

Supplementary Information for

Accurately Predicting Mood Episodes in Mood Disorder Patients Using Wearable Sleep and Circadian Rhythm Features

Dongju Lim^{1,2,†}, Jaegwon Jeong^{3,4,†}, Yun Min Song^{1,2}, Chul-Hyun Cho^{3,4}, Ji Won Yeom^{3,4}, Taek Lee⁵, Jung-Been Lee⁵, Heon-Jeong Lee^{3,4*}, Jae Kyoung Kim^{1,2,6*}

¹ Department of Mathematical Sciences, KAIST, Daejeon 34141, Republic of Korea

² Biomedical Mathematics Group, Institute for Basic Science, Daejeon 34126, Republic of Korea

³ Department of Psychiatry, Korea University College of Medicine, Seoul 02841, Republic of Korea

⁴ Chronobiology Institute, Korea University, Seoul 02841, Republic of Korea

⁵ Division of Computer Science and Engineering, Sun Moon University, Asan 31460, Republic of Korea

⁶ Department of Medicine, College of Medicine, Korea University, Seoul 02841, Republic of Korea

[†]D.L. and J.J. contributed equally

* Correspondence:

Heon-Jeong Lee, MD, PhD, Department of Psychiatry, Anam Hospital, Korea University College of Medicine, 73 Goryeodae-ro, Seongbuk-gu, Seoul 02841, Korea.

Tel: 82-2-920-5815, E-mail: leehjeong@korea.ac.kr

Jae Kyoung Kim, PhD, Department of Mathematical Sciences, KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon, 34141, Korea.

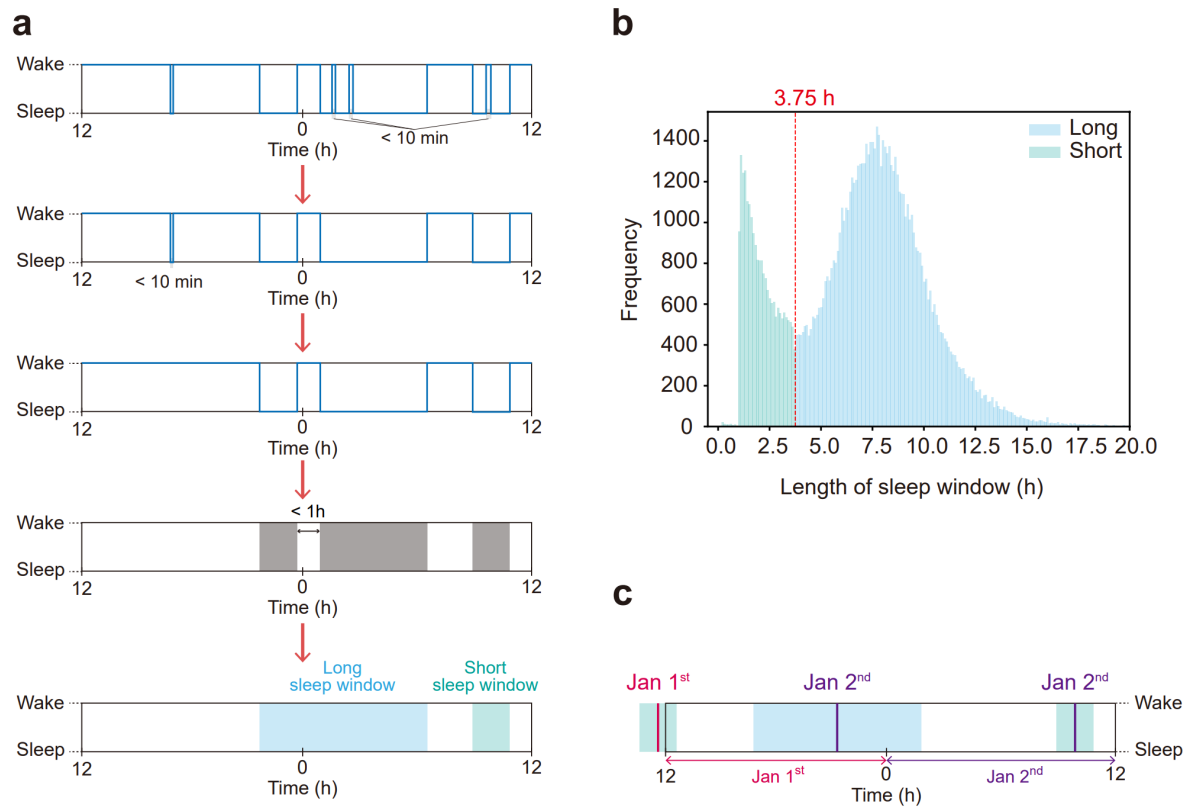
Tel: 82-42-350-2702~4 Email: jaekkim@kaist.ac.kr

