Psychiatr. Pol. ONLINE FIRST Nr 375: 1–15
Published ahead of print 7 September 2025
www.psychiatriapolska.pl
ISSN 0033-2674 (PRINT), ISSN 2391-5854 (ONLINE)
DOI: https://doi.org/10.12740/PP/OnlineFirst/196927

MoodMon application for patients with affective disorders created in Poland – results of the study of the system's effectiveness based on artificial intelligence algorithms

Marlena Sokół-Szawłowska[1], Małgorzata Sochacka[2]

[1] Psychiatric Outpatient Clinic, Institute of Psychiatry and Neurology, Warsaw, Poland
[2] Britenet Med sp. z o.o., Warsaw, Poland

Summary

Aim. In a group of patients diagnosed with BD and MDD, an analysis was conducted to evaluate the efficacy of AI algorithms in detecting mental state changes based on physical voice parameters.

Material and methods. The MoodMon system was developed, including a mobile application for smartphones. In the first stage, the AI was trained using objective data and clinical assessments conducted by psychiatrists, which included 17-item versions of the HDRS (Hamilton Depression Rating Scale) and YMRS (Young Mania Rating Scale) scales, and the CGI (Clinical Global Impression) scale. The second stage was to further refine the AI using individual and population data and generate alerts when subtle changes in mental state were detected. Both stages of the study lasted a total of 944 days.

Results. Physical voice parameters can serve as biomarkers in affective disorders. The most effective in detecting changes in mental state were 19 specific physical voice parameters. The system showed high performance, with the following sensitivity (true positive rate – TPR) and specificity (true negative rate – TNR) values – for both diagnoses: TPR = 89.5%, TNR = 98.8%; BD: TPR = 89.6%, TNR = 98.9%; MDD: TPR = 89.1%, TNR = 98.5%. MoodMon helped to accurately monitor and predict changes in the mental state of patients with affective disorders.

Conclusions. The MoodMon system is an objective AI tool that effectively identifies the initial period of mental state changes in affective disorders based on physical voice parameters.

Key words: affective disorders, artificial intelligence, physical parameters of the voice

Introduction

The global prevalence of bipolar disorder (BD) is approximately 2-3%, and that of major depressive disorder (MDD) is 6.3-10.3%, which is a challenge for healthcare systems [1-4]. At the same time, patients are generally willing to monitor their mental state using mobile applications [5-7]. Individuals suffering from mental disorders actively use the Internet as often as the general population (80%), which supports the development of modern solutions in the field of mental health (e-mental health) [8-12].

The year 2010 marked the beginning of the development of the field of research on the implementation of technological solutions in affective disorders [12-14]. Several systems have been developed and tested based on the combination of two types of data: active-subjective, i.e., reported by users themselves, and passive-objective (behavioral or physiological), i.e., measured by sensors embedded in smartphones or connected devices [12]. Their main advantage lies in their high usefulness for collecting objective data through behavioral and physiological monitoring, such as: physical activity, number of telephone calls and physical parameters of the voice [12, 15-18]. These systems could potentially facilitate the diagnosis of affective states, enable early intervention in the case of prodromal symptoms and support continuous monitoring [12, 19-23].

In addition to collecting data, there is now a technological possibility to test the effectiveness of artificial intelligence (AI) algorithms in predicting changes in mental state in affective disorders [12, 19, 24-30]. The Polish contribution to the development of the use of AI algorithms in this area of psychiatry is the creation of the MoodMon system.

Aim

This article describes one of the goals of an extensive study – the assessment of the effectiveness of sending alerts following the detection of changes in mental state based on physical parameters of the voice in a group of patients with bipolar disorder and recurrent depression.

To our knowledge, we conducted the first study worldwide with the largest cohort in which the system collected, analyzed and used objective voice data from patients with bipolar disorder or recurrent depression. The data collection period was also the longest compared to other studies, exceeding 12 months.

Materials

The study was conducted from 2021 to 2023 (approved by the Bioethics Committee, 170/2021). A total of 100 patients were recruited: 75 with bipolar disorder (type I or II) and 25 with recurrent depressive disorder. Participants provided informed consent. To ensure a large number of observed changes, only patients with relatively rapid phase changes were recruited. Inclusion criteria required at least two phase changes in the previous 12 months for patients with bipolar disorder, and at least one episode in the previous 12 months for patients with recurrent depression.

After the first (observational) stage, 12 patients dropped out of the study. Eighty-eight participants remained: 66 with bipolar disorder and 22 with recurrent depressive disorder. Over a period of more than 18 months, 84 patients (BD-64; MDD-20) participated in the study. Importantly, the number of patients had no direct impact on the quantity and quality of the data collected in the study, as this depended on the number of phase changes and data delivery, especially voice samples. The analytical engine (AI) was trained on data collected from patients during the study period, which lasted 944 days (with the consent of the Bioethics Committee, Phase II of the project was extended).

