw value by the mean of that feature of that person (Supplementary Table S1).

Depressive and manic models were trained separately. Python, scikit-learn, and SHAP packages were utilized for programming, model training, and model explanation respectively. We adopted logistic regression, Decision Tree, K-Nearest Neighbors, Random Forest, Adaptive Boosting, as well as Extreme Gradient Boosting (XGBoost) to predict mood symptoms. Five-fold cross-validation was used to evaluate model performance and hyperparameter selection before making predictions on testing data [16] (Fig. 1C). The grid search strategy for hyperparameter selection during training process was used to ensure model stability, which loops through all candidate parameters leaving only the final best-performing set of parameters.

Model Assessment

Model performance was assessed by accuracy, sensitivity, specificity, precision, F1 score, and Area Under the Receiver Operating Characteristic curve (AUROC). To explain the output of machine learning models [11]. The SHAP method analyzed the prediction from machine learning models and interpreted the contributions of each feature by a Shapley value, which measures the influence of a feature on the prediction.

Results

The features were compared between the depressive and non-depressive labels, and between the manic and non-manic labels by Student's t-test (Supplementary Table S1&S2). We found that the

depressive label had a significantly higher minimal/resting heart rate and lower activity level than the non-depressive label. Besides, the manic label had a significantly lower sleep duration than the non-manic label. Similarly, for individualized data, we found that the manic label had a significantly higher heart rate and activity level and lower sleep duration than the non-manic label.

Prediction Model

Firstly, we used 12 features (without individualized data) for model training. Table 1 shows model performance on the testing set. Compared to other algorithms, the XGBoost performed with the highest accuracy (0.79), AUROC (0.85) and F1 score (0.57) in depressive model. Similarly in manic model, the XGBoost performed superior than others, with the highest accuracy (0.83), AUROC (0.84) and F1 score (0.19).

Table 1. The performance of the models for depressive or manic symptoms without individualized features in the models

	Accuracy	AUROC	Sensitivity	Specificity	Precision	F1 score
Models for a depressive episode						
Logistic Regression	0.63	0.77	0.88	0.57	0.34	0.49
Decision Tree	0.74	0.65	0.34	0.84	0.35	0.35
KNN	0.69	0.58	0.38	0.77	0.30	0.34
Random Forest	0.70	0.83	0.87	0.66	0.39	0.54
AdaBoost	0.78	0.79	0.40	0.87	0.44	0.42
XGBoost	0.79	0.85	0.71	0.81	0.48	0.57
Models for a manic episode						
Logistic Regression	0.65	0.63	0.58	0.65	0.05	0.09
Decision Tree	0.87	0.64	0.25	0.89	0.07	0.10
KNN	0.90	0.51	0.08	0.93	0.03	0.05
Random Forest	0.75	0.78	0.58	0.75	0.07	0.12
AdaBoost	0.94	0.74	0.08	0.97	0.08	0.08
XGBoost	0.83	0.84	0.67	0.84	0.11	0.19

KNN: K-Nearest Neighbors

AdaBoost: Adaptive Boosting

XGBoost: Extreme Gradient Boosting

Next, we examined whether adding individualized features in the models improved prediction. The prediction performance of the 19-feature model with individualized features was successfully improved in both depressive and manic models (Table 2). The accuracy, AUROC, and F1 score in depressive model by the XGBoost increased from 0.79, 0.85, and 0.57 to 0.83, 0.89, and 0.65, while those in manic model increased from 0.83, 0.84, and 0.19 to 0.91, 0.88, and 0.25.

