



Identifying Physiological and Cognitive Indicators of Subthreshold Depression and Major Depressive Disorder Progression Risk

Ayaka Shimizu ¹, Satoshi Yokoyama², Alan SR Fermin¹, Yuki Mitsuyama¹, Keisuke Mizoi¹, Junichiro Yoshimoto ^{3,4}, Toshi A Furukawa⁵, Koki Takagaki⁶, Yuri Okamoto⁶, Yasumasa Okamoto¹, Go Okada¹

¹Department of Psychiatry and Neurosciences, Hiroshima University, Hiroshima, Japan; ²Faculty of Humanities, Niigata University, Niigata, Japan; ³Department of Biomedical Data Science, School of Medicine, Fujita Health University, Aichi, Japan; ⁴Division of Computational Science, International Center for Brain Science, Fujita Health University, Aichi, Japan; ⁵Office of Institutional Advancement and Communications, Kyoto University, Kyoto, Japan; ⁶Health Service Center, Hiroshima University, Hiroshima, Japan

Correspondence: Go Okada, Department of Psychiatry and Neuroscience, Hiroshima University, 1-2-3 Kasumi, Minami-ku, Hiroshima, 734-8553, Japan, Tel +81 82 257 5208, Fax +81 82 257 5209, Email gookada@hiroshima-u.ac.jp

Purpose: Subthreshold depression (StD), depressive symptoms below the diagnostic threshold for major depressive disorder (MDD), is associated with functional impairment and increased risk of MDD onset, but its physiological and cognitive characteristics remain unclear. This study aimed to identify objective markers of StD and validate predictive factors for MDD onset.

Patients and Methods: Among 168 university students, participants were classified into a StD group (n=87) and a healthy control group (n=81) using the Beck Depression Inventory-II (BDI-II). The StD group was further subdivided into “low depression” (BDI-II < 10, n = 21) and “maintained depression” (BDI-II ≥ 10, n = 66) based on their BDI-II scores at baseline. Assessments included voice characteristics, heart rate variability (HRV), peripheral body temperature, emotion recognition, and neurocognitive tasks. Between-group differences were tested using ANOVA and MANOVA, and participants were followed for MDD onset at 3-month intervals until March 2024.

Results: Compared to controls, the StD group showed increased vocal arousal (p = 0.007), elevated peripheral body temperature (p < 0.001), and reduced low-frequency HRV (p = 0.018). Sustained attention and executive function were impaired (p < 0.05). Accuracy in emotion recognition was lower for happy and surprised expressions (p = 0.030; p = 0.038). Subgroup analysis revealed more pronounced impairments in HRV, executive function, and emotion recognition in the maintained depression subgroup, which also had a higher MDD incidence (7 cases). In contrast, the low depression group showed selective attentional deficits and impaired emotion recognition, but no MDD onset.

Conclusion: StD correlates with physiological and cognitive abnormalities reflecting MDD, and persistent symptoms suggest increased progression risk. These findings present candidate markers (voice arousal, HRV, body temperature, emotion recognition) for early identification and risk stratification; however, conclusions should be interpreted cautiously given the exploratory design.

Keywords: subthreshold depression, major depressive disorder, heart rate variability, cognitive function, voice analysis, emotion recognition

Introduction

Major depressive disorder (MDD) is one of the leading causes of disability worldwide, with many cases taking a chronic course.¹ A significant proportion of suicides are related to MDD diagnosis.² Prolonged, untreated periods, worse outcomes, and increased risk of chronicity can occur when depressive symptoms go unrecognized until the individual becomes aware of their distress.³ Therefore, the importance of recognizing these symptoms before MDD onset, as well as early intervention, has been emphasized. MDD is classically viewed as a binary on-off disorder, in which the presence of a minimum of five symptoms, including depressed mood or anhedonia, occurring within the same two-week period is



needed to reach a clinical or non-clinical diagnostic criteria.⁴ This binary diagnostic model of depression is challenged by a growing acceptance of a continuous spectrum model suggesting that depressive conditions exist on a graduated scale of severity and type, ranging from mild mood fluctuations to recurrent, persistent and severe depressive episodes.^{5,6} Consequently, understanding the minimal and mild end of this depressive spectrum is critical, as it enables the timely detection of predictive symptoms that may otherwise progress undetected to more severe manifestations of MDD.

On the mild side of this spectrum is subthreshold depression (StD), a subclinical spectrum between normal health and MDD that does not meet the diagnostic criteria for MDD.^{7–9} Failing to meet the diagnostic criteria for MDD refers to having fewer than five symptoms listed in DSM-5, such as depressed mood or loss of interest, or insufficient duration of symptoms. According to a recent scoping review of StD, the lifetime prevalence rate in the general population is approximately 11.4%, and reports indicate that approximately 16–29% of adolescents and young adults meet the criteria.¹⁰ StD is associated with a high rate of suicide attempts, and patients are reported to exhibit functional impairments comparable to those of MDD.¹¹ Additionally, patients with StD have an increased risk of comorbid anxiety disorders and other psychiatric disorders.^{12,13} Studies show that about 15% of individuals with StD, including adolescents, adults, men and women, convert into MDD within a time span of one year after the initial symptoms.^{9,14–16} It has been recently suggested that a significant proportion of workplace presenteeism due to depression is attributable to StD,^{17,18} and attention is increasingly being directed toward StD-associated performance impairments.¹⁹ However, despite its significant impact on public health and workplace productivity, the biological and physiological predictors of StD remain poorly understood, underscoring an urgent need to identify early markers to prevent progression to MDD and mitigate its medical and societal burden.

MDD diagnosis typically relies on subjective self-reported questionnaires and clinician-led interviews to evaluate mood symptoms;^{20,21} however, recent research is increasingly focused on developing objective, reliable, and accessible measurement tools to enhance diagnostic precision. For example, meta-analysis results indicate that specific indicators of heart rate variability (HRV) measured using an electrocardiogram, such as High-Frequency (HF)-HRV and Low-Frequency (LF)-HRV in the Frequency domain, and the standard deviation of normal-to-normal intervals (SDNN) and the root mean square of successive differences between normal heartbeats (RMSSD) in the Time domain, are reduced in MDD.²² Voice analysis has also been extensively studied, and findings such as “the voices of patients with MDD are flat, soft, and exhibit speech delay” align with clinical observations, with reports indicating that speech rate, pitch variability, and the proportion of pauses during speech correlate with the severity of MDD.^{23,24} Recently, research has progressed in exploring the relationship between indirect indicators—such as zero-cross rate, Hurst index, and mel-frequency cepstral coefficients (MFCC) and MDD, in addition to direct indicators, like speech rate and pitch strength. These studies show that the Indicators calculated from zero-cross rate and Hurst index can distinguish between patients with moderate-to-severe depression and those with mild depression, with an AUC of 0.70.²⁵ Furthermore, the MFCC can differentiate between patients with depression and healthy individuals, with an AUC of 0.88.²⁶ Findings from other studies on physiological indicators show that higher severity of depressive symptoms is also associated with a higher self-reported body temperature during wakefulness and peripheral body temperature measured using wearable sensors.²⁷ Leveraging these physiological indicators holds significant promise for predicting depression onset and severity, offering a pathway to earlier and more targeted interventions. However, debates and challenges remain regarding the use of physiological indicators. For example, there is ongoing discussion about whether the low-frequency (LF) component of heart rate variability (HRV) reflects sympathetic or parasympathetic nervous system activity. In addition, results may differ between short-term and long-term measurements, warranting further investigation. With respect to voice features, although depression has often been associated with a monotonous tone, some studies have reported that individuals with mild depression may instead exhibit a more tense or strained voice, leading to inconsistent findings. While elevated body temperature has also been reported in depression, its specificity remains controversial, as it is strongly influenced by other factors such as stress and insomnia.

In addition to objective physiological markers, numerous studies have extensively investigated neurocognitive markers of MDD through a variety of cognitive tasks. A large-scale UK Biobank study revealed differences in reaction time and performance on the Digit Symbol Substitution Test (DSST) between affected patients and healthy controls.²⁸ A meta-analysis showed that affected patients had worse performance than healthy individuals on tasks such as sustained

attention, set shifting, and spatial working memory.²⁹ Furthermore, a systematic review indicates that during the acute phase of MDD, severe impairments are observed in processing speed, learning and memory, working memory, attention function, verbal fluency, and executive function, that some of these deficits are related to the degree of depression severity and persist even during the remission phase.³⁰ In many studies, the differences in facial emotion recognition between patients with MDD and healthy individuals have been examined as a neurocognitive indicator underlying the negative bias observed in MDD. There is a long history of research in this field, with Ekman's group beginning studies on facial emotion recognition biases in MDD as early as the 1960s.³¹ Various tasks were used in these studies; nevertheless, two broad categories have been identified: emotional recognition tasks (ERT), which require participants to identify basic emotions such as anger, disgust, fear, happiness, sadness, and surprise, and emotional bias tasks (EBT), which require participants to determine whether a presented expression is happy or sad. A meta-analysis of 23 studies comparing those diagnosed with MDD and healthy individuals with no history of psychiatric consultation revealed that those with severe MDD had lower accuracy in recognizing happy expressions than did patients with mild MDD and healthy individuals.³² Furthermore, a systematic review investigating vulnerabilities in emotion recognition among adolescents with MDD identified tendencies such as heightened sensitivity to sadness, underestimation of happiness, and overrecognition of anger.³³ These neurocognitive markers hold significant potential for identifying early cognitive vulnerabilities in MDD, paving the way for targeted preventive strategies. On the other hand, debates remain regarding the use of neurocognitive markers. Although impairments in attention and executive function are frequently reported, there is ongoing discussion as to whether these reflect trait-like or state-dependent deficits, and the potential influence of medication should also be taken into account. Regarding emotion recognition, findings indicating reduced recognition of joy appear relatively consistent, whereas results for other facial expressions vary substantially across studies, likely due to differences in the stimuli and tasks employed.

StD is influenced by genetic factors, psychological stress, social factors, and physical health status, and its course varies from spontaneous improvement over time to long-term persistence or progression to MDD.^{5,9,34–36} However, distinguishing these outcomes based solely on depressive symptoms listed in diagnostic criteria remains challenging due to their subjective nature. Thus, despite the growing recognition of StD as a precursor of MDD, little is known about whether multiple objective physiological, voice, and cognitive markers used to distinguish MDD from normal health can also identify StD within the same sample of individuals.^{6,8} Furthermore, multimodal prospective studies using objective physiological and cognitive markers for MDD, particularly those targeting young individuals, remain limited.

