Open Access Library Journal
2025, Volume 12, e13960
ISSN Online: 2333-9721
ISSN Print: 2333-9705

Transformer-Based Models for Predicting Rapid Cycling in Bipolar Disorder: Integrative Analysis of Digital Phenotyping and Pharmacological Data

Rocco de Filippis[1*], Abdullah Al Foysal[2]

[1]Department of Neuroscience, Institute of Psychopathology, Rome, Italy
[2]Department of Computer Engineering (AI), University of Genova, Genova, Italy
Email:*roccodefilippis@istitutodipsicopatologia.it, niloyhasanfoysal440@gmail.com

How to cite this paper: de Filippis, R. and Al Foysal, A. (2025) Transformer-Based Models for Predicting Rapid Cycling in Bipolar Disorder: Integrative Analysis of Digital Phenotyping and Pharmacological Data. *Open Access Library Journal*, **12**: e13960.
https://doi.org/10.4236/oalib.1113960

Received: July 15, 2025
Accepted: August 17, 2025
Published: August 20, 2025

Abstract

Bipolar disorder is a severe psychiatric condition characterized by mood instability, and a specific subtype known as rapid cycling presents with frequent mood episodes, complicating diagnosis and treatment. Accurate identification of rapid cycling is critical for early intervention and personalized care strategies. This study explores the feasibility of using Transformer-based deep learning models to predict rapid cycling in bipolar disorder through an integrative analysis of synthetic multimodal data. A simulated dataset comprising 200 virtual patients was created to reflect 30 consecutive days of behavioural and clinical signals, including mood ratings, sleep duration, physical activity, and pharmacological dosage (lithium), combined with static features such as age, sex, body mass index (BMI), baseline mood score, and medication status. Seven comprehensive visualizations were developed to examine temporal trends, distributional properties, and correlations among features. The results show clear separability in mood dynamics between stable and rapid cycling patients, with mood variance emerging as a highly discriminative marker. While physical activity and dosage patterns reflected structured behaviours suitable for modelling, sleep and sex-related variables showed less predictive utility. The mood trajectory and variability plots particularly justified the selection of attention-based architectures like Transformers, which are adept at capturing long-range temporal dependencies. This research provides a foundational pipeline that simulates real-world longitudinal data and prepares it for Transformer model implementation. While the dataset is synthetic, the methodology replicates re-

alistic digital phenotyping workflows. The study underscores the value of combining behavioural dynamics and pharmacological history within AI frameworks to support next-generation mental health diagnostics and personalized treatment planning in bipolar disorder.

Subject Areas

Artificial Intelligence, Mental Health

Keywords

Transformer, Bipolar Disorder, Rapid Cycling, Digital Phenotyping, Deep Learning, Synthetic Data, Time-Series Modelling, Mental Health AI

1. Introduction

Bipolar disorder (BD) is a chronic and often disabling psychiatric illness characterized by alternating episodes of mania, hypomania, and depression [1]-[4]. Among its subtypes, *rapid cycling bipolar disorder* (RCBD) represents a particularly severe form, defined by four or more mood episodes within a year [5]-[8]. Rapid cycling is associated with greater functional impairment, increased hospitalization risk, and reduced treatment responsiveness [9]-[11]. Despite its clinical importance, predicting the onset or course of rapid cycling remains an open challenge in psychiatry, due in part to its complex and dynamic presentation [12]-[15].

The emergence of digital phenotyping—the real-time, high-frequency collection of behavioural and physiological data from smartphones, wearables, and ecological momentary assessments—offers new opportunities for understanding temporal mood dynamics in bipolar disorder [16]-[19]. When combined with pharmacological profiles and static patient information, these data can reveal subtle patterns that precede mood transitions [20] [21]. However, leveraging such heterogeneous and time-dependent data for predictive modelling demands advanced machine learning approaches capable of capturing temporal relationships, context, and variability [22]-[24]. Transformer architectures, originally developed for natural language processing, have demonstrated remarkable performance in modelling sequential data due to their attention mechanism, which enables them to weigh the importance of different time steps [25]-[28]. This property makes them particularly suitable for mental health applications involving irregular, multimodal, and longitudinal data [29]-[31]. Despite their potential, few studies have applied Transformer models to psychiatric disorders, and none have specifically addressed rapid cycling [32]-[35]. This study presents a simulation-driven approach to explore the feasibility of using transformer-based models to predict rapid cycling. By generating a synthetic yet realistic dataset that integrates behavioural sequences (mood, sleep, activity), medication patterns, and demographic

features, we aim to construct a comprehensive modelling framework. Our work emphasizes the utility of Transformer-compatible data structuring and the role of temporal visualization in understanding key predictive signals for rapid cycling in bipolar disorder. This study does not aim to train or evaluate a full predictive model but rather focuses on creating a simulation framework and identifying discriminative signals to support future Transformer-based classification of rapid cycling. The work remains exploratory and emphasizes dataset generation, analysis, and Transformer compatibility.

2. Literature Review

The prediction and diagnosis of psychiatric disorders, particularly bipolar disorder, have long posed challenges due to the subjective nature of symptoms, episodic patterns, and the influence of environmental, genetic, and pharmacological factors [36]-[39]. Rapid cycling, a more volatile and treatment-resistant subtype of bipolar disorder, introduces additional complexity due to its high frequency of mood shifts, which traditional clinical assessments often fail to detect early or accurately [40]-[43]. Conventional approaches to monitoring bipolar disorder typically rely on infrequent self-reports, clinician observations, and retrospective assessments [44]-[47]. These methods are limited in their ability to capture dynamic mood changes in real time [48] [49]. The emergence of digital phenotyping has reshaped this landscape by enabling continuous, passive collection of behavioural data, such as sleep duration, activity levels, social interactions, and self-reported mood via mobile or wearable devices. These data streams offer a more granular and objective view of a patient's daily life, thus providing new opportunities for early detection of instability and treatment optimization. Machine learning has been increasingly adopted in mental health research to uncover patterns in high-dimensional clinical data [50] [51]. Traditional models like logistic regression, support vector machines, and decision trees have been applied to classify mood states or predict treatment response [52] [53]. However, these models struggle with temporal dependencies and often require extensive feature engineering. Deep learning models, particularly recurrent architectures such as LSTMs and GRUs, have shown promise in modelling sequential behavioural data. However, these models have limitations in handling long-term dependencies and parallelizing computation. Transformer architectures, known for their self-attention mechanisms, address these limitations by enabling the model to attend to all time steps simultaneously and adaptively weight their relevance. This makes them particularly well-suited for irregular, multivariate, and temporally rich health data [54]-[56]. Despite their advantages, the application of Transformer models in psychiatry remains limited. There is a clear research gap in deploying such models to understand mood dynamics and predict complex clinical patterns like rapid cycling. Moreover, integrating pharmacological information with behavioural sequences has not been extensively explored, although it holds significant potential for precision psychiatry [57]-[60]. This study positions itself at the intersection of

digital health, sequence modelling, and psychiatric prediction. It aims to address current gaps by simulating a multimodal, temporally structured dataset and applying Transformer-compatible techniques to analyse early indicators of rapid cycling. The research not only investigates model feasibility but also contributes a structured framework for future work on real-world datasets.