Table 2. The performance of the models for depressive or manic symptoms with the individualized features included in the models

	Accuracy	AUROC	Sensitivity	Specificity	Precision	F1 score
Models for a depressive episode						
Logistic Regression	0.86	0.87	0.62	0.92	0.68	0.65
Decision Tree	0.79	0.75	0.50	0.87	0.49	0.50
KNN	0.79	0.70	0.54	0.85	0.48	0.51
Random Forest	0.82	0.86	0.79	0.83	0.55	0.65
AdaBoost	0.82	0.83	0.54	0.89	0.56	0.55
XGBoost	0.83	0.89	0.78	0.85	0.56	0.65
Models for a manic episode						
Logistic Regression	0.71	0.84	0.92	0.71	0.09	0.16
Decision Tree	0.90	0.69	0.42	0.92	0.14	0.20
KNN	0.89	0.62	0.33	0.91	0.10	0.15
Random Forest	0.89	0.81	0.25	0.91	0.08	0.12
AdaBoost	0.93	0.78	0.33	0.95	0.16	0.22
XGBoost	0.91	0.88	0.50	0.92	0.17	0.25

KNN: K-Nearest Neighbors

AdaBoost: Adaptive Boosting

XGBoost: Extreme Gradient Boosting

Explanation of Prediction Model

In depressive model, resting heart rate revealed the highest feature importance, followed by deep sleep duration, floors climbed, average heart rate, and steps (Fig. 2A). The force plots explained how resting heart rate (Fig. 2B), number of steps (Fig. 2C) and total sleep duration (Fig. 2D) affect the model. Specifically, steps less than 6,000 steps/day and sleep less than 6h may predict depressive symptoms. Together, a high resting heart rate produces a push to the right, while sufficient activity and sleep produce a push to the left (Fig. 2E); all forces result in a 0.23 predicted probability of a depressive episode.

Discussion

This study constructed prediction models for mild depressive and manic symptoms using digital biomarkers, including activity, sleep state, heart rate, and the individualized parameters of these features. Major findings included, firstly, the accuracy of the prediction achieved 79% (AUROC 0.85) in depressive model, 83% (AUROC 0.84) in manic model. Secondly, adding individualized features improved accuracy to 83% (AUROC 0.89) in depressive model and 91% (AUROC 0.88) in manic model. Thirdly, among the features, higher resting heart rate, lower activity, and shorter sleep

were important in predicting depressive symptoms. These findings supported the utility of digital biomarkers collected from wearable devices in predicting depressive and manic symptoms in the early phase.

The accuracy and AUROC in our depressive model were similar to an earlier study (0.87) [4], with a higher sensitivity in ours (0.78 vs. 0.48). In manic model, their accuracy (0.91) and AUROC (0.91) were close to ours but their sensitivity (0.31) was lower than ours (~0.92). Such discrepancy may be related to the machine learning algorithm (Random Forest in theirs vs. XGBoost in ours) and individualized features included in our model. Additionally, they provided model performance and feature importance, but we further explain the model by utilizing insights from SHAP to explain the clinical implications of the features.

Our findings supported the prediction of depressive symptoms by digital features including activity level, sleep parameters, and heart rates during the seven days before depressive symptoms were reported. Among the features, we found that a higher resting heart rate may contribute to the prediction of depression. Two previous studies also showed that the heart rate was higher in depressive patients [20]. Severe depression is often accompanied by increased heart rate and reduced heart rate variability [14,17]. Our data suggested a turning point of resting heart rate at 60 (Fig. 2E), showing that resting heart rate higher than 60 was linked to a higher chance of predicting depressive symptoms. Regarding sleep parameters, a total sleep duration of fewer than 6 hours may contribute to the prediction of depressive symptoms, consistent with previous evidence that supported sleep duration remained an important correlate for depressive symptoms [9, 25].

Limitations

Several limitations need to be addressed. First, the labeling method relies on self-rated questionnaires completed at least once a week which backfills for seven days given that the questionnaire rated mood symptoms in the past seven days. While participants were instructed to report mood symptoms as soon as they became aware of them, there is possibility that some individuals did not complete the questionnaires promptly when their symptoms first emerged. Timely symptom labeling, immediately upon awareness of depressive or manic symptoms, could potentially enhance data reliability.