Building on the urgent need to address the limitations of subjective diagnostic tools and the scarcity of research on StD, this study aims to identify objective physiological and cognitive markers that characterize StD and predict its progression to MDD. Although extensive research has elucidated physiological and neurocognitive markers of MDD, such as heart rate variability, voice analysis, and facial emotion recognition deficits, studies exploring these markers in StD remain limited despite its role as a potential precursor to MDD. Given that StD represents the minimal to mild side of the depressive continuum with potential to progress into diagnosable MDD, we proposed two hypotheses. First, early disruptions in interoceptive and cognitive regulation—indexed by reduced heart rate variability, altered vocal prosody, and impaired cognitive processing—would characterize StD and precede the emergence of MDD. Specifically, we predicted that subtle but significant deviations in these physiological, vocal, and cognitive markers would serve as early indicators of risk, distinguishing individuals with StD from healthy controls and predicting later progression to MDD. Second, based on the heterogeneity of StD—namely, differing outcomes between groups with transient versus persistent symptoms—we hypothesized that distinct objective indicators would differentiate these subtypes and predict their trajectories. Prior work among university students identified groups within the StD high-symptom cohort whose symptoms either persisted/worsened or improved.³⁷ Moreover, longitudinal studies in adults have shown that approximately 12% of StD cases transition to MDD,⁶ and that persistent depressive symptoms are associated with cognitive decline and adverse brain outcomes in both older adults³⁸ and younger cohorts,³⁹ underscoring the consistent clinical significance of symptom persistence across generations. Based on these hypotheses, we examined whether markers characteristic of MDD can also be identified in StD, and further classified participants according to symptom course to determine objective indicators specific to each subtype. By comparing these markers with those of healthy controls and

assessing the incidence of MDD onset across groups, this study seeks to provide novel insights into StD's biological and cognitive underpinnings, paving the way for targeted early interventions to prevent MDD progression.

Materials and Methods

Participants and Procedure

Students enrolled at Hiroshima University in April 2021 in their first to third year of study were included. The inclusion criteria were as follows.

- 1) Age at university enrollment between 18 and 24 years.
- 2) Beck Depression Inventory-II (BDI-II) score at enrollment was ≥ 18 or < 10 .

The exclusion criteria were as follows:

- 1) In a mental state that makes it difficult to understand the purpose of the study.
- 2) Having a severe physical illness.
- 3) Diagnosis of MDD based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) in the past year.
- 4) A history of manic or hypomanic episodes according to the DSM-IV.
- 5) Current use of psychotropic medications or undergoing psychotherapy.

The participant recruitment process is illustrated in [Figure 1](#). Hiroshima University conducts health checkups for all new students, during which a mental health survey using the BDI-II is administered. Two groups were created based on BDI-II scores at enrollment (StD group: BDI-II score of ≥ 18 ; healthy group: BDI-II score of < 10). The cutoff value for the StD group was set based on a previous study,³⁷ in which the high-risk group (those with a BDI-II score of 18 or higher) was defined as having a higher incidence of MDD during the 1-year follow-up period. A BDI-II score of ≥ 18 was set as the cutoff value for the high-risk group based on a previous study.³⁷ Next, study participants were randomly selected from each group (StD group: 105 participants, healthy group: 101 participants). The selected individuals were contacted using the Email function of the Hiroshima University Student Information System, and an information session was conducted with the students who responded. Among those who provided informed consent, the semi-structured Mini-International Neuropsychiatric Interview (M.I.N.I.) was administered (StD group: 98 participants; control group: 87 participants). Trained interviewers conducted this study. The period between the BDI-II survey conducted during the freshman health checkup and the semi-structured interview varied from approximately two months to two years, depending on the participants' grade level at the initiation of the study in 2021. The semi-structured interview was conducted to confirm exclusion criteria. Among these criteria, "severe physical illness" referred to conditions that could potentially interfere with the research procedures, such as difficulty in visiting the laboratory, inability to complete questionnaires independently, or difficulty in wearing or operating the wearable device. These conditions were also set as exclusion criteria to avoid confounding, as they could directly affect depressive symptoms or physiological indicators; however, no participants in this study met these criteria. Ten participants in the StD group were excluded because they met the diagnostic criteria for MDD. The remaining participants completed the BDI-II at baseline to confirm their current depressive level and to exclude those who exhibited depression at baseline from among participants initially assigned to the healthy control group based on their BDI-II scores at enrollment. Participants who scored ≥ 10 in the control group were excluded from the study, resulting in the exclusion of four additional control participants. The evaluation was discontinued for three participants who were unable to complete the assessment because of poor health. Ultimately, 87 and 81 participants were included in the StD and control groups, respectively. The StD group was further divided into a "low depression subgroup" (BDI-II < 10 at baseline) and a "maintained depression subgroup" (BDI-II ≥ 10 at baseline), based on their BDI-II scores obtained at baseline. The low depression subgroup consisted of individuals who had depressive symptoms at enrollment but had improved to a healthy level by study onset, whereas the maintained depression subgroup consisted of those whose depressive symptoms did not improve to a healthy level and remained

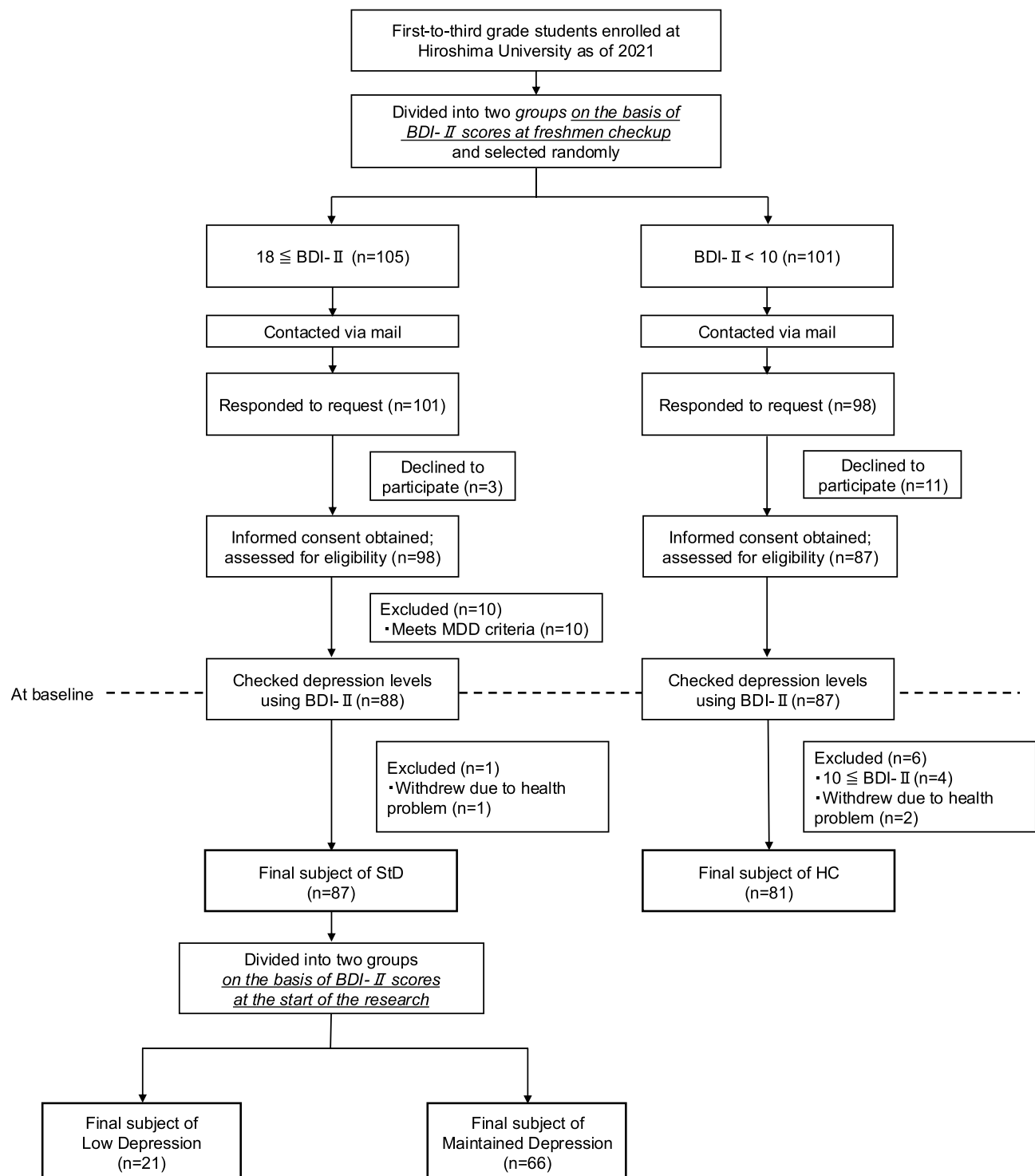


Figure 1 Participant flow.

Abbreviations: BDI-II, Beck Depression Inventory-II; MDD, major depressive disorder; StD, subthreshold depression; HC, healthy control.

persistent. As a result, 21 participants were included in the low depression subgroup and 66 in the maintained depression subgroup.

Measurements were conducted consecutively over 3 days. The first day was designated as T0, and each participant underwent measurements in the laboratory at T0 and T2. T1 involved recording and completing tasks while at home, as usual. At T0, wristband-type activity monitors (Silmee™ W22; TDK Inc., Japan) were distributed and instructed to be

worn by participants. Additionally, instructions on the 3-day testing procedures were provided, and some components of the voice measurement test, pulse measurement, emotional cognitive tasks, and cognitive function tests were conducted. At T1, participants continued their usual daily activities while wearing a wristband-type activity monitor. Measurements using this monitor continued until noon at T2. At T2, participants returned to the laboratory for additional components of the Cognitive Function Test and completed a questionnaire via a web form. The laboratory was established on the Hiroshima University campus, maintained at a comfortable room temperature, and provided an environment free from noise, vibration, and strong light.