Table of contents:

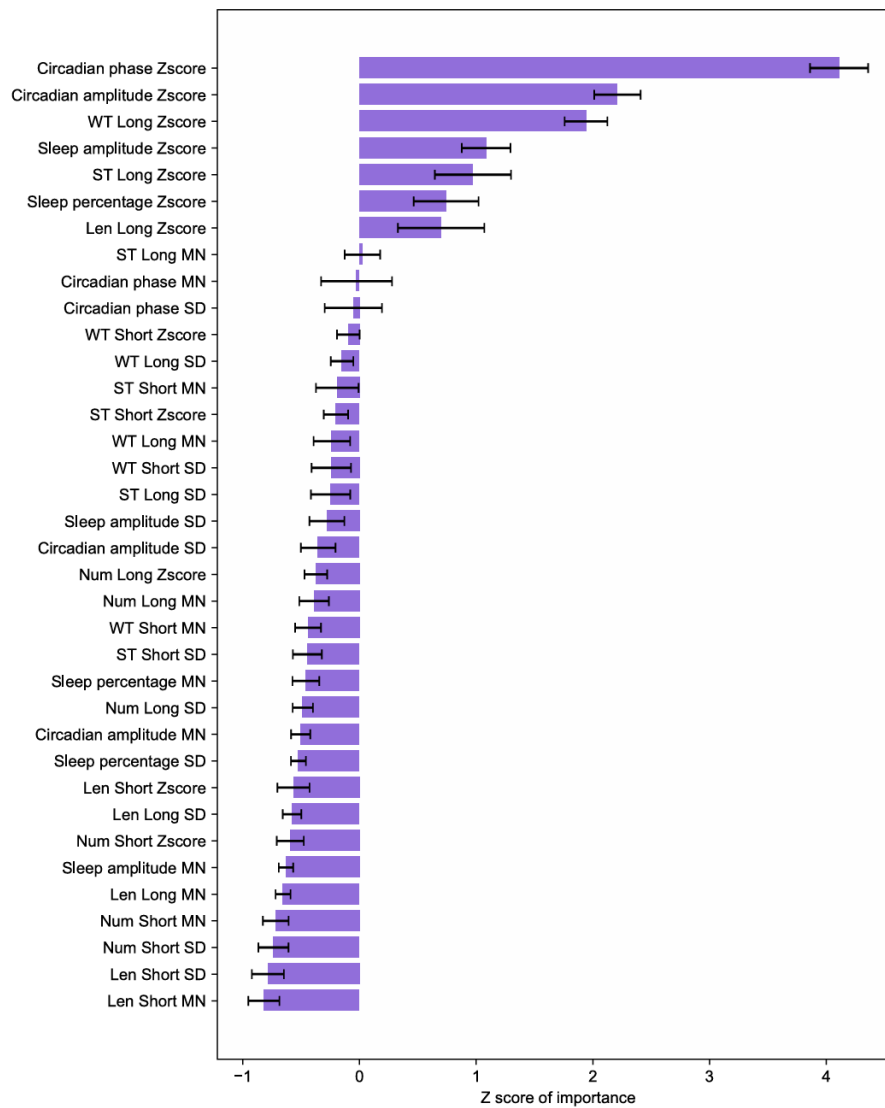
Supplementary Figure 1-6

Supplementary Table 1-7

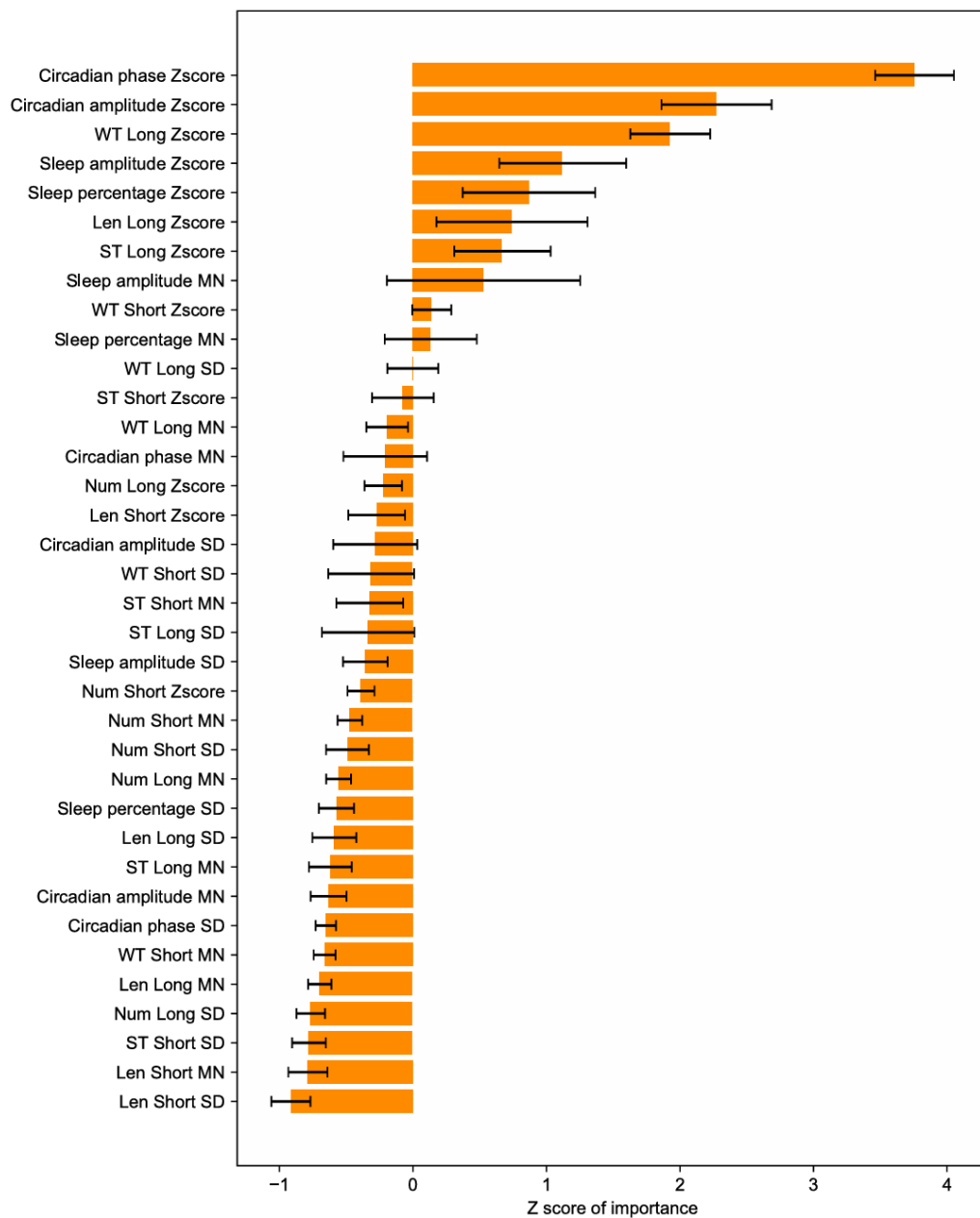
STROBE check list



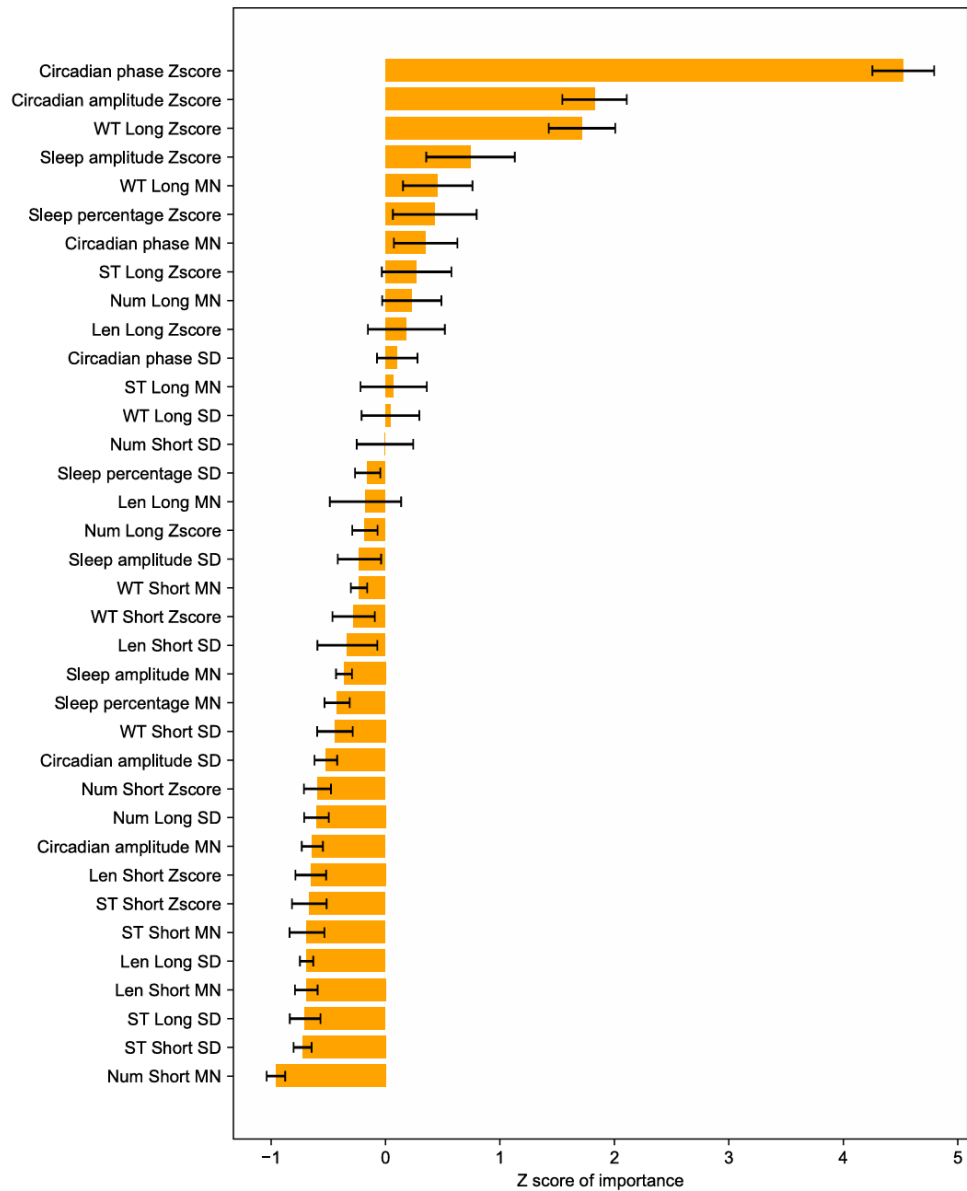
Supplementary Figure 1. Schematic figure for constructing the daily sleep windows. (a) Schematic figure of sleep window construction from the sleep/wake time series data. All wake episodes less than 10 minutes were considered as sleep, and sleep episodes less than 10 minutes were considered as wake to identify comprehensive sleep windows. These windows were