Second, participants in the study were undergoing treatment with mood stabilizers and/or antipsychotics that may reduce symptom relapse. However, it is essential to acknowledge that medical treatment is inevitable in long-term follow-up in a clinical sample. Future studies may consider recording the days on medication to assess its potential impact on sleep and heart rates.

Third, low F1 score in the prediction model for manic symptoms may reflect insufficient events for prediction, which also influences the analysis of feature importance. Nevertheless, this study explored the potential of predicting early signs of mood episodes using digital biomarkers. Towards that end, we employed the SHAP method to interpret the model and included individualized features that better capture the changes in patients' lifestyles and successfully improve the prediction models for both depressive and manic symptoms. Our findings highlighted the potential of applying wearable devices to detect early signs of relapse in BD for early intervention.

Conclusions

This study utilized digital biomarkers obtained from wearable devices to construct machine learning models for prediction of depressive and manic symptoms. By incorporating individualized features into the models, we achieved satisfactory accuracy in predicting depressive symptoms achieves satisfactory accuracy. Furthermore, the application of interpretable SHAP model allowed us to discern that higher resting heart rate, lower activity, and insufficient sleep were indicative of

impending depressive symptoms. Early detection of depressive changes enables the timely introduction of adequate psychoeducation and clinical assessment, facilitating the implementation of intervention can be introduced in time to mitigate upcoming mood episode, thereby reducing the risk of recurrence.

Acknowledgements

This work was funded by a grant from the Ministry of Economic Affairs (113A-10) of Taiwan.

Authors' Contributions

Ding-Shan Liu, Chia-Tung Wu were responsible for data analysis and paper processing; I-Ming Chen, Ming-Hsien Hsieh, Chen-Chun Lu were responsible for the enrollment, contact, and participant follow-ups; Yi-Ling Chien, Feipei Lai directed the project.

Conflicts of Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (5th ed.). American Psychiatric Publishing; 2013 ISBN:0890425574
2. Beck AT, Steer RA, Ball R, Ranieri W. Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. *J Pers Assess* 1996 Dec;67(3):588-97 PMID: 8991972
3. Beck AT, Ward CH, Mendelson M, Mock, J, Erbaugh J. An inventory for measuring depression. *Arch Gen Psychiatry* 1961 Jun;4: 561-71 PMID: 13688369
4. Cho CH, Lee T, Kim MG, In HP, Kim L, Lee HJ. Mood prediction of patients with mood disorders by machine learning using passive digital phenotypes based on the circadian rhythm: prospective observational cohort study. *J Med Internet Res* 2019 Apr 17;21(4):e11029 PMID: 30994461
5. Cretu JB, Culver JL, Goffin KC, Shah S, Ketter TA. Sleep, residual mood symptoms, and time to relapse in recovered patients with bipolar disorder. *J Affect Disord*. 2016 Jan 15;190:162-66 PMID: 26519636
6. Faurholt-Jepsen M, Vinberg M, Frost M, Debel S, Margrethe Christensen E, Bardram JE, Kessing LV. Behavioral activities collected through smartphones and the association with illness activity in bipolar disorder. *Int J Methods Psychiatr Res*. 2016 Dec;25(4):309-23 PMID: 27038019
7. Harrison PJ, Cipriani A, Harmer CJ, Nobre AC, Saunders K, Goodwin GM, Geddes JR. Innovative approaches to bipolar disorder and its treatment. *Ann N Y Acad Sci*. 2016 Feb;1366(1):76-89 PMID: 27111134
8. Kishi T, Matsuda Y, Sakuma K, Okuya M, Mishima K, Iwata N. Recurrence rates in stable bipolar disorder patients after drug discontinuation v. drug maintenance: a systematic review and meta-analysis. *Psychol Med*. 2021 Nov;51(15):2721-29 PMID: 33046156