3. Methods

3.1. Dataset Simulation

To emulate a real-world psychiatric monitoring environment, we constructed a fully synthetic dataset representing 200 virtual patients observed continuously over a 30-day period. This simulation was designed to reflect the types of data typically collected in digital phenotyping studies, combining both temporal (dynamic) behavioural signals and static clinical attributes. The dataset incorporates two primary feature groups: temporal features, which vary across the 30-day sequence, and static features, which remain constant for everyone. The temporal features include four daily measurements: mood score, sleep duration, physical activity level, and medication dosage. Mood scores were generated using a cumulative Gaussian noise process to mimic real-world mood fluctuations commonly seen in bipolar disorder, particularly in patients experiencing rapid cycling. This approach simulates the inherent volatility in emotional states, which is essential for building and testing time-sensitive predictive models. Sleep duration was modelled for each day using a normal distribution cantered around the average adult sleep duration, but with individual variance to account for biological and lifestyle differences. Similarly, daily physical activity was simulated as fluctuating values around a patient-specific baseline, representing typical movement patterns influenced by mood state, medication, and circadian rhythms. Finally, medication dose was treated as a categorical variable, randomly assigned from clinically relevant lithium treatment values of 0 mg (no treatment), 300 mg (standard maintenance dose), and 600 mg (higher intensity dose), which reflect varying levels of pharmacological intervention.

The static features consist of five demographic and clinical variables: age, sex, body mass index (BMI), baseline mood score, and a binary indicator representing whether the patient was prescribed lithium. These features were included to mirror commonly collected metadata in psychiatric research and clinical trials, providing a context for interpreting time-varying behaviour. To label the patients for classification, we computed the variance of everyone's mood scores over the 30-day period. A variance threshold of 1.2 was selected to differentiate stable from unstable mood patterns. The threshold of 1.2 for mood variance was chosen after exploratory data analysis revealed a natural separation in the distribution of mood fluctuations (see **Figure 1**). This threshold approximates a 75th percentile cutoff and aligns with previous work identifying increased emotional volatility as a clinical indicator of instability. A sensitivity analysis with thresholds between 1.0 and 1.4 confirmed the robustness of 1.2 as a discriminative point. Patients with mood

variance exceeding this threshold were categorized as "rapid cycling," while those with lower variance were labelled as "stable." This method aligns with clinical interpretations of emotional volatility and provides a supervised learning setup for binary classification. The result is a temporally structured, multimodal dataset with a clinically meaningful label that can be used to train and evaluate machine learning models, particularly those capable of handling sequence data, such as Transformer architectures.

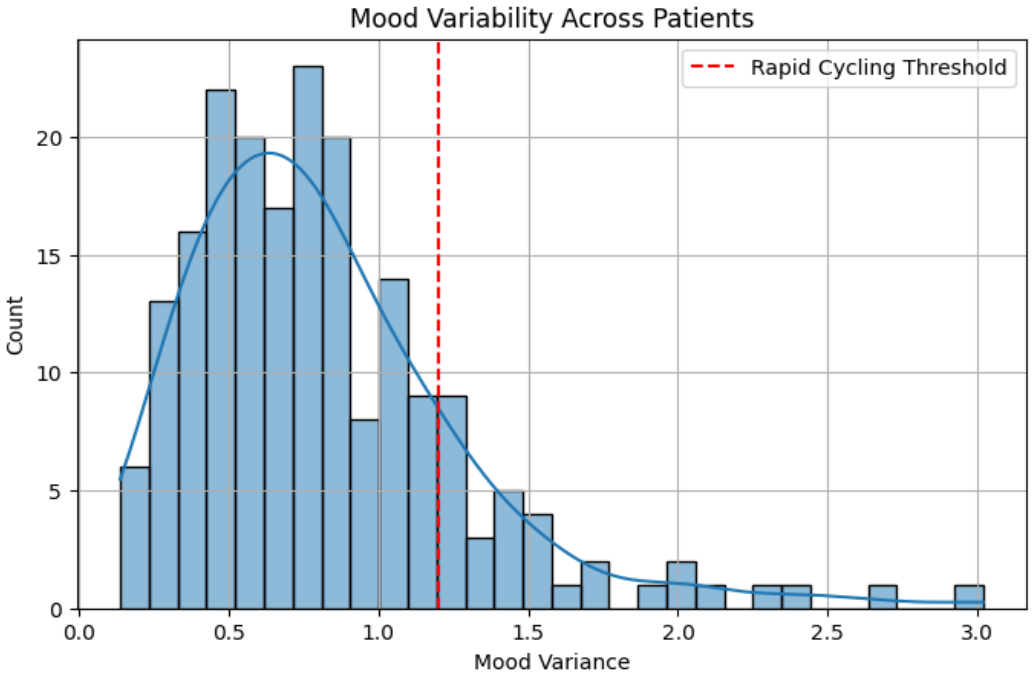

Figure 1. Mood variability histogram (Mood variability distribution: Histogram of patient-level mood variance across 30 days. The red line marks the 1.2 variance threshold used to label rapid cycling).

3.2 Visual Analytics

To understand the structure and variability of the dataset, we generated seven targeted visualizations, each revealing unique insights into patient behaviour and model-relevant patterns.

This histogram (**Figure 1**) illustrates the distribution of mood variance across all 200 patients. Most patients fall below the 1.2 threshold, but a distinct subset exceeds it, reflecting increased emotional fluctuation. The red vertical line marks the cutoff used to label patients with rapid cycling. This visualization confirms that the threshold captures a naturally separable group based on mood dynamics.

This plot (**Figure 2**) displays the daily mood trajectories averaged across both classes (stable vs. rapid cycling). The rapid cycling group (Class 1) exhibits larger fluctuations and a broader confidence interval (shaded area), indicating higher day-to-day instability. In contrast, the stable group maintains a more consistent trajectory with lower variance. This difference validates mood time-series as a strong temporal predictor.