further aggregated if they were within a 1-hour interval, to form a more comprehensive window. Final sleep windows were classified as “Long” or “Short” with a threshold length of 3.75 hours. (b) This threshold was determined by the bimodal distribution of the sleep window length. (c) Long and Short sleep windows were assigned to a specific day based on their midpoints (red and purple vertical lines). We found the nearest midnight from the midpoint and utilized its date as the assigned date for the sleep window.



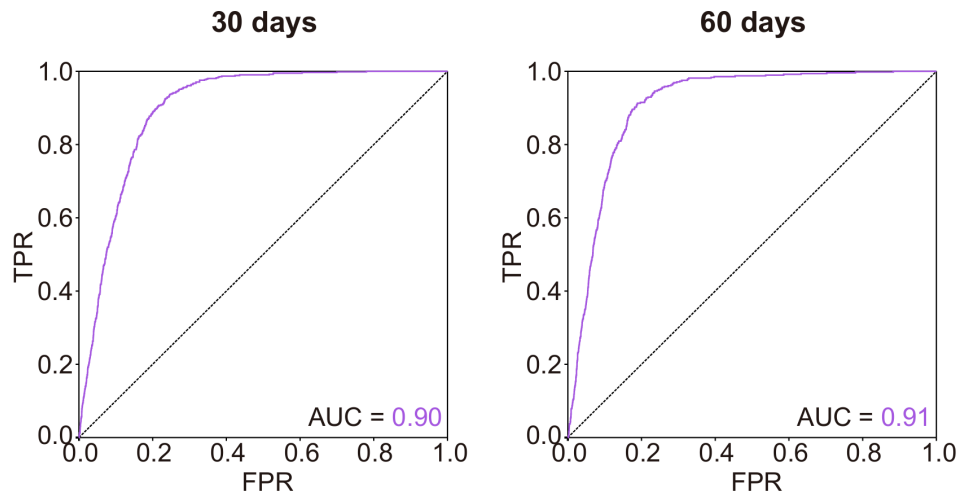
Supplementary Figure 2. Summary plot of the feature importance of 10 different depressive episode prediction models trained with a randomly sampled 80% of the entire dataset. The Z-score of the feature importance was calculated across 10 different depressive episode prediction models, each trained on a randomly sampled 80% of the entire dataset. The purple bar represents the mean of these 10 Z-scores, while the error bar reflects the standard deviation, with its half-length corresponding to the standard deviation among the 10 values. The circadian phase is the most important feature.