Methods

The MoodMon system was developed by a team of psychiatrists and AI specialists on the basis of available research. It consists of:

– mobile application (for Android smartphones; collects and transmits data to secure servers);
– wrist pedometer (records physical activity and sleep parameters; integrated with the application);
– AI-based analytical engine (processes the collected data and learns to send alerts based on objective data from devices and clinical assessments of psychiatrists).

The MoodMon system is schematically presented in Figure 1.

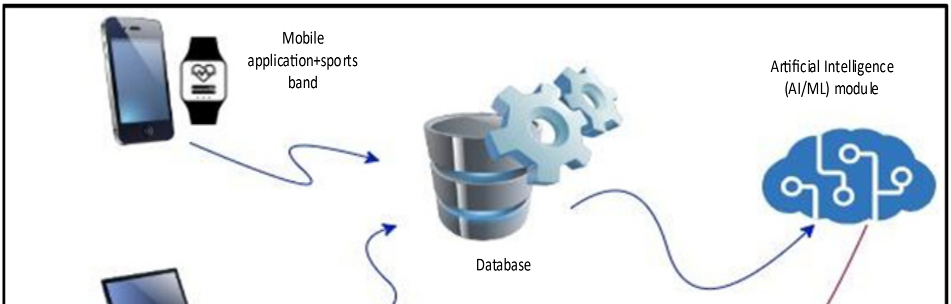

Figure 1. **Flow chart of the MoodMon system and its components**

The daily voice survey consisted of responses to three emotionally neutral questions. At the beginning of the study, each patient selected a convenient time for recording (the content was not analyzed). The medical documentation of the study was stored in an electronic Case Report Form (eCRF).

Mental status was assessed by eight psychiatrists with extensive experience in affective disorders. The voice data collected by the MoodMon system were not available to either the researchers or the patients. Data from the monitoring wristband on physical activity and sleep were available to patients and psychiatrists in graphical form. In the first stage of the study, clinical data (including initial mental state) and demographic data were collected at the initial visit. During subsequent in-person visits (planned once every three months; due to the COVID-19 pandemic, remote visits were permitted), the primary outcome measures used were the 17-item versions of the Hamilton Depression Rating Scale (HDRS) and the Young Mania Rating Scale (YMRS). The Clinical Global Impression (CGI) scale was used for quantitative assessment. The severity of symptoms did not affect the eligibility for the study, but in the case of patients with very severe symptoms (scoring 6 or 7 on the scale), a thorough clinical assessment was required to confirm that consent to participate in the study was given consciously. In this study, the HDRS and YMRS scores were considered binding in the assessment of symptom severity.

During subsequent visits, patients were asked about any significant changes in their general health and about other important and/or sudden events (which could potentially affect activity, mood, or sleep). Every two weeks, the psychiatrist also assessed each patient's condition during brief screening telephone conversations (including a CGI assessment) to determine whether there had been a significant change in mood or activity since the previous contact. To shorten the duration of the telephone conversation, patients answered questions from a predefined visit form. When the questionnaire indicated that such a change had occurred, the patient was invited for an intervention visit with a full assessment using the scales.

According to the records from the examination protocol, the psychiatrist, based on clinical experience, assigned one of seven states observed in affective disorders: mania, hypomania, balanced mood (euthymia), subdepression, depression, severe depression, or mixed state. In the data analysis system, the above assessment by the psychiatrist was recorded as a CGI value and served as a label for the behavioral data provided by the patient on the day of the assessment. In our study, the aim was to identify early/earliest and discrete symptoms of a phase change; therefore, we adopted lower cut-off points than in previous studies: euthymia – HDRS < 8 and YMRS < 6, depression – HDRS ≥ 8 and YMRS < 6, hypomania/mania – HDRS < 8 and YMRS ≥ 6, and mixed state – HDRS ≥ 8 and YMRS ≥ 6 [18, 19, 30].

Behavioral data were continuously collected via smartphones, whereas clinical assessments were much less frequent. Training of the AI predictive model requires the presence of both types of data. Therefore, psychiatric assessments were extrapolated during periods when the patients' mental state was not assessed. If there was no change in CGI between two consecutive visits, all days in that period were assigned the same CGI value as on the initial and final day. If a change in CGI was observed, data obtained between the two consecutive assessments with different CGI values were excluded from the training data set.

In the second stage, visits were scheduled at 3-month intervals, as well as in response to alerts sent to the doctor by the MoodMon system. This was followed by brief telephone visits (CGI assessment and screening questions) or, when clinically necessary, longer in-person visits (including scale assessments). The psychiatrist evaluated the validity of each alert based on the presence or absence of changes in the patient's mental state. The assessment also took into account significant life circumstances reported by the patient.