Follow-up was conducted for participants using the BDI-II every 3 months after the study ended to track changes in depressive symptoms. The follow-up period continued until the participants graduated from the university or until the end of March 2024. If worsening depressive symptoms were observed, the MINI was administered, allowing us to identify the participants who developed MDD during the follow-up period.

Ethics Statement

Written informed consent was obtained from all the participants regarding their participation in the study and the collection of anonymous data. This study was conducted according to the principles of the Declaration of Helsinki, and all the research procedures were reviewed and approved by the Hiroshima University Ethics Committee for Epidemiological and Clinical Research (Approval Number: E2018-1513).

Data Collection

Voice Measurement

Stress is known to influence emotions, and an increase in emotions such as sadness can exacerbate depressive symptoms.⁴⁰ Research has been conducted using sensitivity technology to analyze emotions in a speaker's voice, evaluating mood disorders based on emotional components.⁴¹ The following four indicators, developed in previous studies, were also used in this study. Vivacity was calculated from the joy and sadness components of emotions in speech; relaxation was calculated from the calmness and excitement components, and vitality was calculated from Vivacity and Relaxation. The calculation methods have been previously described.⁴¹ Additionally, Shinohara et al focused on the arousal level of speech in previous studies and hypothesized that the voices of patients with MDD may have lower arousal levels. They developed the following two indicators related to speech arousal: the emotional arousal level voice index (EALVI: calculation method described in a previous study⁴²) using sound pressure change acceleration, and the arousal level voice index (ALVI: calculation method described in a previous study²⁵), focusing on the relationship between the Hurst index and zero-crossing rate of voice waveforms. In addition, we also used the number of utterances (utterances) and the time required for each utterance (duration). Participants sat in a relaxed posture on a chair while holding an iPhone SE (Apple Inc., California, USA) equipped with the MIMOSYS smartphone app (PST Co., Ltd., Yokohama, Japan) in either hand. A lavalier microphone (iRig Mic Lav, IK Multimedia Production Srl, Modena, Italy) was attached to the collar approximately 10 cm below the chin. Participants were instructed to read aloud the standard Japanese sentences displayed on the app. After reading each phrase, they pressed the button on the app to proceed to the next phrase, repeating this procedure until a total of 13 phrases had been recorded. The voice recordings were obtained using MIMOSYS and analyzed within the same system. The Japanese phrases used were selected for their familiarity, ease of pronunciation without emotion, ability to check consonants, and relevance to symptoms in the DSM-5, and were the same set used in previous studies on MIMOSYS.⁴³

HRV Measurement (2 Min)

Participants sat comfortably with their eyes closed, and their resting pulse waves were measured for 2 min using a fatigue stress meter (MF100, Murata Manufacturing Co., Ltd., Kyoto, Japan) and an app (Fatigue Monitor, Fatigue Science Laboratory Inc., Osaka, Japan). Pulse wave was measured using electrical heart rate measurements such as 12-lead electrocardiography and photoplethysmography (PPG), which are implemented in wearable devices, including wrist-watches. Reportedly, PPG can be used to accurately measure heart rate variability comparable to electrocardiography.⁴⁴ Participants sat comfortably on a chair and rested for five minutes before holding the fatigue stress meter (MF100,

Murata Manufacturing Co., Kyoto, Japan) on their knees. The MF100 used in this study is a hybrid measurement device that electrically measures heart rate via sensors attached to the palmar side of the left thumb and the first joint of the right thumb, and optically measures pulse rate via a sensor attached to the palmar side of the right thumb. Participants placed their fingers so that the skin lightly touched all three sensors. The tester launched the Fatigue Monitor app (Fatigue Science Laboratory, Osaka, Japan) on a tablet (iPad, Apple Inc., California, USA) and confirmed that the pulse waveform was displayed correctly. Participants were then instructed to close their eyes and remain relaxed while their resting pulse wave was recorded for two minutes. The following outputs were obtained: mean heart rate (mean HR); low-frequency (LF)-HRV, which reflects sympathetic and parasympathetic nerve activity in the LF component of HRV; high-frequency (HF)-HRV, which reflects parasympathetic nerve activity in the HF component of HRV; and the balance between sympathetic and parasympathetic nerves (LF/HF).

For details on heart rate variability, refer to the European Society of Cardiology Task Force and the North American Society of Pacing and Electrophysiology, 1996.⁴⁵

Emotional Cognitive Tasks

The two types of emotional tasks were performed using an application running on a Windows tablet personal computer (Surface Pro 7; Microsoft Corporation, Washington, D.C., USA). Participants sat in chairs facing desks and performed tasks by touching a tablet placed upright on the desk with their dominant index finger. The tasks were designed based on the results of previous studies. Nonetheless, Japanese facial images were obtained from the “ATR Face Expression Database” (<https://www.atr-p.com/products/face-db.html>) provided by Advanced Telecommunications Research Institute International (ATR)-Promotions Inc. to align with the participants’ living environment. The models used were adult males and females. The contrast and brightness of the facial images were adjusted using Adobe Photoshop CC (Adobe Inc., California, USA), and morphing was performed using the FantaMorph software package (Abrosoft, Beijing, China).

Emotion Recognition Task

ERT is a task where participants are presented with six basic emotional expressions (happiness, sadness, anger, disgust, fear, and surprise⁴⁶) at eight different intensity levels and are asked to select which emotional expression each corresponds to. Following previous research,⁴⁷ we created “ambiguous expressions” by synthesizing the six basic expressions in equal proportions. Subsequently, we morphed these ambiguous expressions with each emotion expression at 100% intensity, resulting in 96 expression stimulus images comprising six expressions and eight intensity levels. In each trial, a fixation point was presented on the monitor for 2000 ms, followed immediately by one of the emotion stimulus images displayed for 300 ms. This was followed by visual masking presented for 300 ms using a noise screen. After this presentation set, six emotion labels were presented as options on the screen, and participants touched the label corresponding to the emotion of the emotion stimulus image presented immediately before the presentation. Because the stimulus images were randomly presented twice for each model, the total number of trials was 192. The accuracy rate was calculated as the unbiased hit rate (Hu,⁴⁸) and used as the outcome measure.

Emotional Bias Task

The EBT task involved showing participants images of faces morphed between two emotions of varying intensities and asking them to select one of two options indicating which emotion the expression corresponded to. Following previous research,⁴⁹ we morphed two emotional expressions (happiness and sadness) for each model, creating seven stages of facial images with varying proportions of sadness and happiness. In each trial, a fixation point was presented on the monitor for 2000 ms, followed immediately by a stimulus image of either intensity for 500 ms, and then a noise screen for 500 ms to mask the visual stimulus. Following this presentation set, two emotion labels were presented on the screen as options, and the participants responded by touching the label that matched the emotion of the stimulus image presented immediately before the monitor. The most ambiguous expression (50% sadness, 50% happiness) and the expressions immediately before and after it (“40% sadness, 60% happiness” and “60% sadness, 40% happiness”) were each presented eight times, while the expressions “30% sadness, 70% happiness” and “70% sadness, 30% happiness” were each

presented six times. The most distinct expressions for each emotion (“sadness 100%” and “happiness 100%”) were presented three times each, resulting in a total of 84 trials. The subjective equivalence point (PSE) was calculated and used as an outcome measure.

Neuropsychological Cognitive Function Tests

The results of a systematic review and meta-analysis of cognitive dysfunction in MDD showed that the patients exhibit significant performance deficits in executive function, memory, and attention compared with those of healthy controls.²⁹ Regarding motor function, reports indicate functional impairments with small to moderate effect sizes.⁵⁰ Based on these reports, we selected and conducted several neuropsychological cognitive tasks associated with MDD from two existing applications equipped with cognitive function test batteries.

1. Cambridge Neuropsychological Test Automated Battery (CANTAB, Cambridge Cognition, Cambridge, UK).^{51,52}

CANTAB is a test set tool validated by over 30 years of neuroscience research and has been used in over 2500 papers. We assessed reaction time (RTI), rapid visual processing (RVP), intra/extradimensional shift task (IED), and spatial working memory (SWM) using a tablet (iPad, Apple Inc., California, USA) with CANTAB installed. Participants sat in chairs facing desks and performed the task by touching a tablet placed upright on the desk with the index finger of their dominant hand.

The RTI is used to evaluate psychomotor speed by measuring the speed at which participants react to targets (reaction speed) and how they move their fingers to touch the targets (motor speed). Five circles were displayed on the screen, and the participants touched those that turned yellow as quickly as possible. The median, mean, and standard deviation of the motor and reaction speeds were analyzed for the 5-choice RTI.

The RVP test requires participants to detect a sequence of target numbers from a series of numbers that appear randomly at a rate of 100 digits per min and evaluate sustained attention, signal detection, and impulsivity. The medians, means, and standard deviations were analyzed. A' is a signal detection indicator that reflects sensitivity to targets, independent of response trends (expected range 0.000–1.000; poor–good).

Similar to the Wisconsin Card Sorting Test, the IED is used to evaluate cognitive flexibility by assessing rule learning and reversal. In the test, stimuli consisting of filled-in shapes and white lines were displayed on the screen, and the participants used feedback to determine which stimulus was correct and devise the rules. After the six correct answers were received, the rules were changed. The adjusted total number of errors and the number of completed stages were analyzed.

SWM is a self-ordered search task in which tokens hidden in one of several boxes are sought. It is a test that is used to evaluate working memory and strategy use. Several boxes were presented on a screen in this test. The participant touched each box individually to find the tokens hidden inside it. As the token is never hidden in the same box twice in a single trial, the participant must repeatedly search while avoiding the location of the previous token. The number of boxes gradually increased to 12. Between errors, that is, the number of times the participant revisited a box where the token had previously been found, was analyzed.

2. THINC-integrated tool (THINC-it; H. Lundbeck A/S, Denmark).

THINC-it is a test set tool that can detect cognitive impairment in patients with MDD with high reliability. In this study, the Choice Reaction Time task (CRT), n-back task, DSST, and Trail Making Test were administered using a Windows tablet PC (Surface Pro 7, Microsoft Corporation, Washington D.C., USA) with THINC-it installed. Participants sat in chairs facing desks and performed the task by touching a tablet placed upright on the desk with the index finger of their dominant hand.

In the CRT, a series of arrows pointing to the left or right appear on the screen. Participants were instructed to select the direction of the arrow as quickly as possible. This test is used to evaluate attention and executive function by analyzing the number of correct responses within the time limit and the average response time.