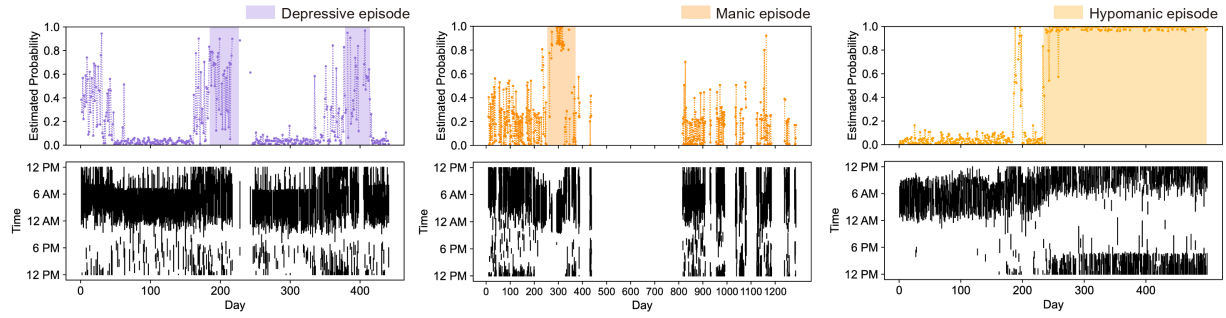
Supplementary Figure 3. Summary plot of the feature importance of 10 different manic episode prediction models trained with a randomly sampled 80% of the entire dataset. The Z-score of the feature importance was calculated across 10 different manic episode prediction models, each trained on a randomly sampled 80% of the entire dataset. The dark orange-colored bar represents the mean of these 10 Z-scores, while the error bar reflects the standard deviation, with its half-length corresponding to the standard deviation among the 10 values. The circadian phase is the most important feature.



Supplementary Figure 4. Summary plot of the feature importance of 10 different hypomanic episode prediction models trained with a randomly sampled 80% of the entire dataset. The Z-score of the feature importance was calculated across 10 different hypomanic episode prediction models, each trained on a randomly sampled 80% of the entire dataset. The orange bar represents the mean of these 10 Z-scores, while the error bar reflects the standard deviation, with its half-length corresponding to the standard deviation among the 10 values. The circadian phase is the most important feature.



Supplementary Figure 5. The trained model can accurately predict depressive episodes without the medication-induced changes in the circadian phase. The depressive episode prediction model trained with 30-day ranges (left) and 60-day range (right) were tested with testing sets consisting of the data from the patients who did not change their medication types and dosage levels after the episode onset (the purple curve). The high AUC value suggests that the accuracy of our prediction models is not due to medication-induced changes in circadian phase.



Supplementary Figure 6. Example for the predicted probability of mood episodes. The predicted probability (dotted curves) of depressive (left), manic (center), and hypomanic (right) episodes for three different patients MDCRC06043, MDCRC01012, MDCRC01070, respectively, and corresponding sleep patterns (black bar indicates the sleep period). Predicted values are well-matched with true episode periods (transparent boxes). The predicted manic episode probabilities of the MDCRC01012 patient are missing from Day 437 (Aug 10th, 2016) to Day 815 (Aug 23th, 2017), as corresponding sleep-wake pattern data were missing. The predicted depressive episode probabilities of MDCRC06043 patient are missing from Day 218 (Aug 21st, 2018) to Day 245 (Sep 17th, 2018) except Day 228 (Aug 31st, 2018) and Day 243 (Sep 15th, 2018), as corresponding sleep-wake pattern data were missing.

Supplementary Table 1. Comparison of clinical characteristics between female and male

	Female	Male	<i>p</i> -value
N (%)	93 (55%)	75 (45%)	
Age at baseline (mean, SD)	23.5 (3.82)	23.7 (3.64)	0.70
Age at onset (mean, SD)	17.4 (4.95)	18.4 (5.09)	0.18
Diagnosis (n, %)			0.50
MDD	28 (30%)	29 (39%)	
BD 1	25 (27%)	17 (22%)	
BD 2	40 (43%)	29 (39%)	
MADRS at baseline (mean, SD)	15.7 (10.9)	14.5 (11.2)	0.47
YMRS at baseline (mean, SD)	2.81 (3.49)	1.87 (2.80)	0.055
CSM at baseline (mean, SD)	30.1 (7.97)	28.6 (6.83)	0.19
GSS at baseline (mean, SD)	6.13 (4.49)	5.53 (5.23)	0.44
Number of patients with recurrence	61 (66%)	47 (63%)	0.82
Clinical fu duration	586 (364)	587 (388)	0.99
Comorbidities			
Anxiety disorder	21 (23)	25 (33)	0.17
OC related disorder	7 (7.5)	10 (13)	0.33
Substance use disorder	5 (5.4)	3 (4.0)	0.73
PTSD	4 (4.3)	3 (5.0)	1
Eating disorder	8 (8.6)	1 (1.3)	0.044*
Somatic symptom and related disorder	2 (2.1)	2 (2.6)	1