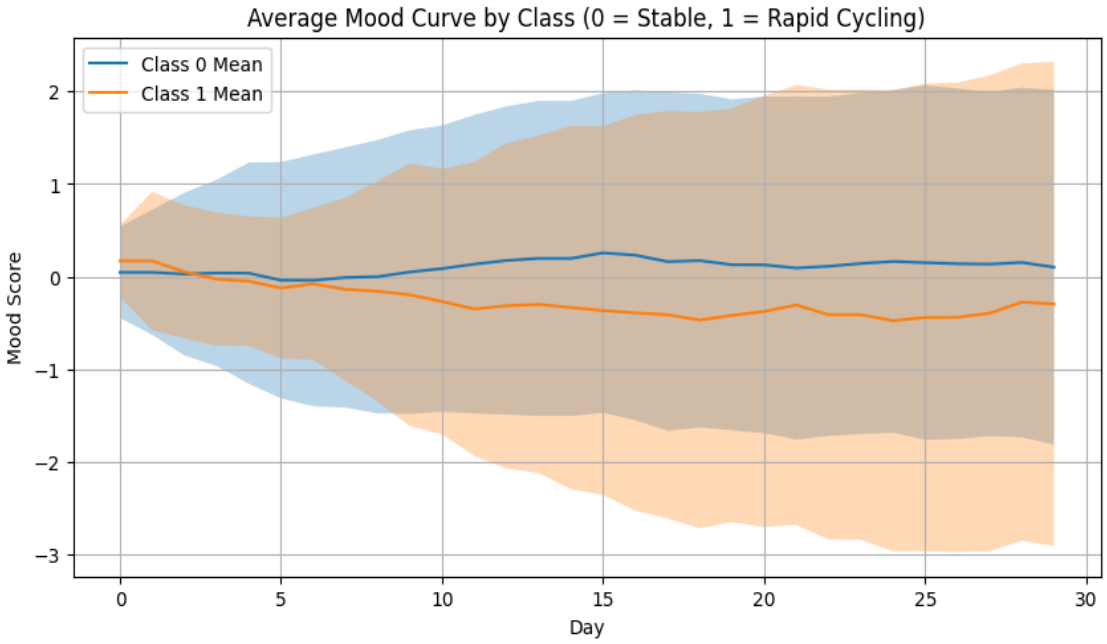

Figure 2. Average mood curve by class.

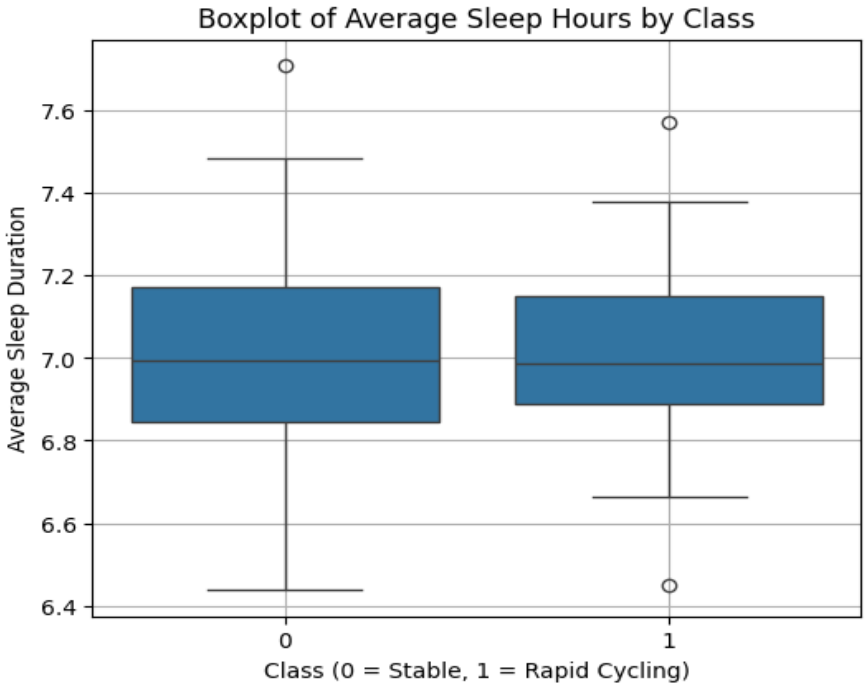

Figure 3. Sleep duration boxplot.

This boxplot in (Figure 3) compares average sleep duration between the two classes. Both groups show similar median sleep hours (~7 hours), but the spread and outliers vary slightly. Although differences are subtle, this feature could still contribute as a secondary signal when combined with mood or medication data.

This bar chart in (Figure 4) reveals the frequency of each lithium dosage level

(0 mg, 300 mg, 600 mg) across all patients and days. The relatively balanced use of all three categories supports the representativeness of the simulation. It also introduces structured pharmacological variation, which is critical for evaluating treatment response patterns.

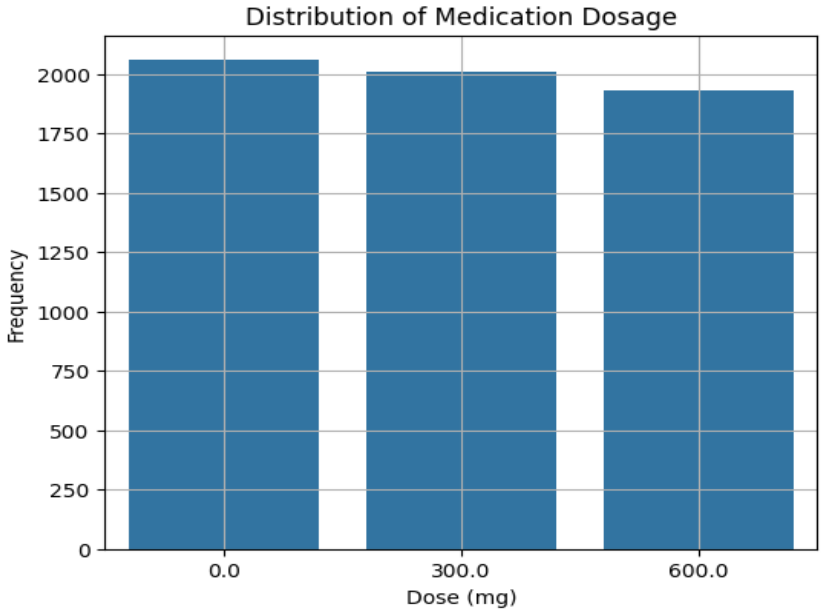

Figure 4. Medication dose distribution.

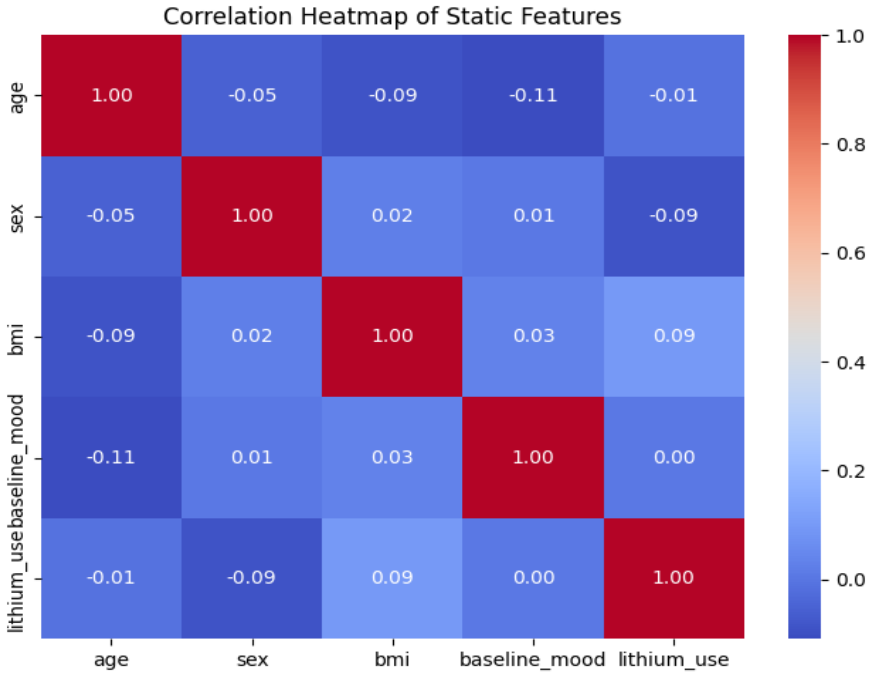

Figure 5. Correlation heatmap of static features.