The data were preprocessed to extract the most relevant information. At the end of stage I, it was shown that physical voice parameters (including pitch, energy, and mel-frequency cepstral coefficients) were the most predictive among all collected data. Nineteen parameters with the highest predictive value were selected and used to monitor the patient's mood, providing alerts about upcoming changes during stage II of the study. In order to predict a possible change in mood and determine whether to issue an alert, acoustic and auxiliary parameters (including seasonality and gender) were analyzed by AI models both at the population level (using data from all patients) and at the individual level. The model was retrained – CGI values assigned by psychiatrists during the assessment of the alert in terms of validity were used for this purpose.

A particularly innovative aspect of the presented system is the alert mechanism implemented in the second phase of the project. This feature enables the psychiatrist to receive notifications of potential changes in the patient's mental state, facilitating preventive interventions if clinically necessary. Upon receiving an alert, and with the patient's consent, the psychiatrist can conduct a comprehensive assessment, evaluating not only the patient's current condition but also the accuracy and urgency of the alert. A notable aspect of this assessment process is *the overlay of human knowledge onto the automated system*. The psychiatrist applies their clinical knowledge and experience to subjectively assess the validity of the MoodMon alert. Based on this assessment, the psychiatrist decides whether therapeutic intervention is warranted. It should be noted that the system does not provide information on the direction of change in the mental state in the sense of "worsening" or "improvement". Additionally, the system detects changes in the mental state caused by other circumstances that may not be related to the disorder (sometimes these are trigger factors for episodes).

After assessing the validity of the alert, the system was provided with data to calculate "success" metrics (a technical term in AI, meaning "effectiveness"), in particular the true positive rate (TPR) and the true negative rate (TNR). The box below presents the general form of these calculations. A more detailed description of the machine learning and rigorous search for the most predictive acoustic voice parameters is beyond the scope of this paper. These details can be found in the technological article by our research team [10].

TPR (sensitivity, true positive rate) – an indicator of correctly predicted mood changes

$$TPR = \frac{\Sigma tp}{\Sigma tp + \Sigma fn}$$

TNR (specificity, true negative rate) – the rate of correctly predicted no changes

$$TNR = \frac{\Sigma tn}{\Sigma tn + \Sigma fp}$$

Labels for each day of all patients included in the observation were marked:

tp – true positive = alert issued and assessed as justified (mood change occurred since the patient's previous clinical assessment),

tn – true negative = no alert and its absence assessed as correct (no mood change since the patient's previous clinical assessment),

fp – false positive = alert sent but assessed as unjustified (no mood change since the patient's previous clinical assessment),

fn – false negative = no alert and decision assessed as incorrect (mood change occurred since the patient's previous clinical assessment).

Results

The average duration of the disorder in the study participants: BD – 8.6 years; MDD – 8 years. Other demographic data are shown in Table 1.

Table 1. **Demographic data**

Category	**MDD** ***N* = 25 (100%)**	**BD** ***N* = 75 (100%)**
Age range		
18-24	6 (24%)	6 (8%)
25-34	6 (24%)	18 (24%)
35-44	6 (24%)	20 (26%)
45-54	6 (24%)	16 (22%)
55-64	1 (4%)	15 (20%)
Gender		
Male	7 (28%)	33 (44%)
Female	18 (72%)	42 (56%)
Martial Status		
Divorced	2 (8%)	11 (15%)

table continued on the next page

Single	13 (52%)	35 (46%)
Married	10 (40%)	27 (36%)
Widowed	0 (0%)	2 (3%)
Education		
Higher	13 (52%)	56 (75%)
Secondary	11 (44%)	18 (24%)
Vocational	1 (4%)	1 (1%)
Professional activity		
Employed	17 (68%)	52 (69%)
Other	3 (12%)	4 (5,5%)
Retiree / disability pensioner	1 (4%)	10 (13,5%)
Student	4 (16%)	7 (9%)
Unemployed	0 (0%)	2 (3%)
Type of household		
Alone	6 (24%)	18 (24%)
With a partner	3 (12%)	13 (17%)
With family	16 (64%)	44 (59%)
Town size		
>200 000	15 (60%)	57 (76%)
5 000-200 000	5 (20%)	13 (17,5%)
<5 000	5 (20%)	5 (6,5%)

During the entire study, 1,394 changes in patients' mental states were recorded. A total of 97,428 sets of behavioral data were obtained, of which 53,934 were complete and included a clinical assessment of the patient's condition on the day of collection. Of these, 1,344 changes were observed for which both complete data and clinical assessment were available. These data sets were used to train the artificial intelligence models. During the entire observation period, 9 patients did not experience any change in their mental state. Completed data sets are also treated as complete data sets. Data completion took place when one or more parameters were missing from a set for a given day. Missing parameters were imputed using the average from the preceding 14 days, provided that no more than 7 days of data were missing during this period. If such an average could not be calculated, the entire set was rejected.