Data are expressed as mean (standard deviation). The *p*-values are results of two-sample t-test for two groups comparison, and chi-squared test and fisher's exact test for sex, diagnosis, and comorbidities. MADRS Montgomery–Åsberg Depression Rating Scale, YMRS Young Mania Rating Scale, CSM Composite Scale of Morningness, GSS Global Seasonality Score, OC obsessive-compulsive, PTSD posttraumatic stress disorder. There were no significant differences in most demographic and clinical characteristics observed between sex; however, the prevalence of eating disorders was higher among females. * *p*-value <0.05

Supplementary Table 2. Annualized relapse rates (100 person-year) by season for each mood episodes with 95% confidence intervals.

	Annualized relapse rate (95% CI, 100 person-year)			
	Spring	Summer	Fall	Winter
Depressive	0.59 (0.42-0.81)	0.84 (0.63-1.10)	0.57 (0.40-0.78)	0.57 (0.41-0.77)
Manic	0.34 (0.14-0.70)	0.25 (0.081-0.58)	0.20 (0.053-0.50)	0.23 (0.074-0.54)
Hypomanic	0.33 (0.18-0.54)	0.11 (0.037-0.26)	0.23 (0.11-0.42)	0.18 (0.082-0.34)

Annualized relapse rates (ARR) by season with 95% confidence interval for depressive (1st row), manic (2nd row), and hypomanic (3rd row) episodes across the seasons. The ARR for depressive episodes is calculated across all patients. The ARR for manic episodes is calculated specifically for patients with bipolar disorder type 1. The ARR for hypomanic episodes is calculated for patients with both bipolar disorder type 1 and 2.

Supplementary Table 3. Descriptive statistics of sleep and circadian indexes

Sleep and Circadian Indexes	Episode type			
	Inter	Depressive	Manic	Hypomanic
ST Long MN	7.04 (1)	6.97 (1.33)	7.1 (0.71)	6.73 (0.8)
ST Long SD	2.56 (0.73)	2.8 (0.63)	2.91 (0.69)	2.89 (0.72)
ST Long Zscore	0.04 (0.98)	0.07 (1.02)	-0.01 (1.19)	-0.11 (0.97)
ST Short MN	0.52 (0.33)	0.54 (0.37)	0.8 (0.53)	0.66 (0.53)
ST Short SD	1.01 (0.28)	1 (0.31)	1.2 (0.38)	1.08 (0.41)
ST Short Zscore	-0.01 (0.98)	0.04 (1.03)	0.04 (1.1)	0.04 (1.01)
WT Long MN	0.82 (0.3)	1.05 (0.58)	0.76 (0.18)	0.74 (0.19)
WT Long SD	0.48 (0.21)	0.69 (0.54)	0.5 (0.11)	0.46 (0.13)
WT Long Zscore	-0.08 (0.98)	0.03 (1.04)	-0.23 (1.12)	-0.08 (0.98)
WT Short MN	0.05 (0.03)	0.07 (0.06)	0.07 (0.05)	0.07 (0.06)
WT Short SD	0.13 (0.07)	0.17 (0.13)	0.14 (0.06)	0.14 (0.07)
WT Short Zscore	-0.03 (0.93)	0.03 (1.05)	0.02 (0.95)	0.01 (0.97)
Len Long MN	7.88 (1.12)	8.05 (1.31)	7.9 (0.73)	7.48 (0.86)
Len Long SD	2.85 (0.81)	3.18 (0.73)	3.22 (0.81)	3.21 (0.85)
Len Long Zscore	0.02 (0.97)	0.07 (1.03)	-0.05 (1.17)	-0.12 (0.98)
Num Long MN	0.97 (0.05)	0.97 (0.05)	0.98 (0.05)	0.95 (0.04)
Num Long SD	0.28 (0.1)	0.3 (0.09)	0.34 (0.14)	0.33 (0.15)
Num Long Zscore	0.02 (0.96)	0.02 (1)	-0.17 (1.26)	-0.02 (0.96)
Len Short MN	0.58 (0.35)	0.62 (0.4)	0.89 (0.59)	0.73 (0.59)
Len Short SD	1.12 (0.31)	1.16 (0.32)	1.33 (0.43)	1.2 (0.45)
Len Short Zscore	-0.02 (0.97)	0.03 (1.03)	0.04 (1.08)	0.04 (1.01)
Num Short MN	0.27 (0.16)	0.29 (0.18)	0.41 (0.25)	0.33 (0.27)
Num Short SD	0.49 (0.13)	0.51 (0.13)	0.59 (0.15)	0.51 (0.19)
Num Short Zscore	-0.01 (0.98)	0.04 (1.03)	0.04 (1.09)	0.03 (0.99)
Sleep percentage MN	0.31 (0.04)	0.31 (0.05)	0.33 (0.02)	0.31 (0.02)
Sleep percentage SD	0.1 (0.03)	0.11 (0.03)	0.11 (0.02)	0.11 (0.03)
Sleep percentage Zscore	0.03 (0.99)	0.09 (1.02)	0 (1.14)	-0.11 (0.99)
Sleep amplitude MN	0.74 (0.08)	0.76 (0.09)	0.76 (0.03)	0.71 (0.06)
Sleep amplitude SD	0.18 (0.06)	0.2 (0.06)	0.2 (0.04)	0.17 (0.03)
Sleep amplitude Zscore	0.02 (0.98)	0.09 (1.04)	-0.04 (1.13)	-0.14 (0.94)
Circadian phase MN	23.39 (1.44)	23.44 (1.4)	23.62 (1.57)	24.53 (1.96)
Circadian phase SD	1.17 (0.58)	1.48 (0.89)	1.52 (0.84)	1.97 (1.82)
Circadian phase Zscore	0.03 (1)	0.09 (0.97)	-0.29 (1.1)	0.09 (1.06)
Circadian amplitude MN	1.03 (0.04)	1.01 (0.04)	1.01 (0.05)	1.01 (0.06)
Circadian amplitude SD	0.08 (0.02)	0.09 (0.02)	0.09 (0.03)	0.09 (0.03)
Circadian amplitude Zscore	0.04 (0.97)	0.03 (0.99)	-0.03 (1.16)	-0.03 (0.95)