9. Li W, Yin J, Cai X, Cheng X, Wang Y. Association between sleep duration and quality and depressive symptoms among university students: A cross-sectional study. *PLoS One*. 2020 Sep 11;15(9):e0238811 PMID: 32915844
10. Liu DS. Using Wearable Device and AI to Predict Mood Symptoms in Bipolar Disorder, Graduate Institute of Computer Science & Information Engineering, College of Electrical Engineering and Computer Science. National Taiwan University, Taipei, TW. 2022
11. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS 2017)* p. 4768 – 77 doi/10.5555/3295222.3295230
12. Maatoug R, Oudin A, Adrien V, Saudreau B, Bonnot O, Millet B, Ferreri F, Mouchabac S, Bourla A. Digital phenotype of mood disorders: A conceptual and critical review. *Front Psychiatry* 2022 Jul 26;13:895860 PMID: 35958638
13. Nardelli M, Lanata A, Bertschy G, Scilingo EP, Valenza G. Heartbeat complexity modulation in bipolar disorder during daytime and nighttime. *Sci Rep*. 2017 Dec 20;7(1):17920 PMID: 29263393
14. Pawlowski MA, Gazea M, Wollweber B, Dresler M, Holsboer F, Keck ME, Steiger A, Adamczyk M, Mikoteit T. Heart rate variability and cordance in rapid eye movement sleep as biomarkers of depression and treatment response. *J Psychiatr Res*. 2017 Sep;92:64-73 PMID: 28411417
15. Priya A, Garg S, Tigga NP. Predicting anxiety, depression and stress in modern life using machine learning algorithms. *Procedia Computer Science* 2020;167:1258-67 doi/10.1016/j.procs.2020.03.442
16. Raschka S. Model evaluation, model selection, and algorithm selection in machine learning. *arXiv 2018 arXiv:1811.12808*
17. Saad M, Ray LB, Bujaki B, Parvaresh A, Palamarchuk I, De Koninck J, Douglass A, Lee EK, Soucy LJ, Fogel S, Morin CM, Bastien C, Merali Z, Robillard R. Using heart rate profiles during sleep as a biomarker of depression. *BMC Psychiatry*. 2019 Jun 7;19(1):168 PMID: 31174510
18. Sajatovic M. Bipolar disorder: disease burden. *Am J Manag Care* 2005 Jun;11(3 Suppl):S80-4 PMID: 16097718
19. Sau A, Bhakta, I. Predicting anxiety and depression in elderly patients using machine learning technology. *Healthc Technol Lett* 2017 Nov;4(6):238-43 doi/10.1049/htl.2016.0096
20. Schiweck C, Lutin E, De Raedt W, Cools O, Coppens V, Morrens M, Van Hoof C, Vrieze E, Claes S. Twenty-Four-Hour Heart Rate Is a Trait but Not State Marker for Depression in a Pilot Randomized Controlled Trial With a Single Infusion of Ketamine. *Front Psychiatry* 2021 Jul;12:696170 PMID: 34393856
21. Simon GE. Social and economic burden of mood disorders. *Biol Psychiatry* 2003 Aug 1;54(3):208-15 PMID: 12893097