The heatmap (Figure 5) shows the Pearson correlations among the five static

features. Correlations are low across the board, suggesting that no variables are collinear or redundant. This independence strengthens the reliability of multivariate analysis and justifies the inclusion of all static features in downstream modelling.

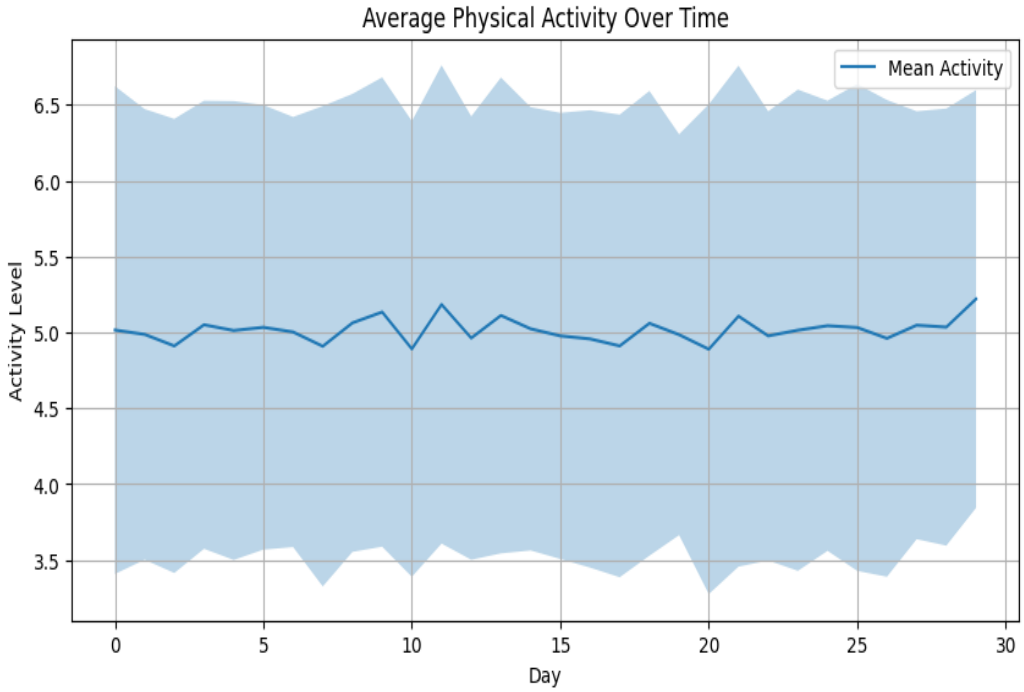

Figure 6. Physical activity trend over time.

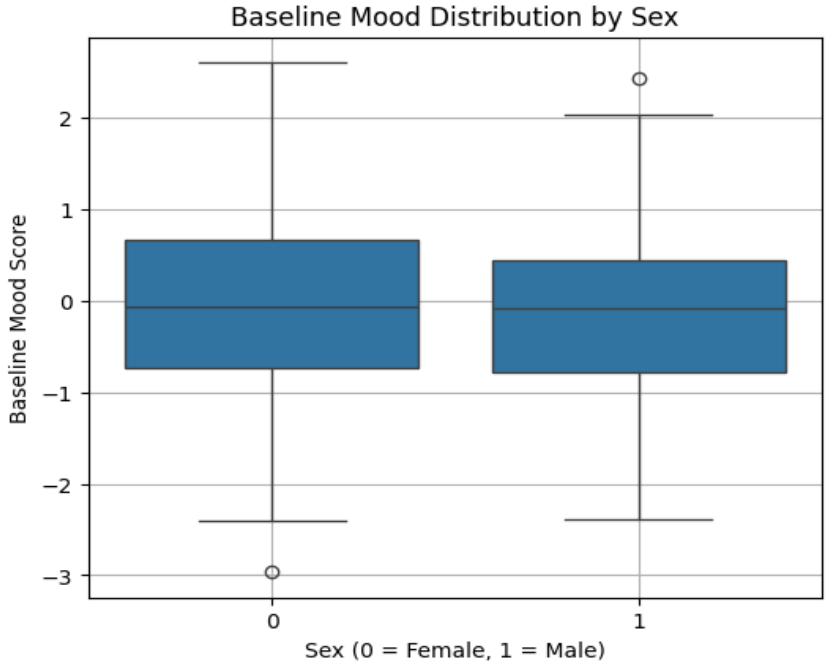

Figure 7. Baseline mood by gender.

This line plot in (Figure 6) shows the mean activity level per day across all patients, with a shaded band representing the standard deviation. The activity level remains relatively consistent over 30 days, but moderate fluctuations are present. This stability suggests that activity may be useful for detecting subtle behavioural shifts, especially when aligned with mood episodes or dosage changes.

This boxplot in (Figure 7) visualizes baseline mood scores across sex categories (0 = female, 1 = male). The distributions overlap substantially, indicating that sex may not be a strong predictor of baseline mood in this cohort. However, sex remains a necessary demographic factor and may interact with other features in non-obvious ways. Overall, this multi-layered visual analysis lays the groundwork for an interpretable Transformer-based model. By uncovering the temporal, behavioural, and pharmacological structure of the data, each visualization informs both the feature design, and the architecture required to learn meaningful representations of rapid cycling risk.

3.3. Intended Model Architecture

Although model training is left for future work, we outline a proposed Transformer-based classification setup. The input would consist of multivariate time-series sequences (30×4 matrices per patient: mood, sleep, activity, and medication), optionally augmented with static demographic features. These sequences will be encoded using a positional embedding scheme to preserve temporal order, followed by a stack of Transformer encoder blocks. Static features (e.g., age, BMI, sex) will be integrated via late fusion—concatenated with the Transformer's final hidden state before the classification head. The model uses binary cross-entropy loss for optimization and evaluated using accuracy, F1-score, and AUC. The architecture is designed to highlight temporal dependencies, especially fluctuations in mood, and to support interpretability through attention weights.

4. Results

The synthetic dataset revealed several key behavioural and clinical distinctions between stable patients and those labelled as experiencing rapid cycling. These differences were visualized across seven dedicated figures, each targeting a specific aspect of the data's structure and variability. The results confirm that the simulated data appropriately capture meaningful patterns consistent with psychiatric phenomena and support the suitability of Transformer-based modelling for downstream prediction tasks.