Data are expressed as mean (standard deviation). There is no unit for features ending with 'Zscore'. For other features, the unit is 'hours' for features starting with 'ST', 'WT', or 'Len'. Unit of 'Circadian phase MN' is time. For example, Circadian phase MN of 23 means that the estimated circadian phase (DLMO) was 11 pm.

Supplementary Table 4. Demographic and baseline clinical characteristics of patients included in subgroup analysis for depressive episodes without changes in medication

Characteristic	Patients with recurrences	Patients without recurrences	<i>p</i> -value
N (%)	15 (14%)	93 (86%)	
Sex, female (n, %)	10 (67%)	46 (49%)	0.27
Ethnicity			
Korean (n, %)	15 (100%)	93 (100%)	
Age at baseline, years (mean, SD)	23.9 (3.60)	23.7 (4.01)	0.64
Age of onset, years (mean, SD)	17.9 (6.30)	18.6 (5.03)	0.58
Durations of clinical follow-up, d (mean, SD)	739 (345)	548 (374)	0.040*
Diagnosis (n, %)			0.64
MDD	4 (27%)	35 (38%)	
BD1	5 (33%)	22 (24%)	
BD2	6 (40%)	36 (39%)	
MADRS at baseline (mean, SD)	14.0 (10.4)	13.1 (10.5)	0.66
YMRS at baseline (mean, SD)	1.73 (2.96)	2.34 (3.38)	0.55
CSM at baseline (mean, SD)	28.9 (6.86)	29.8 (7.81)	0.65
GSS at baseline (mean, SD)	6.20 (4.66)	5.46 (4.53)	0.58
Comorbidities			
Substance use disorders	1 (6.7%)	4 (4.3%)	0.53
Anxiety disorders	6 (40%)	22 (24%)	0.21
OC related disorders	1 (6.7%)	5 (5.4%)	1
PTSD	0 (0%)	2 (2.2%)	1
Eating disorder	1 (6.7%)	2 (2.2%)	0.36
Somatic symptom and related disorder	1 (6.7%)	2 (2.2%)	0.36

Data are expressed as means (standard deviation). The *p*-values are results of wilcoxon rank sum test and Fisher's exact test for two groups comparison. MDD major depressive disorder, BD1 bipolar I disorder, BD2 bipolar II disorder, MADRS Montgomery–Åsberg Depression Rating Scale, YMRS Young Mania Rating Scale, CSM Composite Scale of Morningness, GSS Global Seasonality Score, OC obsessive-compulsive, PTSD posttraumatic stress disorder. * *p*-value <0.05.