In the n-back task, a series of symbols moves horizontally across the screen, and the first symbol disappears. The participants were required to recall and respond to missing symbols as quickly and accurately as possible. This test is used to evaluate working memory, executive function, attention, and concentration, with the average number of correct responses and response time within the time limit analyzed.

In the DSST, six symbols and their corresponding numbers are displayed at the bottom of the screen, and the participant is required to match the correct symbol to the number shown in the center of the screen as quickly as possible. This test is used to evaluate executive function, processing speed, attention, and concentration. The number of correct answers, errors, and average response time within the time limit were analyzed.

In the Trail Making Test, letters and numbers are scattered across the screen, and the participant is required to draw lines connecting the letter “a” to the number “1”, then the letter “i” to the number “2”, and so on, in order. The participants were required to connect all points as quickly as possible, ensuring that the letters and numbers were alternated. This test is used to evaluate executive function and analyze the time taken to complete a task.

Questionnaires

Questionnaires included the Japanese version of the BDI-II, validated by Kojima et al⁵³ and the Presenteeism Scale for Students (PSS⁵⁴). The BDI-II is the most widely used self-report inventory for measuring the severity of MDD. Higher scores (0–63) indicate a greater depression severity. The PSS is used to assess presenteeism among students. It calculates the Work Impairment Score (WIS) based on responses to the question “To what extent have health issues affected your usual academic (work) performance?”

Monitoring Physiological Signals with Wearable Devices

During the experiment, participants wore a commercially available wristband-type activity meter (Silmeew22, TDK Inc., Japan) on their non-dominant wrist to measure heart rate variability and peripheral body temperature. The sensor was affixed to ensure contact with the skin on the dorsal side of the wrist. They removed the device only when exposed to water and wore it at all times, including during sleep. The measurement period used for the analysis was from 9:00 PM at T0 to noon at T2.

Heart rate variability was measured by detecting the peak intervals of the pulse waves using a PPG sensor built into the device. In the time domain, the SDNN and the RMSSD were analyzed. In the frequency domain, HF, LF, and Very Low Frequency (VLF; providing information on sympathetic nervous system load) were analyzed. Periods during which the device was not in contact with the skin due to belt loosening or other causes were output as a heart rate of 0 and were removed as measurement failures. In addition, periods with a recorded heart rate below 40 beats per minute were excluded as measurement errors. These periods were regarded as non-wear time, and participants whose non-wear time exceeded 20% of the total measurement duration were excluded from the analysis. Missing data were not imputed; all missing values were handled by listwise deletion. The analysis was performed using the Kubios HRV Premium software (Kubios, Kuopio, Finland).

Peripheral body temperature was determined by measuring the wrist temperature every minute. To exclude extreme outliers, measurements in the upper and lower 5% of the distribution were removed. One-minute periods during which heart rate was not recorded were considered non-wear time, and participants whose non-wear time exceeded 20% of the total measurement duration were excluded from the analysis. Missing data were not imputed; all missing values were handled by listwise deletion.

Statistical Analysis

For demographic and clinical characteristics, age and sex distributions were assessed using chi-squared tests. BDI-II scores were compared using a one-way analysis of variance (ANOVA). For variables other than those analyzed in the ERT, multivariate ANOVA (MANOVA) was performed with sex as a covariate, comparing the StD group and healthy controls.

For the ERT, following previous studies, we conducted a two-way ANOVA with emotion (anger, disgust, fear, happiness, sadness, and surprise) and intensity as factors to compare the accuracy of emotion recognition between groups and intensity levels. We compared the overall StD group with the healthy controls, as well as each StD subgroup with the healthy controls. Following findings from previous studies, facial expression intensity was analyzed at six levels (intensities 3–8), excluding low-intensity stimulus images.⁴⁷

The above indicators were initially compared between the StD ($n = 87$) and healthy control ($n = 81$) groups. Next, focusing on the heterogeneity of StD, the group was divided into two subgroups based on BDI-II scores at the initiation of the study: “BDI-II < 10 ; low depression group ($n = 21$)” and “BDI-II ≥ 10 ; maintained depression group ($n = 66$)”. For indicators showing significant differences between the StD group and the healthy control group, comparisons between the low depression subgroup and the healthy control group, and between the maintained depression subgroup and the healthy control group, were evaluated using MANOVA/ANOVA between two groups with gender as a covariate. For pairwise group comparisons, multiple comparison adjustments were performed using the Bonferroni method implemented in SPSS.

We also performed a survival analysis (Kaplan–Meier method) to compare the proportion of participants who developed MDD during the follow-up period. We compared the results between the subgroups and the healthy control group.

All statistical analyses were performed using the IBM SPSS Statistics version 27 for Windows (SPSS Japan Inc., Tokyo, Japan).

The primary analyses in this study were conducted as two-group comparisons between each StD subtype and the healthy control group to test the hypothesis that objective indicators specific to each subtype can be identified. In addition, supplementary ANOVA/MANOVA analyses were performed among the three groups (healthy control, low depression, and maintained depression), with detailed results provided in [Supplementary Tables S3–S5](#).

Results

Overall Summary

In this study, the StD group showed significant differences compared with the healthy control group in physiological indicators (increased vocal ALVI, decreased LF ratio in continuous heart rate variability, and elevated peripheral body temperature), neurocognitive indicators (greater standard deviation of RTI movement time, lower RVP A' , fewer DSST correct responses, and longer mean reaction time), and emotional–cognitive indicators (reduced accuracy in recognizing happiness and surprise facial expressions). Furthermore, the PSS-WIS, an indicator of presenteeism, was significantly higher in the StD group.

In the subgroup analysis, the maintained depression group showed increased ALVI, decreased LF ratio, and elevated peripheral temperature among physiological indicators; reduced DSST correct responses and prolonged reaction time among neurocognitive indicators; and reduced accuracy in recognizing happiness and surprise expressions among emotion-recognition indicators. In the low depression group, physiological indicators showed elevated peripheral temperature; neurocognitive indicators showed increased standard deviation of RTI reaction time and decreased RVP A' ; and emotional–cognitive indicators showed decreased accuracy in recognizing high-intensity happiness expressions. PSS scores were elevated in both subgroups. At follow-up, the incidence of MDD was significantly higher in the maintained depression group.

Comparison Between the StD Group and the Healthy Control Group

[Table 1](#) shows the demographic and clinical characteristics of participants in the StD ($n = 87$) and control ($n = 81$) groups. Significant differences were observed in the following indicators when both groups were compared: ALVI in Voice Measurement, SD of movement time in CANTAB’s RTI, A' in RVP, correct number and mean response time in the DSST of THINC-it, WIS in the PSS, LF (%) in continuous heart rate variability measurements, peripheral body temperature, and unbiased hit rate (Hu) for happy and surprised expressions in the ERT ([Tables 1 and 2](#), [Figure 2](#)).

ALVI was significantly higher in the StD group than in the healthy control group [$F(1, 165) = 7.45, p = 0.007$]. The SD of the movement time in the RTI group was significantly higher in the StD group than in the healthy control group [$F(1, 165) = 5.32, p = 0.022$]. A' in RVP was significantly lower in the StD group than in the control group [$F(1, 165) = 5.84, p = 0.017$]. The number of correct responses on the DSST was significantly lower in the StD group than in the control group [$F(1, 165) = 4.36, p = 0.038$], and the mean response time was significantly longer in the StD group than in

Table 1 Participants' characteristics of healthy control group and subthreshold depression group

	Healthy Controls n=81		Subthreshold Depression n=87				
	Mean	(SD)	Mean	(SD)	95% CI	p-value	Cohen's d
Sex (Male/Female)	33/48		50/37			0.032	
Age	19.5	0.9	19.6	1.0		0.213	
BDI-II	3.06	2.56	14.77	8.24	[-13.604, -9.809]	<0.001	-1.89
Voice							
utterance duration (ms)	13.01	0.19	13.02	0.21	[-0.076, 0.05]	0.688	-0.05
vitality	27488.01	10286.42	27261.46	8278.33	[-2511.732, 3251.221]	0.800	0.02
vivacity	0.26	0.07	0.26	0.09	[-0.015, 0.033]	0.452	0.05
relaxation	0.22	0.13	0.19	0.12	[-0.029, 0.038]	0.783	0.19
EALVI	0.63	0.15	0.66	0.15	[-0.004, 0.035]	0.126	-0.20
ALVI	0.40	0.03	0.39	0.03	[-0.001, 0.018]	0.081	0.33
	0.38	0.03	0.39	0.04	[-0.025, -0.004]	0.007	-0.27
HRV (2 min)							
Mean HR	80.22	13.10	81.83	11.80	[-16.662, 58.98]	0.338	-0.13
LF (ms ²)	600.10	514.06	613.39	658.83	[-190.997, 260.064]	0.940	-0.02
HF (ms ²)	660.01	680.19	551.86	522.48	[-109.471, 176.956]	0.422	0.18
LF/HF ratio	1.80	2.63	1.67	1.73	[-0.443, 0.911]	0.495	0.06
RTI							
Median movement time (ms)	214.85	49.43	221.74	51.40	[-22.688, 8.612]	0.376	-0.14
Median reaction time (ms)	329.83	23.17	330.14	28.28	[-8.849, 7.216]	0.841	-0.01
Mean movement time (ms)	217.76	48.04	223.69	49.62	[-21.091, 9.221]	0.441	-0.12
Mean reaction time (ms)	335.31	23.72	334.72	30.35	[-8.698, 8.197]	0.953	0.02
SD of movement time (ms)	37.42	14.71	44.66	21.49	[-12.411, -0.96]	0.022	-0.39
SD of reaction time (ms)	36.00	12.41	36.15	15.78	[-5.146, 3.626]	0.733	-0.01
RVP							
A'	0.96	0.03	0.95	0.04	[0.002, 0.023]	0.017	0.28
Median response latency (ms)	415.85	68.29	425.56	66.25	[-31.62, 10.043]	0.308	-0.14
Mean response latency (ms)	455.87	75.17	461.62	75.19	[-32.627, 13.474]	0.413	-0.08
SD of response latency (ms)	146.56	75.34	135.11	64.74	[-14.988, 27.375]	0.565	0.16
IED							
Total errors adjusted	16.15	18.99	15.59	22.63	[-6.232, 6.755]	0.937	0.03
Stages completed	8.77	0.78	8.82	0.91	[-0.297, 0.228]	0.796	-0.06
SWM							
Between errors 4 boxes	0.25	0.80	0.26	0.77	[-0.24, 0.247]	0.976	-0.01
Between errors 6 boxes	1.19	2.08	1.08	2.20	[-0.572, 0.761]	0.780	0.05
Between errors 8 boxes	3.08	4.57	3.51	4.99	[-2.199, 0.725]	0.321	-0.09
Between errors 12 boxes	14.61	13.01	15.38	13.97	[-5.951, 2.181]	0.361	-0.06
CRT							
Correct	39.16	1.23	38.93	1.39	[-0.135, 0.677]	0.189	0.18
Mean response time (ms)	640.16	110.46	640.93	112.88	[-33.95, 35.297]	0.969	-0.01

(Continued)

Table 1 (Continued).