During the study, 100 initial visits, 410 in-person visits, 1,908 telephone visits, 760 intervention visits, and 100 final visits were conducted. A total of 2,656 alarms were recorded. Of these, 1,587 were assessed by psychiatrists. The number of phases recorded for patients during subsequent contacts across the two stages of the study was 4,589.

The effectiveness of the system for both indications together was: TPR = 89.5%, TNR = 98.8%. For bipolar affective disorder: TPR = 89.6%, TNR = 98.9%. For recurrent depressive disorder: TPR = 89.1%, TNR = 98.5%.

Discussion of results

Clinical studies analyzing changes in spoken language in affective disorders depending on the stage of the illness date back to 1938 [31]. Healthy individuals were also studied, though without the advanced technologies currently available. During these analyses, it was proven that a specific vocal affect correlates with characteristic patterns of voice modulation. During the experience of different emotions, breathing, phonation and articulation patterns are modified – this is reflected in the speech signal. Even basic features of speech such as tempo, loudness, intonation and rhythm can differentiate affective states. Experiencing anger or fear leads to an acceleration of speech tempo with high values of fundamental frequency ($F0$) and a wide range of intonation, due to activation of the sympathetic nervous system, rapid heart rate and increased blood pressure, which can sometimes be accompanied by dry mouth and muscle tremors. The opposite occurs in situations of boredom or sadness [32, 33]. When speech is slow and monotonous, $F0$ is lowered without major changes in intonation. This is due to stimulation of the parasympathetic nervous system, slowing of the heart rate, decreased blood pressure and increased saliva production. However, more complex patterns have also shown promise in recognizing emotional states. These include linear predictive coding (LPC) parameters and Mel-frequency cepstral coefficients (MFCC) [34].

Technological advances in recent years have contributed to the exploration of the variability of physical voice parameters during remission periods and affective disorder episodes. In the vast majority of cases, these have been short studies on small or very small clinical populations, in which a variety of methodologies have been used.

In a pilot study by Karam et al. from 2014, speech data were collected over a period of 6-12 months from six participants (four women and two men) diagnosed with bipolar I disorder and a history of rapid mood changes (four or more episodes of mania, hypomania or depression per year). Voice samples were recorded during both clinical and non-clinical interactions (a methodology for continuous and unobtrusive collection of unstructured speech by recording everyday telephone conversations was described). The study concluded that manic and depressive states could be recognized based on physical speech data (some of which were particularly predictive). However, it was noted that the system had difficulty detecting depression based on voice recordings during non-clinical interactions. In this study, the most informative features for classifying bipolar states were mean binary voiced activity, *SD* of pitch, mean segmental zero crossing rate, and mean segmental smoothed voiced activity [35].

Muaremi et al. in 2014 described the feasibility of voice analysis during telephone conversations to predict bipolar episodes. Natural telephone conversations of 12 patients over 12 weeks were analyzed. Additionally, psychiatric interviews were conducted every three weeks in the hospital. The most important speech features for predicting bipolar states were shown to be the harmonic-to-noise ratio (HNR), the number of short turns, and the pitch variance $F0$ [36].

In 2015, Guidi et al. described a case study (data from a 36-year-old male patient with bipolar disorder and using an Android smartphone – the most widely used phone worldwide). In this study, an application was designed to analyze current speech using a smartphone device. Thirty-second voice samples from the patient were compared to samples from a 25-year-old female and a 30-year-old male without psychiatric disorders. The study lasted 14 weeks. The task for the participants consisted of commenting on a picture (done 15 times at home). The patient's mood state was assessed by a psychiatrist on the day preceding each voice recording session. The application could record audio samples and estimate the fundamental speech frequency $F0$ and its variations. The study concluded that changes in voice pitch characteristics could be extracted using smartphones and that they significantly correlated with changes between balanced, depressed and hypomanic mood states in a patient with bipolar disorder [37].

In the study by Gideon et al. described in 2016, real-life speech patterns were monitored. It was assumed that, in a naturalistic setting, patients use different smartphone models, which causes acoustic changes during voice recordings. Data from 37 participants with bipolar disorder were analyzed over an average period of 29 weeks, during which 780 weekly clinical assessments were recorded. Only these structured calls were used for analysis. The article presents methods to improve the comparability of the collected data through different approaches to preprocessing, feature extraction and data modeling. The authors were able to increase the discriminative power of their algorithms for predicting the patient's current mood state. A major limitation of the study was that only structured interviews from different devices were analysed [38].

The study by Maxhuni et al. was based on the assumption that speech-related motor activity is one type of motor activity, and this is the most consistent indicator of bipolar disorder. Audio, accelerometer, and self-report data were obtained from five patients with bipolar disorder over a 12-week period during real-life activities. The study presented a system capable of classifying the state of patients with bipolar disorder using sensory information from smartphones. No previous studies have focused on the naturalistic observation of daily telephone conversations to classify impaired life functioning in people with bipolar disorder. The results indicated the system could classify the course of mood episodes or relapses in patients with high confidence. The use of audio features (emotional and spectral) resulted in >80% accuracy in classifying the mood state of patients with bipolar disorder. Testing emotional and spectral features separately and jointly resulted in similar accuracy [39].