Supplementary Table 5. Comparison of Circadian phase Z-score across each type of mood episode and diagnosis

	MDD	BD2	BD1
Inter	-0.02 (0.99)	-0.05 (0.99)	0.2 (1)
Depressive	0.15 (1.06)	0 (0.91)	0.24 (0.91)
Hypomanic		0.07 (0.98)	0.12 (1.21)
Manic			-0.29 (1.1)

Data are expressed as mean (standard deviation). The Circadian phase Zscore during depressive episodes in MDD and BD1 patients is higher than that of during the inter-episode period. In addition, the Circadian phase Zscore is lower during manic episode in BD1 patients.

Supplementary Table 6. Effective size of Circadian phase Zscore between two seasons.

	Spring	Summer	Fall	Winter
Spring		-0.03	-0.01	0.11
Summer	0.03		0.02	0.14
Fall	0.01	-0.02		0.13
Winter	-0.11	-0.14	-0.13	

The pairwise Cohen's d effective size between two different seasons was calculated for the Circadian phase Zscore. Positive values indicate that the Circadian phase Zscore of the comparison group (column) is higher than that of the reference group (row). All absolute values of effective sizes were below 0.2, implying that no distinct seasonal pattern of Circadian phase Zscore exists in our dataset.

Supplementary Table 7. Hyperparameters of trained XGBoost models

Parameters	Search range	Depressive	Manic	Hypomanic
objective		binary:logistic	binary:logistic	binary:logistic
colsample_bytree	(0.5, 0.99)	0.5	0.99	0.5
gamma	(0.01, 1)	1	0.01	1
learning_rate	(0.01, 1)	0.01	0.01	0.01
max_delta_step	(0, 0.5)	0	0	0
max_depth	(2, 10)	10	10	10
min_child_weight	(1, 10)	7.43	10	10
n_estimators	(500, 1000)	560	987	648
scale_pos_weight	(1, 30)	8.45	1	19.32
subsample	(0.7, 0.8)	0.8	0.8	0.7

Hyperparameters for training prediction models of Depressive, Manic, and Hypomanic episodes. Other parameters were set to the default values of Python ‘xgboost’ package version 1.7.3.

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No.	Recommendation	Page No.	Relevant text from manuscript
Title and abstract	1	(a) Indicate the study’s design with a commonly used term in the title or the abstract	1	Abstract
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	1	Abstract
Introduction				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	1	Introduction: paragraphs 1-3
Objectives	3	State specific objectives, including any prespecified hypotheses	1,2	Introduction: paragraph 4
Methods				
Study design	4	Present key elements of study design early in the paper	1,2 9	Introduction: paragraph 4 Methods: section 1
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	9,10	Methods: sections 1,2
Participants	6	(a) Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case-control study—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants (b) Cohort study—For matched studies, give matching criteria and number of exposed and unexposed Case-control study—For matched studies, give matching criteria and the number of controls per case	9,10	Cohort study: Methods: sections 2,3
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	9,10	Methods: section 2
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	9,10	Methods: section 2
Bias	9	Describe any efforts to address potential sources of bias	10	Methods: section 8
Study size	10	Explain how the study size was arrived at	2	Figure 2

Continued on next page

Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	10 2,3	Methods: section 4 Results: section 2
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	10,11	Methods: sections 5-9
		(b) Describe any methods used to examine subgroups and interactions	5	Results: section 6
		(c) Explain how missing data were addressed	2,3	Results: section 2
		(d) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed	9	Methods: section 1
		<i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed		
		<i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy		
		(e) Describe any sensitivity analyses		Not applicable
Results				
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	2 2 10	Results: section 1 Figure 2 Methods: section 3
		(b) Give reasons for non-participation at each stage	2 10	Results: section 1 Methods: section 3
		(c) Consider use of a flow diagram	2	Figure 2
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	2	Results: section 1 Supplementary Table 1
		(b) Indicate number of participants with missing data for each variable of interest		Not applicable All missing data were excluded
		(c) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)	2	Results: section 1
Outcome data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time	2	Results: section 1
		<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure		
		<i>Cross-sectional study</i> —Report numbers of outcome events or summary measures		
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	3,4	Results: section 3
		(b) Report category boundaries when continuous variables were categorized		Not applicable
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period		Not applicable

Continued on next page

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	5 8	Results: section 5,6 Figure 6
Discussion				
Key results	18	Summarise key results with reference to study objectives	5	Discussion: paragraphs 1,2
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	9	Discussion: paragraphs 7,8
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	6,7	Discussion: paragraphs 3,4,5
Generalisability	21	Discuss the generalisability (external validity) of the study results	8,9	Discussion: paragraph 6
Other information				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	13	Acknowledgement

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at www.strobe-statement.org.