	Healthy Controls n=81		Subthreshold Depression n=87				
	Mean	(SD)	Mean	(SD)	95% CI	p-value	Cohen's d
N-back task							
Correct	35.37	6.03	36.18	3.46	[-1.918, 1.025]	0.550	-0.17
Mean response time (ms)	864.07	153.94	833.02	139.88	[-20.726, 69.524]	0.287	0.21
DSST							
Correct	83.99	14.20	79.16	14.45	[0.253, 9.141]	0.038	0.34
Errors	4.06	3.59	4.11	3.33	[-1.03, 1.106]	0.944	-0.01
Mean response time (ms)	1455.94	237.76	1551.31	261.01	[-170.224, -15.181]	0.019	-0.38
TMT							
Timesum (ms)	15831.28	4033.24	16709.14	4481.16	[-2246.637, 401.462]	0.171	-0.21
EBT							
PSE	44.41	5.13	44.91	5.72	[-1.69, 1.584]	0.949	-0.09
PSS							
Work impairment score	7.54	9.07	19.48	13.61	[-15.964, -8.785]	<0.001	-1.03

Notes: Bold values indicate statistically significant differences at $p < 0.05$.

Abbreviations: SD, standard deviation; CI, confidence interval; BDI-II, Beck Depression Inventory-II; EALVI, emotional arousal level voice index; ALVI, arousal level voice index; HRV, heart rate variability; HR, heart rate; HF, high frequency; LF, low frequency; RTI, reaction time; ms, milliseconds; RVP, rapid visual processing; IED, intra/extradimensional shift task; SWM, spatial working memory; CRT, choice reaction time task; DSST, digit symbol substitution test; TMT, trail making test; EBT, emotional bias task; PSE, subjective equivalence point; PSS, presenteeism scale for students.

Table 2 Physiological characteristics of healthy control group and subthreshold depression group

	Healthy Controls n=68		Subthreshold Depression n=72				
	Mean	(SD)	Mean	(SD)	95% CI	p-value	Cohen's d
HRV (Continuous measurement)							
Time Domain							
Mean HR	64.77	5.57	65.72	7.00	[-3.605, 0.647]	0.171	-0.15
SDNN (ms)	60.75	10.34	59.45	10.87	[-0.001, 0.006]	0.139	0.12
RMSSD (ms)	70.48	12.06	69.51	12.78	[-0.001, 0.007]	0.210	0.08
Frequency Domain							
VLF(%)	14.87	1.37	15.23	2.09	[-1.132, 0.047]	0.071	-0.20
LF(%)	46.28	2.01	45.51	1.56	[0.129, 1.354]	0.018	0.43
HF(%)	38.78	2.74	39.18	2.82	[-1.132, 0.741]	0.680	-0.14
LF/HF ratio	1.20	0.13	1.17	0.11	[-0.014, 0.066]	0.203	0.28
Skin Temperature							
Skin Temperature (°C)	31.82	1.09	32.98	0.68	[-1.41, -0.798]	<0.001	-1.28

Notes: Bold values indicate statistically significant differences at $p < 0.05$.

Abbreviations: SD, standard deviation; CI, confidence interval; HRV, heart rate variability; HR, heart rate; ms, milliseconds; SDNN, standard deviation of NN intervals; RMSSD, root mean square of successive differences between normal heartbeats; VLF, very low frequency; HF, high frequency; LF, low frequency.

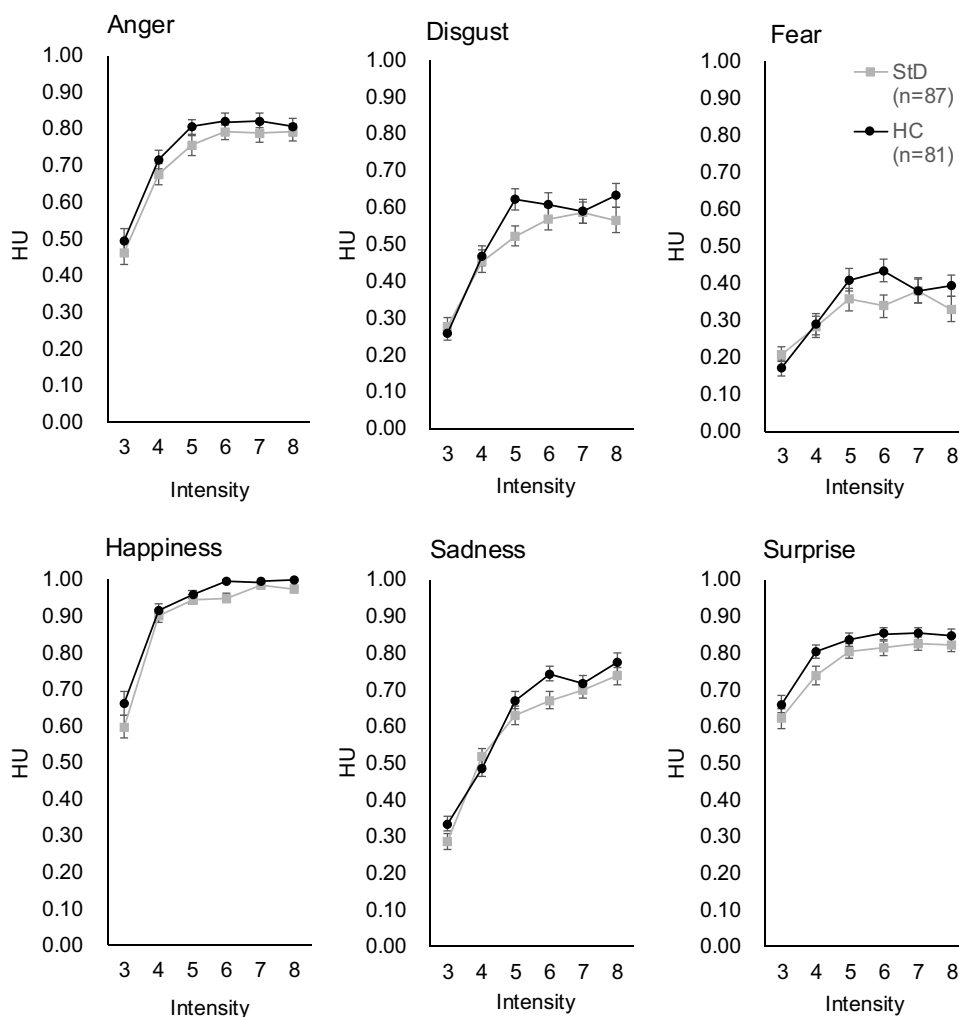


Figure 2 Unbiased hit rate for each emotion at each intensity level for each group.

Notes: Error bars represent standard error.

Abbreviations: HU, unbiased hit rate; StD, subthreshold depression; HC, healthy control.

the control group [$F(1, 165) = 5.58, p = 0.019$]. The PSS-WIS of the PSS was significantly higher in the StD group than in the control group [$F(1, 165) = 46.33, p < 0.01$] (Table 1).

The LF (%) in the continuous heart rate variability measurements was significantly lower in the StD group than in the healthy group [$F(1, 137) = 5.73, p = 0.018$]. The peripheral body temperature at the wrist was significantly higher in the StD group than in the control group [$F(1, 137) = 50.92, p < 0.001$] (Table 2).

The results of ERT are shown in a graph of HU values for expression intensities of 3–8 for each emotion (Figure 2). Regarding the main effects by group, no significant differences were found for anger [$F(1, 165) = 0.70, p = 0.403$], disgust [$F(1, 165) = 0.76, p = 0.384$], fear [$F(1, 165) = 0.47, p = 0.493$] and sadness [$F(1, 165) = 1.65, p = 0.200$]. A main group effect was observed for happiness [$F(1, 165) = 4.77, p = 0.030$] and surprise [$F(1, 165) = 4.40, p = 0.038$], with the StD group showing lower accuracy in all cases. No interaction effects between groups or intensities were observed for any emotion. Detailed ANOVA results for the ERT comparing the StD and healthy control groups are provided in [Supplementary Table S1](#).

Comparison Between StD Subgroups and the Control Group

The StD group was divided into two subgroups: the “low depression subgroup (BDI-II < 10; n=21)” and the “maintained depression subgroup (BDI-II ≥ 10; n=66)”. The demographic and clinical characteristics of the participants are presented in [Table 3](#). Indicators showing significant differences between the StD (n = 87) and healthy control (n = 81) groups were compared between each subgroup and the healthy control group using two-group ANOVA ([Tables 3 and 4](#), [Figure 3](#)).

ALVI was significantly different between the maintained depression and healthy control groups [$F(1, 144) = 5.37$, $p = 0.022$]. The SD of the RTI movement time differed significantly between the low depression and healthy control groups [$F(1, 99) = 6.07$, $p = 0.015$]. The A' in RVP significantly differed between the low depression and healthy control groups [$F(1, 99) = 7.68$, $p = 0.007$]. The number of correct responses in the DSST [$F(1, 144) = 5.91$, $p = 0.016$] and mean response time [$F(1, 144) = 7.95$, $p = 0.005$] significantly differed between the maintained depression and healthy control groups. PSS scores also significantly differed between the two subgroups and the healthy control group, with the maintained depression group [$F(1, 144) = 35.61$, $p < 0.001$] having a higher score than the low depression group [$F(1, 99) = 34.92$, $p < 0.001$] ([Table 3](#)).