Faurholt-Jepsen et al. in 2016 reported an accuracy ranging from 0.61 to 0.74 in classifying mood states based on physical voice features. These results were obtained by analyzing data from 28 outpatients with bipolar disorder, collected daily in natural settings over a period of 12 weeks. Voice features (extracted during daily telephone conversations), automatically generated objective smartphone data on behavioral activity, and electronic self-monitoring data were used. Voice features were found to be more accurate, sensitive, and specific in classifying manic or mixed states than in depressive states. Combining voice features with automatically generated objective smartphone data on behavioral activity and electronic self-monitoring data slightly increased the accuracy, sensitivity, and specificity of classifying affective states. Therefore, it was concluded that voice features collected in natural settings using smartphones can serve as objective state markers in patients with bipolar disorder [19].

The analyses cited above, along with a summary of the review of several other (methodologically diverse) studies in depression and bipolar disorder, indicate that there are many physical features of speech that can be identified as potential biomarkers of affective states. A preliminary consensus (based on the available fragmentary studies) suggests that changes in speech activity and voice characteristics may be potentially sensitive and accurate measures of objective prodromal symptoms in affective episodes. Moreover, the continued application of AI technology in this area offers opportunities for clinical use (which is helpful and anticipated by both patients and psychiatrists) [12, 19, 27].

The most frequently used acoustic features of speech for diagnosing depression were: prosodic features (such as pitch, fundamental frequency (*F*0), energy, speaking rate, and pauses), voice quality features (such as tremor and jitter), vocal tract descriptors (formants), and spectral features (such as MFCC and LPC) [40-49]. Recently, the most technologically advanced results of research on the physical parameters of voice in depression have been published. They use – like our study – modern analytical capabilities based on AI [50-52].

Based on even fewer studies, several potential biomarkers of mania have been identified in the form of physical features of speech. These include prosody, such as pitch, intensity (energy), speech rate and pause duration, LPC, and spectral MFCC coefficients [53, 54].

In summary, several similar solutions, but with different functionalities, were developed and tested before MoodMon. They were based data collected over short periods from small populations.

The innovative nature of MoodMon makes direct comparison to other solutions difficult, as the studies cited above collected objective "static" data indicating remission or episode, without a machine learning component to predict changes in mental state. In our project, AI was used to predict and "suggest" to the clinician discrete changes in mental state, indicating (via alerts) the need to contact the patient for clinical evaluation in case of the onset of an affective episode. After clinical evaluation, the machine learning system improved and became increasingly accurate in distinguishing changes in mental state resulting from a relapse of depression or an episode of bipolar disorder.

The MoodMon system collected a uniquely large amount of data in the form of physical voice parameters from a large population. The study is distinguished by testing AI simultaneously in recurrent depressive disorder and bipolar disorder [12, 27, 34-37]. The largest amount of behavioral data and data from visits, defined as a full data set [12, 19, 27], was analyzed. During the study, as many as 4,570 phases were captured and labeled with one of seven affective states.

During the study, AI algorithms combined numerous objective parameters with clinical assessment. Statistical analyses showed that in the set of objective data collected by the MoodMon system from mobile devices in the studied population, 19 physical voice parameters were most strongly correlated with mental state [10]. It was therefore confirmed that changes in physical voice characteristics reflect changes in mental state and can be treated as biomarkers.

A psychiatrist assessed on an ongoing basis whether the AI was accurately performing its assigned task of sending an alert when the mental state changed – these assessments were the most numerous compared to other studies. This contributed to improving the AI's. The analysis system operated on both individual and group-level data – the amount of data for training was larger than in previous studies [12, 19, 27]. The results, showing high sensitivity and specificity of the alerts for both disorders, indicate that MoodMon has great clinical potential.

More detailed descriptions differentiating the groups in terms of, among others, sociodemographic factors and users' subjective evaluation of the MoodMon application according to the type of affective disorder are beyond the scope of this study and are therefore presented in separate publications [55, 56]. A separate technological publication, providing the necessary technical explanations for AI specialists, contains details on the most predictive set of 19 physical features of the patients' voices [10].

Conclusions

1. Our study using AI confirmed that physical parameters of the voice are biomarkers in affective disorders.
2. The high sensitivity and specificity of MoodMon alerts indicates that the system can be used in the very early phase of a change in mental state.
3. The human factor is still essential and superior – the psychiatrist's role consists of assessing the clinical validity of AI-generated alerts and, based on this assessment, making therapeutic decisions.