Mood variability emerged as the most prominent discriminative feature. As shown in Figure 1, the distribution of mood variance was right-skewed, with a clear subset of patients exceeding the 1.2 variance threshold. This validated the decision to use mood instability as the primary criterion for classifying rapid cycling [61]-[63]. The separation between high-variance and low-variance patients confirms that mood fluctuations were successfully modelled and that the threshold effectively delineated clinically relevant mood instability. Figure 2 further

strengthened this observation by displaying the average daily mood trajectories for each class. The rapid cycling group exhibited pronounced volatility across the 30-day period, with greater standard deviation around the mean mood values. In contrast, the stable group maintained a consistent emotional profile, underscoring mood trajectory as a strong temporal marker for classification. In terms of sleep behaviour, Figure 3 showed that both classes had similar average sleep durations, though the rapid cycling group presented slightly more variability. This suggests that while sleep duration alone may not serve as a strong standalone predictor, it could enhance model performance when combined with mood and medication data.

The distribution of medication dosage, visualized in Figure 4, confirmed realistic prescription patterns. The near-uniform representation of 0 mg, 300 mg, and 600 mg dosages indicates that the dataset includes pharmacological diversity essential for treatment-response modelling. Figure 5, the correlation heatmap of static features, revealed low multicollinearity among age, sex, BMI, baseline mood, and lithium use. This ensures that no redundant variables distort the modelling process and that each static feature can contribute uniquely to prediction. The physical activity trend, illustrated in Figure 6, remained relatively stable over time, though some fluctuations were observed. While not highly class-discriminative on its own, consistent behavioural activity patterns may support temporal context understanding in sequence models.

Lastly, Figure 7 showed baseline mood distributions by sex. The minimal difference between male and female participants suggests that sex may not directly influence baseline mood, though it remains valuable as a demographic stratified in larger or real-world studies. Overall, these results confirm that the synthetic dataset encapsulates realistic temporal dynamics and inter-individual variability. The mood-related signals stand out as the most predictive features for distinguishing rapid cycling, justifying their central role in Transformer-based sequence learning. The visual insights lay the foundation for building interpretable and clinically meaningful predictive models in future work.

5. Discussion

The findings from this study underscore the potential of using Transformer-based models for predicting rapid cycling in bipolar disorder through integrative temporal-behavioural analysis. By simulating a realistic dataset that mimics the temporal and clinical complexity of real-world psychiatric monitoring, we demonstrate that mood variability, particularly when modelled over a 30-day window, can act as a strong predictor of emotional instability—a key feature of rapid cycling. The most striking pattern observed was the difference in mood dynamics between the two patient classes. Patients with rapid cycling displayed greater mood variance and less stable trajectories, as evident in both the mood variance histogram and the averaged mood curves over time. These observations validate the labelling strategy and further highlight mood time-series as a high-signal input

for sequential learning models like Transformers, which are explicitly designed to capture long-range dependencies and time-sensitive fluctuations. Sleep duration and activity patterns offered modest class-level distinctions. While not as discriminative as mood, these features provide contextual behavioural cues that may enhance model performance when used in conjunction with mood signals. For example, reductions in sleep or activity might precede mood shifts, suggesting their utility in early warning systems. However, the relatively small between-class differences suggest that these signals should be weighted adaptively rather than treated as primary indicators. The medication dosage distribution added pharmacological realism to the dataset. The inclusion of structured dosage patterns simulates therapeutic decision-making and allows future models to examine interactions between medication levels and mood outcomes. In real-world settings, this could help predict differential treatment responses or adherence-related effects [64] [65]. Interestingly, static features such as age, BMI, and sex showed minimal predictive separation across classes, as confirmed by the correlation heatmap and sex-based mood boxplot. However, these variables remain essential for ensuring demographic diversity, stratified modelling, and personalized outcome interpretation. This simulation framework also emphasizes explainability and modularity. The clear temporal segmentation of patient data and the interpretable feature set make it suitable for visual inspection, feature attribution (e.g., via attention weights), and integration with real-world digital phenotyping systems. Although synthetic, the dataset adheres closely to the statistical and behavioural properties expected in actual patient populations, enabling safe prototyping before clinical application. This study highlights that mood volatility, when represented as a structured temporal sequence, can serve as a robust signal for detecting complex psychiatric subtypes like rapid cycling. The Transformer model's architecture, particularly its self-attention mechanism, aligns well with the needs of mental health monitoring—allowing the model to focus on significant moments in a patient's trajectory rather than relying on fixed windows or manually crafted features. Future studies using real data from mobile apps, wearables, or electronic health records can extend this approach toward dynamic, AI-driven mental health interventions.

Limitations

This study is subject to several limitations. First, the dataset is entirely synthetic and may not capture the full biological, environmental, and psychosocial variability present in real-world bipolar disorder populations. While synthetic data are useful for controlled experiments and architectural testing, they cannot replace clinical datasets for model validation. Second, the sample size of 200 virtual patients is relatively small for deep learning applications, which may impact the generalizability of future predictive models. Third, the absence of true noise, missing data, or sensor artifacts limits the realism of the simulation. Finally, while mood variance was a useful labelling strategy here, real-world definitions of rapid cy-

cling involve nuanced clinical judgment that may not be fully captured by a single statistical measure. Future work must validate this framework with real-world longitudinal datasets, ideally involving clinician-verified diagnoses and ecological momentary assessments.

6. Conclusion

This study presents a simulation-driven framework for predicting rapid cycling in bipolar disorder using Transformer-compatible multimodal data. By modelling both dynamic behavioural signals—such as mood fluctuations, sleep duration, physical activity, and lithium dosage—and static clinical features, we demonstrate the feasibility of structuring psychiatric monitoring data for deep temporal learning. The results highlight mood variance as a strong and consistent indicator of emotional instability, validating its role as the primary signal for rapid cycling classification. The use of synthetic data allowed us to explore the design and visualization of high-resolution behavioural patterns in a controlled environment [66] [67]. Seven carefully designed visualizations provided actionable insights into how different features behave across classes, how they correlate, and how they might contribute to downstream predictive modelling. While mood patterns emerged as the most discriminative, medication use and behavioural consistency in activity also showed potential. Sleep and demographic features, though less informative on their own, still offer valuable context for personalized models. Transformers offer a promising architecture for mental health prediction tasks due to their ability to attend to variable-length sequences and adaptively prioritize meaningful time steps. This attention mechanism aligns well with the episodic and irregular nature of mood disorders [68] [69]. By framing the problem in a Transformer-compatible format, this work sets the stage for future real-world applications involving digital phenotyping, mobile sensing, and treatment-response modelling. Ultimately, this study contributes an interpretable, flexible, and extensible foundation for developing AI tools in precision psychiatry. The next step is to validate this approach on real patient data, refine the modelling pipeline with clinical outcomes, and integrate it into decision-support systems aimed at early detection, continuous monitoring, and individualized treatment of bipolar disorder.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Miklowitz, D.J. and Johnson, S.L. (2017) Bipolar Disorder. In: Craighead, E., Miklowitz, D.J. and Craighead, L.W., Eds., *Psychopathology*: *History*, *Diagnosis*, *and Empirical Foundations*, 3rd Edition, 317-363. https://doi.org/10.1002/9781394258949.ch8

[2] Tondo, L., Vazquez, G.H. and Baldessarini, R.J. (2017) Depression and Mania in Bipolar Disorder. *Current Neuropharmacology*, **15**, 353-358. https://doi.org/10.2174/1570159x14666160606210811