The LF (%) in continuous heart rate variability measurements showed a significant difference between the maintained depression and control groups [$F(1, 120) = 6.27$, $p = 0.014$]. The peripheral body temperature was significantly different from that of the healthy group in both subgroups: the low depression group [$F(1, 82) = 19.10$, $p < 0.001$] and the maintained depression group [$F(1, 120) = 36.80$, $p < 0.001$] ([Table 4](#)).

Regarding the ERT results, a group main effect was observed for happiness and surprise when comparing the StD group (n = 87) and the healthy control group (n = 81). Subgroup comparisons were performed with healthy controls. Happiness expression showed a main effect of group in the maintained depression group [$F(1, 144) = 5.28$, $p = 0.023$], with both groups having lower accuracy than the control group. In the low depression group, the expression of intensities 6, 7, and 8 had lower accuracy than in the control group, intensity 6 [$F(1, 99) = 17.42$, $p < 0.001$], intensity 7 [$F(1, 99) = 6.94$, $p = 0.010$], and intensity 8 [$F(1, 99) = 5.59$, $p = 0.020$]. Surprised expressions showed a main effect of group in the maintained depression group; however, lower accuracies were found in this group [$F(1144) = 4.91$, $p = 0.028$] ([Figure 3](#)).

Finally, the incidence of MDD in each subgroup was compared to that in the control group. From the end of the measurement period until March 2024, no participants in the healthy control or low depression groups developed MDD. However, seven participants in the maintained depression group did. The number of onset cases of MDD was zero in both the healthy control group and the low depression group, so a statistical comparison between groups could not be performed. Survival analysis comparing the incidence rates of MDD between the healthy control and maintained depression groups revealed a significant difference ($P = 0.002$) ([Figure 4](#)). Detailed ANOVA results for the subgroup comparisons between the healthy control group and each StD subtype are presented in [Supplementary Table S2](#).

The results of the supplementary three-group ANOVA/MANOVA analyses are presented in [Supplementary Tables S3–S5](#). Although these analyses revealed some differences, such as certain indicators showing a shift from significance to a trend toward significance, the main conclusion of this study—namely, the primary differences between the healthy control group and the StD subtypes—was largely maintained.

Discussion

In this study, our primary goal was to identify objective physiological and cognitive markers characterizing StD and to explore their potential in predicting its progression into MDD. This approach addresses the critical gap in early identification of individuals at risk of developing MDD, highlighted by the limitations of subjective diagnostic tools. We hypothesized that markers previously associated with MDD—such as abnormalities in HRV, voice features, and neurocognitive performance—would also be present in StD, reflecting a continuum between these conditions, and that specific markers could differentiate StD subtypes based on symptom persistence. Our findings broadly support this hypothesis, revealing significant differences between the StD and healthy control groups across multiple objective indicators, including voice analysis (ALVI), cognitive task performance (RTI, RVP, DSST), heart rate variability (LF %), peripheral body temperature, and emotional recognition accuracy (ERT for happiness and surprise). Furthermore,

Table 3 Participants' characteristics of healthy control group and subgroups of subthreshold depression

	Healthy Controls n=81		Low Depression Group n=21					Maintained Depression Group n=66				
	Mean	(SD)	Mean	(SD)	95% CI	p-value	Cohen's d	Mean	(SD)	95% CI	p-value	Cohen's d
Sex (Male/Female)	33/48		9/12			1.000		41/25			0.013	
Age	19.5	0.9	19.6	0.8		0.623		19.7	1.1		0.673	
BDI-II	3.06	2.56	4.90	2.36	[-3.075, -0.61]	0.004	-0.73	17.90	6.83	[-16.455, -13.122]	<0.001	-3.00
Voice												
ALVI	0.38	0.03	0.39	0.04	[-0.022, 0.009]	0.406	-0.19	0.39	0.04	[-0.024, -0.002]	0.022	-0.34
RTI												
SD of movement time (ms)	37.42	14.71	48.55	28.64	[-20.102, -2.168]	0.015	-0.61	43.42	18.76	[-10.432, 0.645]	0.083	-0.36
RVP												
A'	0.96	0.03	0.94	0.03	[0.005, 0.032]	0.007	0.65	0.95	0.04	[0, 0.022]	0.056	0.26
DSST												
Correct	83.99	14.20	82.19	13.98	[-5.126, 8.705]	0.609	0.13	78.20	14.57	[1.11, 10.78]	0.016	0.40
Mean response time (ms)	1455.94	237.76	1496.81	301.71	[-163.381, 82.451]	0.515	-0.16	1568.66	246.70	[-197.398, -34.693]	0.005	-0.47
PSS												
Work impairment score	7.54	9.07	22.00	13.43	[-19.395, -9.644]	<0.001	-1.43	18.68	13.67	[-15.33, -7.701]	<0.001	-0.98

Notes: Bold values indicate statistically significant differences at $p < 0.05$.

Abbreviations: SD, standard deviation; CI, confidence interval; BDI-II, Beck Depression Inventory-II; ALVI, arousal level voice index; HRV, heart rate variability; RTI, reaction time; ms, milliseconds; RVP, rapid visual processing; DSST, digit symbol substitution test; PSS, presenteeism scale for students.

Table 4 Physiological characteristics of healthy control group and subgroups of subthreshold depression

	Healthy Controls n=68		Low Depression Group n=17					Maintained Depression Group n=55				
	Mean	(SD)	Mean	(SD)	95% CI	p-value	Cohen's d	Mean	(SD)	95% CI	p-value	Cohen's d
HRV (Continuous measurement)												
Frequency Domain												
LF(%)	46.28	2.01	45.81	2.25	[-0.67, 1.564]	0.428	0.23	45.42	1.28	[0.169, 1.445]	0.014	0.50
Skin Temperature												
Skin Temperature (°C)	31.82	1.09	33.04	0.55	[-1.745, -0.653]	<0.001	-1.20	32.96	0.71	[-1.398, -0.71]	<0.001	-1.21

Notes: Bold values indicate statistically significant differences at $p < 0.05$.

Abbreviations: SD, standard deviation; CI, confidence interval; HRV, heart rate variability; LF, low frequency.

subgroup analyses demonstrated distinct marker profiles for transient (low-depression) and persistent (maintained-depression) StD, with the latter subgroup showing a significantly higher incidence of MDD onset. These results underscore the potential of these markers for early risk stratification and intervention.

In individuals with StD, sustained attention—as indicated by lower A' scores in the Rapid Visual Information Processing (RVP) task—was impaired, mirroring deficits observed in MDD and supporting the notion of a shared cognitive continuum. Similarly, performance on the Digit Symbol Substitution Test (DSST) revealed fewer correct responses and longer average response times in the StD group compared to healthy controls, indicating deficits in executive functions, including motor speed, attention, planning, strategizing (as involved in learning symbol-number pairs), and working memory (required to maintain task rules). These findings suggest that specific cognitive impairments characteristic of MDD are also evident in StD, underscoring their potential as early markers of risk.⁵⁵ In contrast, tasks

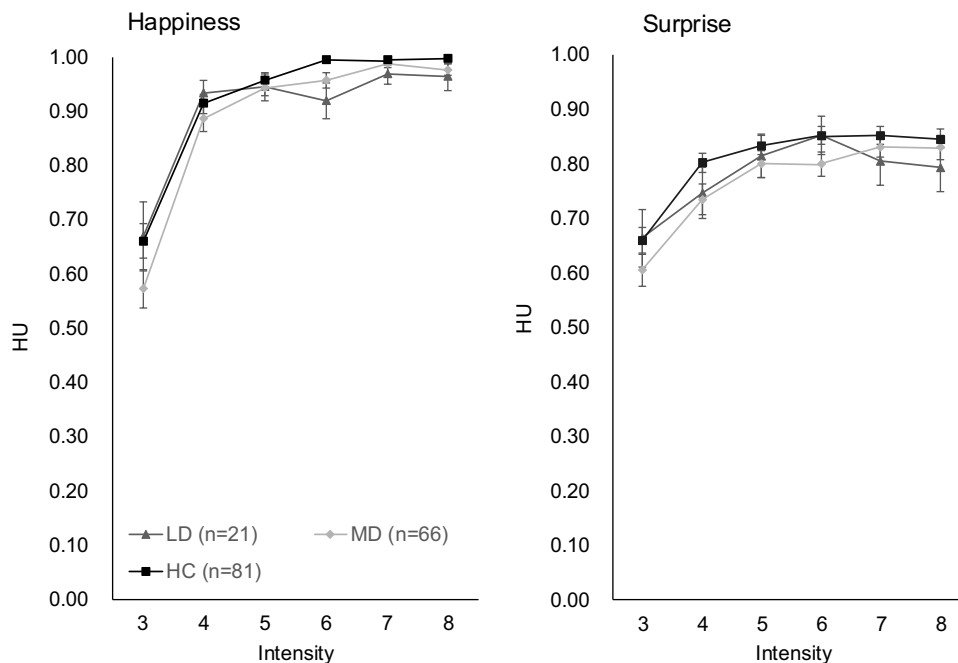


Figure 3 Unbiased hit rate for happiness and Surprise emotions at each intensity level for subgroups of StD and HC.

Notes: Error bars represent standard error.

Abbreviations: HU, unbiased hit rate; LD, low depression; MD, maintained depression; HC, healthy control.