References

1. Nierenberg AA, Agustini B, Köhler-Forsberg O, Cusin C, Katz D, Sylvia LG et al. *Diagnosis and treatment of bipolar disorder: A review.* JAMA 2023; 330(14): 1370–1380.
2. Grande I, Berk M, Birmaher B, Vieta E. *Bipolar disorder.* Lancet 2016; 387(10027): 1561–1572.
3. Baumeister H, Härter M. *Prevalence of mental disorders based on general population surveys.* Soc. Psychiatry Psychiatr. Epidemiol. 2007; 42(7): 537–546.
4. Hayes JF, Miles J, Walters K, King M, Osborn DP. *A systematic review and meta-analysis of premature mortality in bipolar affective disorder.* Acta Psychiatr. Scand. 2015; 131(6): 417–425.
5. Daus H, Kislicyn N, Heuer S, Backenstrass M. *Disease management apps and technical assistance systems for bipolar disorder: Investigating the patients' point of view.* J. Affect. Disord. 2018; 229: 351–357.

6. Van Til K, McInnis MG, Cochran A. *A comparative study of engagement in mobile and wearable health monitoring for bipolar disorder.* Bipolar Disord. 2020; 22(2): 182–190. Doi: 10.1111/bdi.12849.
7. Dickerson RF, Gorlin EI, Stankovic JA. *Empath: A continuous remote emotional health monitoring system for depressive illness.* 2011. https://www.cs.virginia.edu/~stankovic/psfiles/robempath.pdf.
8. Trefflich F, Kalckreuth S, Mergl R, Rummel-Kluge C. *Psychiatric patients' internet use corresponds to the Internet use of the general public.* Psychiatry Res. 2015; 226(1): 136–141.
9. Jędrzejewska B, Borycka A, Kotulska M, Laskus P, Lichman M, Lubczyńska Z et al. *Efficiency of mobile apps for monitoring and management of mental health – Review article.* J. Educ. Health Sport 2021; 35(1): 62–80.
10. Kamińska D, Kamińska O, Sochacka M, Sokół-Szawłowska M. *The role of selected speech signal characteristics in discriminating unipolar and bipolar disorders.* Sensors 2024; 24(14): 4721. Doi: 10.3390/s24144721.
11. Hidalgo-Mazzei D, Mateu A, Reinares M, Matic A, Vieta E, Colom F. *Internet-based psychological interventions for bipolar disorder: Review of the present and insights into the future.* J. Affect. Disord. 2015; 188: 1–13.
12. Dogan E, Sander C, Wagner X, Hegerl U, Kohls E. *Smartphone-based monitoring of objective and subjective data in affective disorders: Where are we and where are we going?* Systematic review. J. Med. Internet Res. 2017; 19(7): e262.
13. Høifødt RS, Lillevoll KR, Griffiths KM, Wilsgaard T, Eisemann M, Waterloo K et al. *The clinical effectiveness of web-based cognitive behavioral therapy with face-to-face therapist support for depressed primary care patients: Randomized controlled trial.* J. Med. Internet Res. 2013; 15(8): e153.
14. Festersen P, Corradini A. *Re: Mind: A mobile application for bipolar disorder patients.* In: *4th International Conference on Wireless Mobile Communication and Healthcare – Transforming Healthcare Through Innovations in Mobile and Wireless Technologies (MOBIHEALTH), December 5, 2014, Athens, Greece.* 2014. pp. 343–346.
15. Crisan C. *Lack of insight in bipolar disorder: The impact on treatment adherence, adverse clinical outcomes and quality of life.* In: *Psychotic disorders – An update.* Budapest: IntechOpen; 2018.
16. Faurholt-Jepsen M, Vinberg M, Frost M, Christensen EM, Bardram JE, Kessing LV. *Smartphone data as an electronic biomarker of illness activity in bipolar disorder.* Bipolar Disord. 2015; 17(7): 715–728.
17. Grünerbl A, Muaremi A, Osmani V, Bahle G, Ohler S, Tröster G et al. *Using smartphone mobility traces for the diagnosis of depressive and manic episodes in bipolar patients.* In: *Proceedings of the 5th Augmented Human International Conference, March, 2014, Kobe, Japan.* 2014. pp. 1–8.
18. Grünerbl A, Muaremi A, Osmani V, Bahle G, Ohler S, Tröster G et al. *Smartphone-based recognition of states and state changes in bipolar disorder patients.* IEEE J. Biomed. Health Inform. 2015; 19(1): 140–148.
19. Faurholt-Jepsen M, Busk J, Frost M, Vinberg M, Christensen EM, Winther O et al. *Voice analysis as an objective state marker in bipolar disorder.* Transl. Psychiatry 2016; 6(7): e856.
20. Vieta E, Salagre E, Grande I, Carvalho AF, Fernandes BS, Berk M et al. *Early intervention in bipolar disorder.* Am. J. Psychiatry 2018; 175(5): 411–426.
21. Altamura AC, Buoli M, Caldiroli A, Caron L, Cumerlato Melter C, Dobrea C et al. *Misdiagnosis, duration of untreated illness (DUI) and outcome in bipolar patients with psychotic symptoms: A naturalistic study.* J. Affect. Disord. 2015; 182: 70–75. Doi: 10.1016/j.jad.2015.04.024.