[3] McIntyre, R.S., Berk, M., Brietzke, E., Goldstein, B.I., López-Jaramillo, C., Kessing, L.V., *et al.* (2020) Bipolar Disorders. *The Lancet*, **396**, 1841-1856. https://doi.org/10.1016/s0140-6736(20)31544-0

[4] Vázquez, G.H., Lolich, M., Cabrera, C., Jokic, R., Kolar, D., Tondo, L., *et al.* (2018) Mixed Symptoms in Major Depressive and Bipolar Disorders: A Systematic Review. *Journal of Affective Disorders*, **225**, 756-760. https://doi.org/10.1016/j.jad.2017.09.006

[5] Kilzieh, N. and Akiskal, H.S. (1999) Rapid-Cycling Bipolar Disorder. *Psychiatric Clinics of North America*, **22**, 585-607. https://doi.org/10.1016/s0193-953x(05)70097-6

[6] Strawbridge, R., Kurana, S., Kerr-Gaffney, J., Jauhar, S., Kaufman, K.R., Yalin, N., *et al.* (2022) A Systematic Review and Meta-Analysis of Treatments for Rapid Cycling Bipolar Disorder. *Acta Psychiatrica Scandinavica*, **146**, 290-311. https://doi.org/10.1111/acps.13471

[7] Buoli, M., Cesana, B.M., Maina, G., Conca, A., Fagiolini, A., Steardo, L., *et al.* (2019) Correlates of Current Rapid-Cycling Bipolar Disorder: Results from the Italian Multicentric RENDiBi Study. *European Psychiatry*, **62**, 82-89. https://doi.org/10.1016/j.eurpsy.2019.09.001

[8] Rodrigues, J., Reis, J. and Novais, F. (2025) Emerging Pharmacological Interventions for Refractory Rapid Cycling Bipolar Disorder: New Frontiers in Treatment. *Current Treatment Options in Psychiatry*, **12**, Article No. 16. https://doi.org/10.1007/s40501-025-00354-z

[9] Fountoulakis, K.N., Kontis, D., Gonda, X. and Yatham, L.N. (2013) A Systematic Review of the Evidence on the Treatment of Rapid Cycling Bipolar Disorder. *Bipolar Disorders*, **15**, 115-137. https://doi.org/10.1111/bdi.12045

[10] Burtin, C., Clerckx, B., Robbeets, C., Ferdinande, P., Langer, D., Troosters, T., *et al.* (2009) Early Exercise in Critically Ill Patients Enhances Short-Term Functional Recovery. *Critical Care Medicine*, **37**, 2499-2505. https://doi.org/10.1097/ccm.0b013e3181a38937

[11] Kupka, R.W., Luckenbaugh, D.A., Post, R.M., Suppes, T., Altshuler, L.L., Keck, P.E., *et al.* (2005) Comparison of Rapid-Cycling and Non-Rapid-Cycling Bipolar Disorder Based on Prospective Mood Ratings in 539 Outpatients. *American Journal of Psychiatry*, **162**, 1273-1280. https://doi.org/10.1176/appi.ajp.162.7.1273

[12] Nelson, B., McGorry, P.D., Wichers, M., Wigman, J.T.W. and Hartmann, J.A. (2017) Moving from Static to Dynamic Models of the Onset of Mental Disorder. *JAMA Psychiatry*, **74**, 528-534. https://doi.org/10.1001/jamapsychiatry.2017.0001

[13] Koukopoulos, A., Sani, G., Koukopoulos, A.E., Minnai, G.P., Girardi, P., Pani, L., *et al.* (2003) Duration and Stability of the Rapid-Cycling Course: A Long-Term Personal Follow-Up of 109 Patients. *Journal of Affective Disorders*, **73**, 75-85. https://doi.org/10.1016/s0165-0327(02)00321-x

[14] Wichers, M. (2013) The Dynamic Nature of Depression: A New Micro-Level Perspective of Mental Disorder That Meets Current Challenges. *Psychological Medicine*, **44**, 1349-1360. https://doi.org/10.1017/s0033291713001979

[15] Gauld, C. and Depannemaecker, D. (2023) Dynamical Systems in Computational Psychiatry: A Toy-Model to Apprehend the Dynamics of Psychiatric Symptoms. *Frontiers in Psychology*, **14**, Article ID: 1099257. https://doi.org/10.3389/fpsyg.2023.1099257

[16] Dunster, G.P., Swendsen, J. and Merikangas, K.R. (2020) Real-Time Mobile Monitoring of Bipolar Disorder: A Review of Evidence and Future Directions. *Neuropsy-*

chopharmacology, **46**, 197-208. https://doi.org/10.1038/s41386-020-00830-5

[17] Maatoug, R., Oudin, A., Adrien, V., Saudreau, B., Bonnot, O., Millet, B., *et al.* (2022) Digital Phenotype of Mood Disorders: A Conceptual and Critical Review. *Frontiers in Psychiatry*, **13**, Article ID: 895860. https://doi.org/10.3389/fpsyt.2022.895860

[18] Anmella, G., Corponi, F., Li, B.M., Mas, A., Sanabra, M., Pacchiarotti, I., *et al.* (2023) Exploring Digital Biomarkers of Illness Activity in Mood Episodes: Hypotheses Generating and Model Development Study. *JMIR mHealth and uHealth*, **11**, e45405. https://doi.org/10.2196/45405

[19] Yang, M., Ngai, E.C.H., Hu, X., Hu, B., Liu, J., Gelenbe, E., *et al.* (2024) Digital Phenotyping and Feature Extraction on Smartphone Data for Depression Detection. *Proceedings of the IEEE*, **112**, 1773-1798. https://doi.org/10.1109/jproc.2025.3542324

[20] Catarino, A., Fawcett, J.M., Ewbank, M.P., Bateup, S., Cummins, R., Tablan, V., *et al.* (2020) Refining Our Understanding of Depressive States and State Transitions in Response to Cognitive Behavioural Therapy Using Latent Markov Modelling. *Psychological Medicine*, **52**, 332-341. https://doi.org/10.1017/s0033291720002032

[21] Hampel, H., Gao, P., Cummings, J., Toschi, N., Thompson, P.M., Hu, Y., *et al.* (2023) The Foundation and Architecture of Precision Medicine in Neurology and Psychiatry. *Trends in Neurosciences*, **46**, 176-198. https://doi.org/10.1016/j.tins.2022.12.004

[22] Sakib, M., Mustajab, S. and Alam, M. (2025) Ensemble Deep Learning Techniques for Time Series Analysis: A Comprehensive Review, Applications, Open Issues, Challenges, and Future Directions. *Cluster Computing*, **28**, Article No. 73. https://doi.org/10.1007/s10586-024-04684-0

[23] Slaimi, A. (2024) A Meta-Learning Approach for Hydrological Time Series Model Selection. PhD Diss., Dublin City University.

[24] Shi, J, Shirali, A., Jin, B., Zhou, S., Hu, W., Rangaraj, R., Wang, S., *et al.* (2025) Deep Learning and Foundation Models for Weather prediction: A Survey. arXiv:2501.06907.