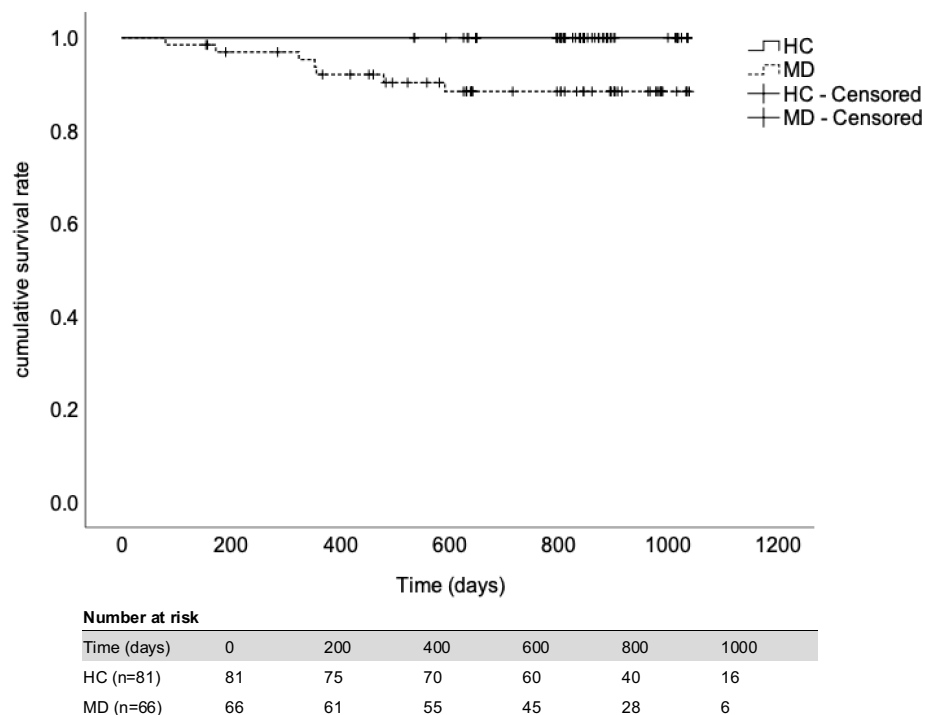


Figure 4 Survival analysis comparing the incidence rates of MDD between the healthy control group and the maintained depression group. **Abbreviations:** MDD, major depressive disorder; HC, healthy control; MD, maintained depression.

assessing motor and reaction times (Reaction Time Task, RTI), set-shifting, spatial working memory, N-back, and the Trail Making Test (TMT) showed no abnormalities in StD, despite reported deficits in MDD.⁵⁵ This highlights that certain cognitive functions remain preserved in StD as a precursor state. Collectively, these results elucidate selective cognitive vulnerabilities in StD, offering critical insights for identifying at-risk individuals and informing targeted interventions to prevent progression to MDD.

Elevated peripheral body temperature, measured at the wrist using a wearable device, was observed in the StD group, consistent with patterns reported in MDD. Several mechanisms may underlie body temperature elevation in depression, including imbalances in sympathetic and parasympathetic activity contributing to impaired heat dissipation due to reduced sweating capacity, disrupted thermoregulation from neuroimmune interactions, chronic inflammation, stress, HPA axis dysregulation, and abnormalities in temperature-sensitive channels.^{27,56,57} For example, a recent study identified higher peripheral body temperature in individuals with mild depressive symptoms but no clinical diagnosis of MDD compared to those without depressive symptoms. This finding aligns with our results and suggests that thermoregulatory abnormalities may emerge early in the StD-MDD continuum.

Regarding HRV, our study found no significant difference in HF between the StD and control groups; however, LF was significantly reduced in the StD group. Although HF—an indicator of parasympathetic nervous activity—is known to decrease in MDD,²² our findings suggest that this alteration may not yet be pronounced in StD.⁵⁸ LF, traditionally linked to sympathetic nervous activity, is now understood to reflect a broader balance between autonomic sympathetic and parasympathetic functions or neural responses to postural changes rather than purely sympathetic activity.^{59,60} Recent studies reported similar LF reductions in StD without significant HF differences, indicating that impaired baroreceptor function—sensors in the carotid sinus and aortic arch that regulate heart rate and vascular tone via autonomic reflexes—may contribute to these changes. This is supported by meta-analyses showing reduced LF in adults with MDD—particularly in younger (30–40 years) and older populations—suggesting a consistent autonomic dysregulation across the depressive spectrum.^{22,61} Notably, our study detected LF reductions through continuous HRV measurements using wearable devices, but not in 2-min pulse wave assessments conducted before

cognitive tasks, possibly due to anticipatory stress from upcoming tasks influencing short-term measurements. These findings emphasize the value of continuous HRV monitoring as a sensitive, objective tool for detecting autonomic changes in StD, offering insights into its potential as an early marker for MDD risk and highlighting the importance of measurement context in psychophysiological research.

Assessment of emotional recognition with the ERT revealed reduced accuracy in identifying happy and surprised expressions in the StD group, suggesting early deficits in processing positive emotions that align with patterns observed in MDD. Previous meta-analyses have reported that individuals with MDD exhibit lower recognition accuracy for all facial expressions compared to healthy controls, with particularly pronounced deficits for happy expressions.³²

In line with the present findings, a previous study found that MDD patients receiving a placebo showed significantly poorer recognition of happy and surprised—often categorized as positive or non-negative emotions alongside anger, disgust, fear, and sadness—compared to healthy controls.^{62–64} Our findings in StD mirror these deficits in positive emotion recognition, supporting the hypothesis that cognitive biases emerge early in the StD-MDD continuum. In contrast, emotional bias assessed with the EBT showed no evidence of a negative bias in StD, unlike in MDD, where patients frequently misinterpret ambiguous expressions as sad, with this bias diminishing as symptoms improve.⁶⁵ The absence of a detectable negative bias in our StD cohort likely reflects their subthreshold depressive symptoms, highlighting a key distinction between StD and MDD and underscoring the potential of ERT as an objective marker for early identification of at-risk individuals.

The WIS scores on the PSS were significantly higher in the StD group, indicating that presenteeism, a hallmark of workplace impairment, is evident in this precursor state to MDD. Previous studies have established a strong association between MDD severity and presenteeism, with workplace productivity losses linked to depressive symptoms.¹⁸ Research targeting student populations further supports this, showing that students with emotional problems, including those exhibiting depressive mood and anxiety symptoms, exhibit significantly higher WIS scores than their peers.⁵⁴ Our findings extend these observations to StD, demonstrating that presenteeism emerges before the onset of clinical MDD, consistent with prior reports suggesting that workplace impairments originate in StD.^{17,18} This highlights the functional significance of StD and reinforces the need for early detection of its cognitive and physiological markers to mitigate its societal and economic impacts through targeted interventions.

Notably, some indicators in StD exhibited patterns distinct from those typically observed in MDD, particularly in voice arousal, as measured by the ALVI, which was higher in the StD group compared to healthy controls. In contrast, previous studies have reported lower ALVI in MDD, with more severe depressive symptoms associated with reduced vocal intensity.²⁵ It has been suggested that elevated speech arousal in the early stages of MDD may reflect heightened stress levels, a finding supported by comparisons showing higher voice arousal in MDD patients with milder symptoms.⁴² Our results align with this, suggesting that StD, as a precursor to MDD, may be characterized by a high-stress state that leads to increased ALVI compared to healthy controls. However, the validity and reliability of ALVI as a marker of vocal arousal in MDD remain limited. Likewise, although EALVI has been proposed as a marker of vocal arousal, it only showed a marginal trend in the present study. Therefore, it may be premature to draw firm conclusions about the clinical significance of these markers, and further studies need to investigate this issue.

Although no reports have specifically reported on the SD of movement time in the RTI in MDD, our findings revealed a significant increase in RTI movement time variability in the StD group, suggesting inattention or inconsistent motor performance as an early cognitive marker. A meta-analysis confirmed impaired attention in MDD, supporting the notion that attentional deficits may emerge in StD.²⁹ However, simple motor speed in the RTI was not impaired in StD in this study, consistent with a meta-analysis indicating that motor speed deficits in MDD have a smaller effect size compared to attention.⁵⁰ These findings highlight selective cognitive vulnerabilities in StD, distinguishing it from MDD and underscoring the potential of RTI variability as a novel marker for identifying individuals at risk.

The StD group was further divided into transient (low depression, BDI-II < 10) and persistent (maintained depression, BDI-II ≥ 10) subgroups to explore heterogeneity in marker profiles and MDD risk. In the low depression group, increased RTI movement time variability and lower A' scores in the Rapid Visual Information Processing (RVP) task indicated deficits in sustained attention, suggesting that attentional impairments may predispose individuals to reactive depression rather than directly causing it. Conversely, the maintained depression group exhibited poorer performance on the DSST, reflecting deficits in executive function such as attention, planning, and working memory. Previous research comparing

“endogenous” and “neurotic” MDD subgroups significantly lower DSST performance in the more severe endogenous group, aligning with our observation that persistent depressive symptoms in StD are associated with pronounced executive dysfunction.⁶⁶

Physiologically, low-frequency (LF) heart rate variability (HRV) was significantly reduced in the maintained depression group, consistent with baroreceptor dysfunction observed in MDD, but not in the low depression group, indicating that autonomic dysregulation may be specific to persistent StD. In contrast, elevated peripheral body temperature was observed in both subgroups, suggesting that thermoregulatory abnormalities are a robust feature of StD, potentially driven by neuroimmune or stress-related mechanisms, though further research is needed to elucidate these pathways.

Facial expression recognition also varied by subgroup: the maintained depression group showed reduced accuracy for happy and surprised expressions across all intensities, while the low depression group exhibited deficits only for more straightforward happy expressions. This persistent difficulty in recognizing positive emotions, even after symptom improvement in the low depression group, suggests underlying cognitive vulnerability that may predispose individuals to reactive depression. However, the recovery of ambiguous expression recognition to normal levels in the low depression group requires further investigation to clarify its mechanisms.

The maintained depression group demonstrated a significantly higher incidence of MDD onset compared to the low depression and healthy control groups, with no MDD cases observed in the latter two during follow-up, consistent with prior studies linking higher depressive symptom severity to increased MDD risk.³⁷ Maintained depression is a subtype characterized by the long-term persistence of depressive symptoms rather than transient episodes, leading not only to functional impairment but also to a substantially increased risk of progression to MDD. This persistent course is thought to reflect physiological and cognitive characteristics that are continuous with MDD, such as autonomic dysregulation and cognitive impairment. These vulnerabilities may interact with stressful environments and cognitive biases, resulting in symptom chronicity and, ultimately, the development of MDD that meets diagnostic criteria.

The most clinically significant finding of this study is that the prospective design allowed us to track the risk of MDD onset. This supports the notion that StD is not merely mild depression but a continuum encompassing heterogeneous subtypes, indicating it could be a target for clinical intervention.