22. Menculini G, Verdolini N, Gobbicchi C, Del Bello V, Serra R, Brustenghi F et al. *Clinical and psychopathological correlates of duration of untreated illness (DUI) in affective spectrum disorders.* Eur. Neuropsychopharmacol. 2022; 61: 60–70.

23. Francis A, Gasparo P. *Interval between symptom onset and hospitalization in mania.* J. Affect. Disord. 1994; 31(3): 179–185.

24. Faurholt-Jepsen M, Vinberg M, Christensen EM, Frost M, Bardram J, Kessing LV. *Daily electronic self-monitoring of subjective and objective symptoms in bipolar disorder – The MONARCA trial protocol (MONitoring, treAtment and pRediCtion of bipolAr disorder episodes): A randomised controlled single-blind trial.* BMJ Open 2013; 3(7): e003353.

25. Gentili C, Valenza G, Nardelli M, Lanatà A, Bertschy G, Weiner L et al. *Longitudinal monitoring of heartbeat dynamics predicts mood changes in bipolar patients: A pilot study.* J. Affect. Disord. 2017; 209: 30–38.

26. Tams S, Hill K, Ortiz de Guinea A, Thatcher J, Grover V. *NeuroIS-alternative or complement to existing methods? Illustrating the holistic effects of neuroscience and self-reported data in the context of technostress research.* J. Assoc. Inf. Syst. 2014; 15(10): 723–753.

27. Flanagan O, Chan A, Roop P, Sundram F. *Using acoustic speech patterns from smartphones to investigate mood disorders: Scoping review.* JMIR Mhealth Uhealth. 2021; 9(9): e24352.

28. Dey A, Wac K, Ferreira D, Tassini K, Hong J, Ramos J. *Getting closer: An empirical investigation of the proximity of user to their smart phones.* In: *Proceedings of the 13th International Conference on Ubiquitous Computing.* New York: USA: ACM; 2011. S. A.

29. Gravenhorst F, Muaremi A, Bardram J, Grünerbl A, Mayora O, Wurzer G et al. *Mobile phones as medical devices in mental disorder treatment: An overview.* Pers. Ubiquit. Comput. 2014; 19(2): 335–353.

30. Faurholt-Jepsen M, Vinberg M, Frost M, Debel S, Christensen EM, Bardram JE et al. *Behavioral activities collected through smartphones and the association with illness activity in bipolar disorder.* Int. J. Methods Psychiatr. Res. 2016; 25(4): 309–323.

31. Newman S, Mather VG. *Analysis of spoken language of patients with affective disorders.* Am. J. Psychiatry 1938; 94: 913–942.

32. Picard RW. *Affective computing.* Cambridge, MA: MIT Press; 2000.

33. Williams CE, Stevens KN. *Emotions and speech: Some acoustical correlates.* J. Acoust. Soc. Am. 1972; 52(4): 1238–1250.

34. Nakatsu R, Nicholson J, Tosa N. *Emotion recognition and its application to computer agents with spontaneous interactive capabilities.* In: *Proceedings of the 7th ACM International Conference on Multimedia (Part 1), 1999 Oct 30–Nov 5, Orlando, FL.* New York: ACM; 1999. pp. 343–351.

35. Karam ZN, Provost EM, Singh S, Montgomery J, Archer C, Harrington G et al. *Ecologically valid long-term mood monitoring of individuals with bipolar disorder using speech.* Proc. IEEE Int. Conf. Acoust. Speech Signal Process. 2014; 2014: 4858–4862.

36. Muaremi A, Gravenhorst F, Grünerbl A, Arnrich B, Tröster G. *Assessing bipolar episodes using speech cues derived from phone calls.* In: Cipresso P, Matic A, Lopez G, eds. *Pervasive computing paradigms for mental health. MindCare 2014.* Cham: Springer; 2014. pp. 103–114.

37. Guidi A, Salvi S, Ottaviano M, Gentili C, Bertschy G, de RD et al. *Smartphone application for the analysis of prosodic features in running speech with a focus on bipolar disorders: System performance evaluation and case study.* Sensors (Basel) 2015; 15(11): 28070–2887.

38. Gideon J, Provost EM, McInnis M. *Mood state prediction from speech of varying acoustic quality for individuals with bipolar disorder.* Proc IEEE Int. Conf. Acoust. Speech Signal Process 2016; 2016: 2359–2363.