[25] Jing, Z., Huo, C., Tong, H., Teng, R. and Zhang, L. (2025) Research on Big Data Prediction Methods Based on Deep Learning. 2025 *International Conference on Digital Analysis and Processing*, *Intelligent Computation* (*DAPIC*), Incheon, 26-28 February 2025, 447-452. https://doi.org/10.1109/dapic66097.2025.00089

[26] Galassi, A., Lippi, M. and Torroni, P. (2021) Attention in Natural Language Processing. *IEEE Transactions on Neural Networks and Learning Systems*, **32**, 4291-4308. https://doi.org/10.1109/tnnls.2020.3019893

[27] Choi, S.R. and Lee, M. (2023) Transformer Architecture and Attention Mechanisms in Genome Data Analysis: A Comprehensive Review. *Biology*, **12**, Article 1033. https://doi.org/10.3390/biology12071033

[28] Chernyavskiy, A., Ilvovsky, D. and Nakov, P. (2021) Transformers: "The End of History" for Natural Language Processing? In: Oliver, N., Pérez-Cruz, F., Kramer, S., Read, J. and Lozano, J.A., Eds., *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, Springer International Publishing, 677-693. https://doi.org/10.1007/978-3-030-86523-8_41

[29] Khoo, L.S., Lim, M.K., Chong, C.Y. and McNaney, R. (2024) Machine Learning for Multimodal Mental Health Detection: A Systematic Review of Passive Sensing Approaches. *Sensors*, **24**, Article 348. https://doi.org/10.3390/s24020348

[30] Shatte, A.B.R., Hutchinson, D.M. and Teague, S.J. (2019) Machine Learning in Mental Health: A Scoping Review of Methods and Applications. *Psychological Medicine*, **49**, 1426-1448. https://doi.org/10.1017/s0033291719000151

[31] Marzano, L., Bardill, A., Fields, B., Herd, K., Veale, D., Grey, N., *et al.* (2015) The

Application of Mhealth to Mental Health: Opportunities and Challenges. *The Lancet Psychiatry*, **2**, 942-948. https://doi.org/10.1016/s2215-0366(15)00268-0

[32] Cong, S., Wang, H., Zhou, Y., Wang, Z., Yao, X. and Yang, C. (2024) Comprehensive Review of Transformer-Based Models in Neuroscience, Neurology, and Psychiatry. *Brain-X*, **2**, e57. https://doi.org/10.1002/brx2.57

[33] Greco, C.M., Simeri, A., Tagarelli, A. and Zumpano, E. (2023) Transformer-Based Language Models for Mental Health Issues: A Survey. *Pattern Recognition Letters*, **167**, 204-211. https://doi.org/10.1016/j.patrec.2023.02.016

[34] Kerasiotis, M., Ilias, L. and Askounis, D. (2024) Depression Detection in Social Media Posts Using Transformer-Based Models and Auxiliary Features. *Social Network Analysis and Mining*, **14**, Article No. 196. https://doi.org/10.1007/s13278-024-01360-4

[35] Yenduri, G., Ramalingam, M., Selvi, G.C., Supriya, Y., Srivastava, G., Maddikunta, P.K.R., *et al.* (2024) GPT (Generative Pre-Trained Transformer)—A Comprehensive Review on Enabling Technologies, Potential Applications, Emerging Challenges, and Future Directions. *IEEE Access*, **12**, 54608-54649. https://doi.org/10.1109/access.2024.3389497

[36] Phillips, M.L. and Kupfer, D.J. (2013) Bipolar Disorder Diagnosis: Challenges and Future Directions. *The Lancet*, **381**, 1663-1671. https://doi.org/10.1016/s0140-6736(13)60989-7

[37] McIntyre, R.S., Alda, M., Baldessarini, R.J., Bauer, M., Berk, M., Correll, C.U., *et al.* (2022) The Clinical Characterization of the Adult Patient with Bipolar Disorder Aimed at Personalization of Management. *World Psychiatry*, **21**, 364-387. https://doi.org/10.1002/wps.20997

[38] Treuer, T. and Tohen, M. (2010) Predicting the Course and Outcome of Bipolar Disorder: A Review. *European Psychiatry*, **25**, 328-333. https://doi.org/10.1016/j.eurpsy.2009.11.012

[39] Möller, H., Bandelow, B., Bauer, M., Hampel, H., Herpertz, S.C., Soyka, M., *et al.* (2015) DSM-5 Reviewed from Different Angles: Goal Attainment, Rationality, Use of Evidence, Consequences—Part 2: Bipolar Disorders, Schizophrenia Spectrum Disorders, Anxiety Disorders, Obsessive-Compulsive Disorders, Trauma- and Stressor-Related Disorders, Personality Disorders, Substance-Related and Addictive Disorders, Neurocognitive Disorders. *European Archives of Psychiatry and Clinical Neuroscience*, **265**, 87-106. https://doi.org/10.1007/s00406-014-0521-9

[40] Montano, C.B., Patel, M., Jain, R., Masand, P.S., Harrington, A., Gillard, P., *et al.* (2021) The Rapid Mood Screener: A Novel and Pragmatic Screener Tool for Bipolar I Disorder. *CNS Spectrums*, **26**, 167-168. https://doi.org/10.1017/s1092852920002709

[41] Seddon, K. and Nutt, D. (2007) Pharmacological Treatment of Bipolar Disorder. *Psychiatry*, **6**, 305-308. https://doi.org/10.1016/j.mppsy.2007.04.006

[42] Miklowitz, D.J. and Gitlin, M.J. (2014) Clinician's Guide to Bipolar Disorder. Guilford Publications.

[43] Sree Hari Rao, V., Raghvendra Rao, C. and Yeragani, V.K. (2006) A Novel Technique to Evaluate Fluctuations of Mood: Implications for Evaluating Course and Treatment Effects in Bipolar/Affective Disorders. *Bipolar Disorders*, **8**, 453-466. https://doi.org/10.1111/j.1399-5618.2006.00374.x

[44] Miller, C.J., Johnson, S.L. and Eisner, L. (2009) Assessment Tools for Adult Bipolar Disorder. *Clinical Psychology: Science and Practice*, **16**, 188-201. https://doi.org/10.1111/j.1468-2850.2009.01158.x

[45] Ortiz, A. and Grof, P. (2016) Electronic Monitoring of Self-Reported Mood: The Re-

turn of the Subjective? *International Journal of Bipolar Disorders*, **4**, Article No. 28. https://doi.org/10.1186/s40345-016-0069-x

[46] Lam, D.H., Jones, S.H. and Hayward, P. (2010) Cognitive Therapy for Bipolar Disorder: A Therapist's Guide to Concepts, Methods and Practice. Wiley. https://doi.org/10.1002/9780470970256

[47] Zhuparris, A., Maleki, G., van Londen, L., Koopmans, I., Aalten, V., Yocarini, I.E., *et al.* (2023) A Smartphone- and Wearable-Based Biomarker for the Estimation of Unipolar Depression Severity. *Scientific Reports*, **13**, Article No. 18844. https://doi.org/10.1038/s41598-023-46075-2