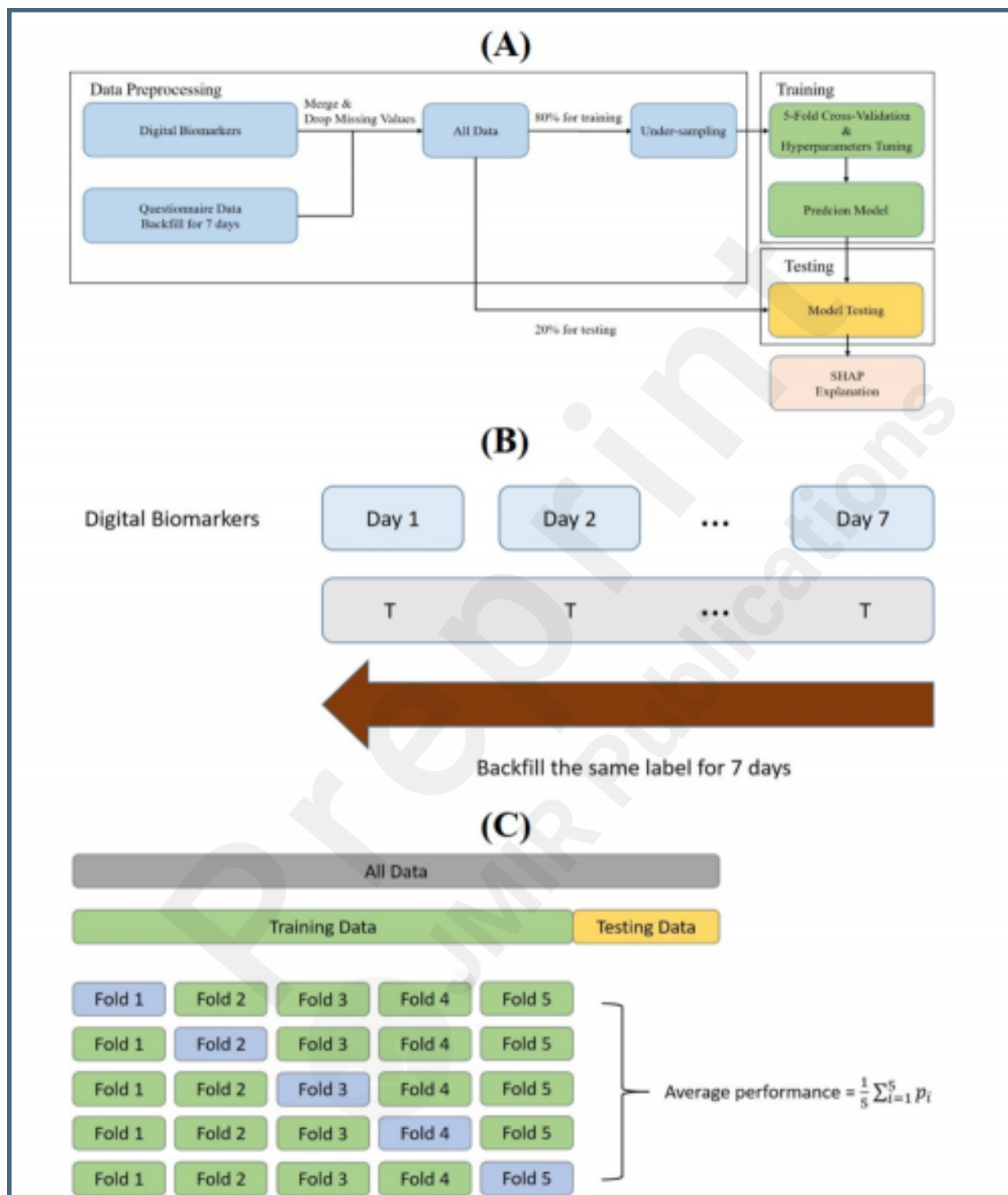
22. Vieta E, Salagre E, Grande I, Carvalho AF, Fernandes BS, Berk M, Birmaher B, Tohen M, Suppes T. Early intervention in bipolar disorder. *Am J Psychiatry* 2018 May 1;175(5):411-26 PMID: 29361850
23. Vázquez GH, Holtzman JN, Lolic M, Ketter TA, Baldessarini RJ. Recurrence rates in bipolar disorder: Systematic comparison of long-term prospective, naturalistic studies versus randomized controlled trials. *Eur Neuropsychopharmacol* 2015 Oct;25(10):1501-12 PMID: 26238969
24. Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. *Br J Psychiatry* 1978 Nov;133:429-35 PMID: 728692
25. Zhai L, Zhang H, Zhang D. Sleep duration and depression among adults: A meta-analysis of prospective studies. *Depress Anxiety* 2015 Sep;32(9):664-70 PMID: 26047492
26. Zheng YP, Wei LA, Goa LG, Zhang GC, Wong CG. Applicability of the Chinese Beck Depression Inventory. *Compr Psychiatry*. 1988 Sep-Oct;29(5):484-9 PMID: 3180758