39. Maxhuni A, Muńoz-Meléndez A, Osmani V, Perez H, Mayora O, Morales EF. *Classification of bipolar disorder episodes based on analysis of voice and motor activity of patients.* Pervasive Mob. Comput. 2016; 31: 50–66.
40. Cohn JF, Kruez TS, Matthews I, Yang Y, Nguyen MH, Padilla MT et al. *Detecting depression from facial actions and vocal prosody.* In: *Proceedings of the 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, 2009 Sep 10–12, Amsterdam, Netherlands.* IEEE; 2009. pp. 1–7.
41. Yang Y, Fairbairn C, Cohn JF. *Detecting depression severity from vocal prosody.* IEEE Trans. Affect. Comput. 2013; 4(2): 142–150.
42. Alghowinem S, Goecke R, Wagner M, Epps J, Breakspear M, Parker G. *Detecting depression: A comparison between spontaneous and read speech.* In: *Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing; 2013 May 26–31, Vancouver, BC, Canada.* IEEE; 2013. pp. 7547–7551.
43. Cummins N, Vlasenko B, Sagha H, Schuller B. *Enhancing speech-based depression detection through gender dependent vowel-level formant features.* In: *Proceedings of the Conference on Artificial Intelligence in Medicine in Europe, 2017 Jun 21–24, Vienna, Austria.* Springer; 2017. pp. 209–214.
44. Moore E, Clements M, Peifer J, Weisser L. *Comparing objective feature statistics of speech for classifying clinical depression.* In: *Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2004 Sep 1–5, San Francisco, CA.* IEEE; 2004. pp. 17–20.
45. Rejaibi E, Komaty A, Meriaudeau F, Agrebi S, Othmani A. *MFCC-based recurrent neural network for automatic clinical depression recognition and assessment from speech.* Biomed. Signal Process. Control. 2022; 71(Part A): 103107.
46. Cummins N, Epps J, Breakspear M, Goecke R. *An investigation of depressed speech detection: Features and normalization.* In: *Proceedings of the 12th Annual Conference of the International Speech Communication Association, 2011 Aug 27–31, Florence, Italy.* Interspeech; 2011.
47. Taguchi T, Tachikawa H, Nemoto K, Suzuki M, Nagano T, Tachibana R et al. *Major depressive disorder discrimination using vocal acoustic features.* J. Affect. Disord. 2018; 225: 214–220.
48. Lopez-Otero P, Docio-Fernandez L, Garcia-Mateo C. *Assessing speaker independence on a speech-based depression level estimation system.* Pattern Recognit. Lett. 2015; 68(Part 2): 343–350.
49. Ringeval F, Schuller B, Valstar M, Gratch J, Cowie R, Scherer S et al. *AVEC 2017: Real-life depression and affect recognition workshop and challenge.* In: *Proceedings of the 7th Annual Workshop on Audio/Visual Emotion Challenge, 2017 Oct 23, Mountain View, CA.* ACM; 2017. pp. 3–9.
50. Weisenburger RL, Mullarkey MC, Labrada J, Labrousse D, Yang MY, MacPherson AH et al. *Conversational assessment using artificial intelligence is as clinically useful as depression scales and preferred by users.* J. Affect. Disord. 2024; 351: 489–498.
51. Zhang Y, Folarin AA, Dineley J, Conde P, Angel de V, Sun S et al. *Identifying depression-related topics in smartphone-collected free-response speech recordings using an automatic speech recognition system and a deep learning topic model.* J. Affect. Disord. 2024; 355: 40–49.
52. Ravi V, Wang J, Flint J, Alwan A. *Enhancing accuracy and privacy in speech-based depression detection through speaker disentanglement.* Comput. Speech Lang. 2024; 86: 101605.
53. Farrús M, Codina-Filbà J, Escudero J. *Acoustic and prosodic information for home monitoring of bipolar disorder.* Health Informatics J. 2021; 27(1): 1460458220972755.

54. Bedi G, Carrillo F, Cecchi GA, Slezak DF, Sigman M, Mota NB et al. *Automated analysis of free speech predicts psychosis onset in high-risk youths.* NPJ Schizophr. 2015; 1: 15030.

55. Sokół-Szawłowska M, Kamińska O, Sochacka M. *MoodMon: novel optimization of bipolar disorder monitoring through patient-driven voice parameter submission and AI technology.* Advances in Psychiatry and Neurology/Postępy Psychiatrii i Neurologii. 2024; 33(4): 230-240.

56. Sokół-Szawłowska M, Sochacka M. *MoodMon – the first Polish application using AI in bipolar disorder and recurrent depressive disorders. Patients' preliminary assessment.* Psychiatr. Pol. 2025: 1-16. https://doi.org/10.12740/PP/ OnlineFirst/207280

Corresponding author: Marlena Sokół-Szawłowska
e-mail: marlenasokolsz@gmail.com