[48] Leon, E., Clarke, G., Callaghan, V. and Sepulveda, F. (2004) Real-Time Detection of Emotional Changes for Inhabited Environments. *Computers & Graphics*, **28**, 635-642. https://doi.org/10.1016/j.cag.2004.06.002

[49] van Genugten, C.R., Schuurmans, J., van Ballegooijen, W., Hoogendoorn, A.W., Smit, J.H. and Riper, H. (2021) Discovering Different Profiles in the Dynamics of Depression Based on Real-Time Monitoring of Mood: A First Exploration. *Internet Interventions*, **26**, Article 100437. https://doi.org/10.1016/j.invent.2021.100437

[50] Rutledge, R.B., Chekroud, A.M. and Huys, Q.J. (2019) Machine Learning and Big Data in Psychiatry: Toward Clinical Applications. *Current Opinion in Neurobiology*, **55**, 152-159. https://doi.org/10.1016/j.conb.2019.02.006

[51] Koppe, G., Meyer-Lindenberg, A. and Durstewitz, D. (2020) Deep Learning for Small and Big Data in Psychiatry. *Neuropsychopharmacology*, **46**, 176-190. https://doi.org/10.1038/s41386-020-0767-z

[52] Ermers, N.J., Hagoort, K. and Scheepers, F.E. (2020) The Predictive Validity of Machine Learning Models in the Classification and Treatment of Major Depressive Disorder: State of the Art and Future Directions. *Frontiers in Psychiatry*, **11**, Article ID: 472. https://doi.org/10.3389/fpsyt.2020.00472

[53] Patel, M.J., Andreescu, C., Price, J.C., Edelman, K.L., Reynolds, C.F. and Aizenstein, H.J. (2015) Machine Learning Approaches for Integrating Clinical and Imaging Features in Late-Life Depression Classification and Response Prediction. *International Journal of Geriatric Psychiatry*, **30**, 1056-1067. https://doi.org/10.1002/gps.4262

[54] Moor, M., Horn, M., Rieck, B., Roqueiro, D. and Borgwardt, K. (2019) Early Recognition of Sepsis with Gaussian Process Temporal Convolutional Networks and Dynamic Time Warping. *Machine Learning for Healthcare Conference*, 2-26.

[55] Martín-Palomeque, P., Escudero-Arnanz, Ó., Álvarez-Rodríguez, J. and Soguero-Ruiz, C. (2024) Irregular Temporal Classification of Multidrug Resistance Development in Intensive Care Unit Patients. 2024 *IEEE International Conference on Bioinformatics and Biomedicine* (*BIBM*), Lisbon, 3-6 December 2024, 5016-5023. https://doi.org/10.1109/bibm62325.2024.10822248

[56] Yang, X., Qi, X. and Zhou, X. (2023) Deep Learning Technologies for Time Series Anomaly Detection in Healthcare: A Review. *IEEE Access*, **11**, 117788-117799. https://doi.org/10.1109/access.2023.3325896

[57] Pardiñas, A.F., Owen, M.J. and Walters, J.T.R. (2021) Pharmacogenomics: A Road Ahead for Precision Medicine in Psychiatry. *Neuron*, **109**, 3914-3929. https://doi.org/10.1016/j.neuron.2021.09.011

[58] Zanardi, R., Prestifilippo, D., Fabbri, C., Colombo, C., Maron, E. and Serretti, A. (2020) Precision Psychiatry in Clinical Practice. *International Journal of Psychiatry in Clinical Practice*, **25**, 19-27. https://doi.org/10.1080/13651501.2020.1809680

[59] Gandal, M.J., Leppa, V., Won, H., Parikshak, N.N. and Geschwind, D.H. (2016) The

Road to Precision Psychiatry: Translating Genetics into Disease Mechanisms. *Nature Neuroscience*, **19**, 1397-1407. https://doi.org/10.1038/nn.4409

[60] Huys, Q.J.M., Maia, T.V. and Frank, M.J. (2016) Computational Psychiatry as a Bridge from Neuroscience to Clinical Applications. *Nature Neuroscience*, **19**, 404-413. https://doi.org/10.1038/nn.4238

[61] Bauer, M., Beaulieu, S., Dunner, D.L., Lafer, B. and Kupka, R. (2008) Rapid Cycling Bipolar Disorder—Diagnostic Concepts. *Bipolar Disorders*, **10**, 153-162. https://doi.org/10.1111/j.1399-5618.2007.00560.x

[62] Carvalho, A.F., Dimellis, D., Gonda, X., Vieta, E., McIntyre, R.S. and Fountoulakis, K.N. (2014) Rapid Cycling in Bipolar Disorder. *The Journal of Clinical Psychiatry*, **75**, e578-e586. https://doi.org/10.4088/jcp.13r08905

[63] Alarcon, R.D. (1985) Rapid Cycling Affective Disorders: A Clinical Review. *Comprehensive Psychiatry*, **26**, 522-540. https://doi.org/10.1016/0010-440x(85)90019-7

[64] Haberer, J.E., Sabin, L., Amico, K.R., Orrell, C., Galárraga, O., Tsai, A.C., *et al.* (2017) Improving Antiretroviral Therapy Adherence in Resource-Limited Settings at Scale: A Discussion of Interventions and Recommendations. *Journal of the International AIDS Society*, **20**, Article 21371. https://doi.org/10.7448/ias.20.1.21371

[65] Mellor, A., Hamill, K., Jenkins, M.M., Baucom, D.H., Norton, P.J. and Drummond, S.P.A. (2019) Partner-Assisted Cognitive Behavioural Therapy for Insomnia versus Cognitive Behavioural Therapy for Insomnia: A Randomised Controlled Trial. *Trials*, **20**, Article No. 262. https://doi.org/10.1186/s13063-019-3334-3

[66] Bolaños, L.A., Xiao, D., Ford, N.L., LeDue, J.M., Gupta, P.K., Doebeli, C., *et al.* (2021) A Three-Dimensional Virtual Mouse Generates Synthetic Training Data for Behavioral Analysis. *Nature Methods*, **18**, 378-381. https://doi.org/10.1038/s41592-021-01103-9

[67] Man, K. and Chahl, J. (2022) A Review of Synthetic Image Data and Its Use in Computer Vision. *Journal of Imaging*, **8**, Article 310. https://doi.org/10.3390/jimaging8110310

[68] Panchal, P., Kaltenboeck, A. and Harmer, C.J. (2019) Cognitive Emotional Processing across Mood Disorders. *CNS Spectrums*, **24**, 54-63. https://doi.org/10.1017/s109285291800130x

[69] Rich, B.A., Carver, F.W., Holroyd, T., Rosen, H.R., Mendoza, J.K., Cornwell, B.R., *et al.* (2011) Different Neural Pathways to Negative Affect in Youth with Pediatric Bipolar Disorder and Severe Mood Dysregulation. *Journal of Psychiatric Research*, **45**, 1283-1294. https://doi.org/10.1016/j.jpsychires.2011.04.006