Abbreviations

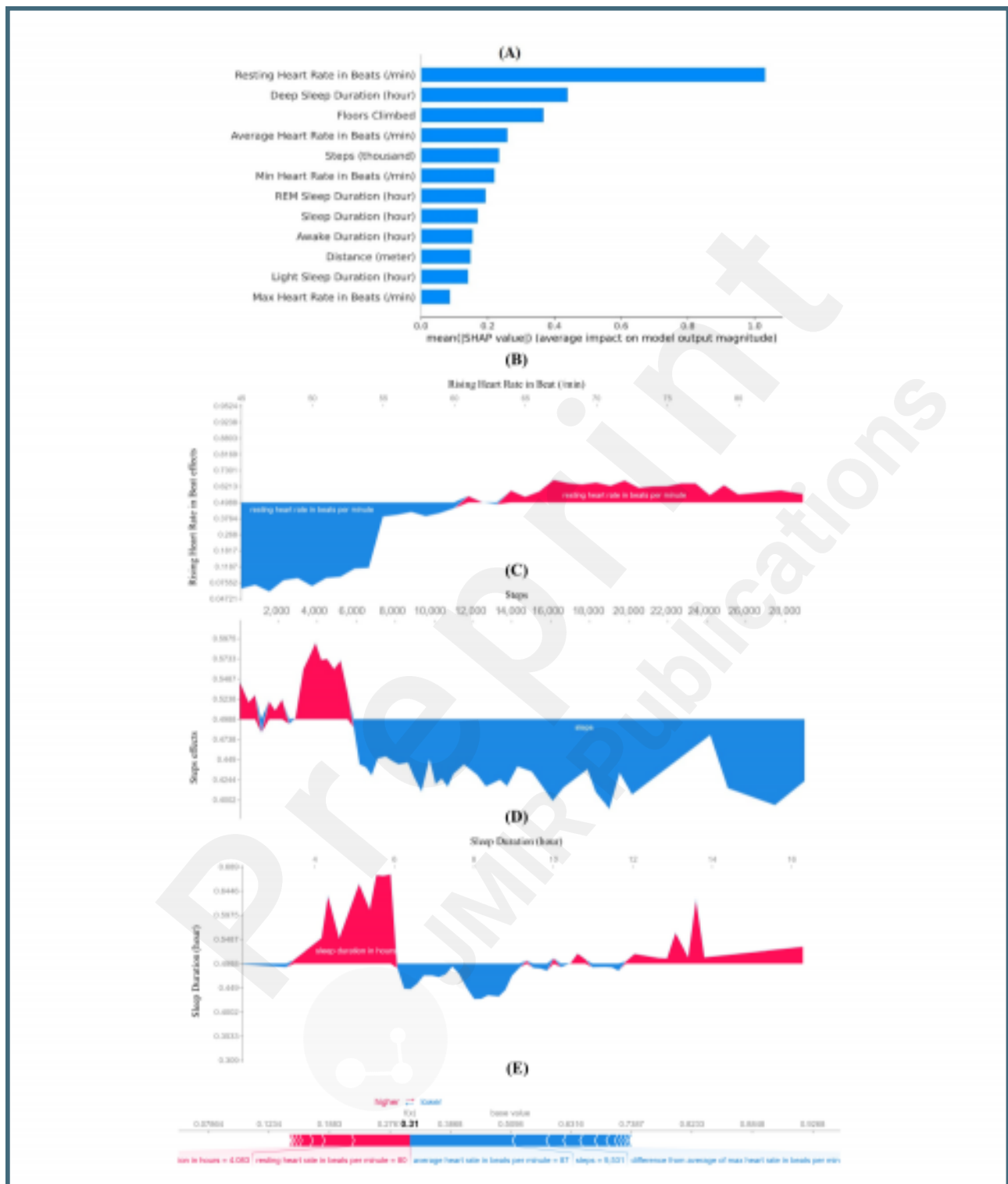
Supplementary Files

Figures

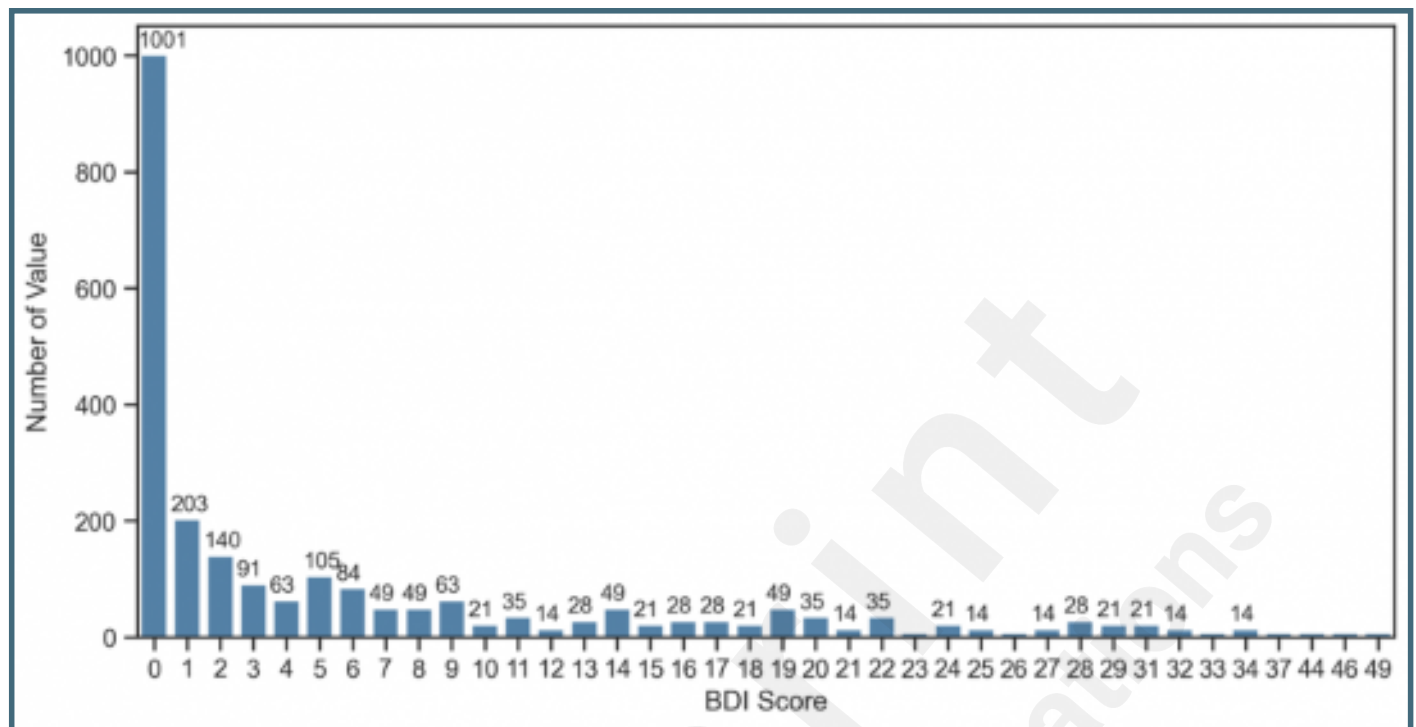
Machine learning algorithms.



Explanation of prediction model.



Distribution of BDI scores.



Distribution of YMRS scores.