Limitations in our study design— This study was exploratory and based on a relatively small sample; therefore, caution is warranted when interpreting results with small effect sizes. Overall, effect sizes were generally small, although some subgroup comparisons showed moderate effects. This pattern suggests early or subtle alterations in StD that warrant replication in larger samples before drawing firm clinical conclusions. In addition, several limitations should be noted. First, regarding methodological constraints, participants were instructed to wear a wristband-type activity monitor from 9:00 PM at T0 until noon at T2. However, factors such as forgetting to wear the device or inadequate sensor contact with the skin may have affected the accuracy of the data. Measures such as regular reminder emails could have further improved compliance with experimental requirements. Second, due to design limitations, the interval between the BDI-II survey conducted during freshman health checkups and the semi-structured interviews varied considerably by academic year, ranging from approximately two months to two years. Furthermore, the follow-up period after study completion was not uniform, ending either upon university graduation or on March 31, 2024. Consequently, variability may have been introduced in evaluating the persistence of depressive symptoms. Third, subgroup sizes were unequal, which may have influenced the prediction of MDD onset risk. In particular, because only seven participants developed MDD, conclusions regarding predictive factors for MDD should be interpreted with great caution. Fourth, as this study exploratively examined a wide range of physiological and cognitive markers, the possibility of false-positive findings due to multiple comparisons cannot be excluded. Fifth, the study sample consisted solely of Japanese university students with relatively homogeneous characteristics, which may limit the generalizability of the findings. Moreover, this study was designed to investigate subgroup differences in physiological and cognitive markers within the StD group, as well as whether these subgroups differ in their risk of progression to MDD. However, due to sample size limitations and the exploratory nature of the research, it was not possible to directly evaluate the predictive validity of each marker. Larger-scale longitudinal studies are warranted to clarify the roles of specific physiological and cognitive markers in predicting depression onset.

Considering these limitations, future research should include larger and more diverse samples and employ standardized longitudinal designs to distinguish more clearly between fluctuating and stable depressive states, and to refine the predictive models for MDD onset.

Finally, the markers used in this study—such as voice features, heart rate variability, and activity monitoring—are more easily applicable to clinical and daily-life settings than traditional neuroimaging methods and are also cost-effective. However, as these data are highly personal, careful attention to privacy and ethical considerations is essential. Future work should therefore develop robust frameworks for data governance and ethical oversight, while also optimizing feasibility and cost-efficiency in real-world applications.

Conclusion

The results of this study provide preliminary objective indicators of risk for MDD onset and should not be regarded as an established diagnostic tool at this time. Nevertheless, these findings constitute an important stepping-stone toward the early identification of depression and the development of prevention strategies.

Several limitations should be noted. First, the small number of MDD cases (seven) limits predictive power, and conclusions should be interpreted with great caution. Second, the study population was restricted to Japanese university students with relatively homogeneous characteristics, limiting generalizability. Third, because numerous physiological and cognitive markers were examined exploratorily, the possibility of false-positive findings due to multiple comparisons cannot be excluded.

Accordingly, future longitudinal studies with larger and more diverse samples are needed to validate these markers and to develop and refine predictive models. Furthermore, advancing clinical application will require careful consideration of measurement feasibility, cost, and data privacy/ethical governance, alongside rigorous evaluation of real-world implementation.

Acknowledgments

The authors would like to thank all those who participated in this study.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study was supported (in part) by the Japan Agency for Medical Research and Development (AMED) under grant number JP24wm0625204 and the JST- Mirai Program under grant number JPMJMI20D6.

Disclosure

Dr Junichiro Yoshimoto reports grants from the Japan Science and Technology Agency during the conduct of the study, and grants from the Japan Agency for Medical Research and Development (AMED), the Japan Society for the Promotion of Science, and Honda R&D, Japan, outside the submitted work. Dr Toshi A Furukawa reports personal fees from Boehringer Ingelheim, Daiichi Sankyo, DT Axis, Micron, Shionogi, SONY, UpToDate, and grants from DT Axis and Shionogi, outside the submitted work. Dr Go Okada reports grants from the Japan Agency for Medical Research and Development (AMED) during the conduct of the study. The authors report no conflicts of interest in this work.

References

1. Wittchen HU, Jacobi F, Rehm J, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol*. 2011;21(9):655–679. doi:10.1016/j.euroneuro.2011.07.018
2. Kraus C, Kadriu B, Lanzenberger R, Zarate CA Jr, Kasper S. Prognosis and improved outcomes in major depression: a review. *Transl Psychiatry*. 2019;9(1):127. doi:10.1038/s41398-019-0460-3
3. Ghio L, Gotelli S, Marcenaro M, Amore M, Natta W. Duration of untreated illness and outcomes in unipolar depression: a systematic review and meta-analysis. *J Affect Disord*. 2014;152-154:45–51. doi:10.1016/j.jad.2013.10.002

4. American Psychiatric Association, American Psychiatric Association. *DSM-5 Task Force. Diagnostic and Statistical Manual of Mental Disorders: DSM-5*. 5th ed. Washington, D.C: American Psychiatric Association; 2013.
5. Cuijpers P, Smit F. Subthreshold depression as a risk indicator for major depressive disorder: a systematic review of prospective studies. *Acta Psychiatr Scand*. 2004;109(5):325–331. doi:10.1111/j.1600-0447.2004.00301.x
6. Tuithof M, Ten HM, van Dorselaer S, Kleinjan M, Beekman A, de Graaf R. Course of subthreshold depression into a depressive disorder and its risk factors. *J Affect Disord*. 2018;241:206–215. doi:10.1016/j.jad.2018.08.010
7. Pincus HA, Davis WW, McQueen LE. ‘Subthreshold’ mental disorders. A review and synthesis of studies on minor depression and other ‘brand names’. *Br J Psychiatry*. 1999;174:288–296. doi:10.1192/bjp.174.4.288
8. McElroy SL, Guerdjikova AI, Romo-Nava F. Diagnosing and treating major depressive episodes that lie along the mood disorders spectrum: focus on depression with mixed features. *Cns Spectrums*. 2021;26(2):133–139. doi:10.1017/S1092852921000262
9. Zhang R, Peng X, Song X, et al. The prevalence and risk of developing major depression among individuals with subthreshold depression in the general population. *Psychol Med*. 2023;53(8):3611–3620. doi:10.1017/S0033291722000241
10. Volz HP, Stirnweiß J, Kasper S, Möller HJ, Seifritz E. Subthreshold depression - concept, operationalisation and epidemiological data. A scoping review. *Int J Psychiatry Clin Pract*. 2023;27(1):92–106. doi:10.1080/13651501.2022.2087530
11. Wesselhoeft R, Sorensen MJ, Heiervang ER, Bilenberg N. Subthreshold depression in children and adolescents - a systematic review. *J Affect Disord*. 2013;151(1):7–22. doi:10.1016/j.jad.2013.06.010
12. Lewinsohn PM, Solomon A, Seeley JR, Zeiss A. Clinical implications of “subthreshold” depressive symptoms. *J Abnorm Psychol*. 2000;109(2):345–351. doi:10.1037/0021-843X.109.2.345
13. Karsten J, Hartman CA, Smit JH, et al. Psychiatric history and subthreshold symptoms as predictors of the occurrence of depressive or anxiety disorder within 2 years. *Br J Psychiatry*. 2011;198(3):206–212. doi:10.1192/bjp.bp.110.080572
14. Bertha EA, Balázs J. Subthreshold depression in adolescence: a systematic review. *Eur Child Adolesc Psychiatry*. 2013;22(10):589–603. doi:10.1007/s00787-013-0411-0
15. Jeurig HW, Huisman M, Comijs HC, Stek ML, Beekman ATF. The long-term outcome of subthreshold depression in later life. *Psychol Med*. 2016;46(13):2855–2865. doi:10.1017/S0033291716001549
16. Lee YY, Stockings EA, Harris MG, et al. The risk of developing major depression among individuals with subthreshold depression: a systematic review and meta-analysis of longitudinal cohort studies. *Psychol Med*. 2019;49(1):92–102. doi:10.1017/S0033291718000557
17. Yamabe K, Liebert R, Flores N, Pashos CL. Health-related quality of life outcomes, economic burden, and associated costs among diagnosed and undiagnosed depression patients in Japan. *Clinicoecon Outcomes Res*. 2019;11:233–243. doi:10.2147/CEOR.S179901
18. Johnston DA, Harvey SB, Glozier N, Calvo RA, Christensen H, Deady M. The relationship between depression symptoms, absenteeism and presenteeism. *J Affect Disord*. 2019;256:536–540. doi:10.1016/j.jad.2019.06.041
19. Johns G. Presenteeism in the workplace: a review and research agenda. *J Organizational Behav*. 2010;31(4):519–542. doi:10.1002/job.630
20. Beck AT, Steer RA, Ball R, Ranieri W. Comparison of beck depression inventories -IA and -II in psychiatric outpatients. *J Pers Assess*. 1996;67(3):588–597. doi:10.1207/s15327752jpa6703_13
21. Hamilton M. A rating scale for depression. *J Neurol Neurosurg*. 1960;23(1):56–62. doi:10.1136/jnnp.23.1.56
22. Koch C, Wilhelm M, Salzmann S, Rief W, Euteneuer F. A meta-analysis of heart rate variability in major depression. *Psychol Med*. 2019;49(12):1948–1957. doi:10.1017/S0033291719001351
23. Nilsson A, Sundberg J, Ternstrom S, Askenfelt A. Measuring the rate of change of voice fundamental frequency in fluent speech during mental depression. *J Acoust Soc Am*. 1988;83(2):716–728. doi:10.1121/1.396114
24. Cannizzaro M, Harel B, Reilly N, Chappell P, Snyder PJ. Voice acoustical measurement of the severity of major depression. *Brain Cogn*. 2004;56(1):30–35. doi:10.1016/j.bandc.2004.05.003
25. Shinohara S, Toda H, Nakamura M, et al. Evaluation of the severity of major depression using a voice index for emotional arousal. *Sensors*. 2020;20(18):5041. doi:10.3390/s20185041
26. Taguchi T, Tachikawa H, Nemoto K, et al. Major depressive disorder discrimination using vocal acoustic features. *J Affect Disord*. 2018;225:214–220. doi:10.1016/j.jad.2017.08.038
27. Mason AE, Kasl P, Soltani S, et al. Elevated body temperature is associated with depressive symptoms: results from the TemPredict Study. *Sci Rep*. 2024;14(1):1884. doi:10.1038/s41598-024-51567-w
28. Lyll LM, Cullen B, Lyll DM, et al. The associations between self-reported depression, self-reported chronic inflammatory conditions and cognitive abilities in UK biobank. *Eur Psychiatry*. 2019;60:63–70. doi:10.1016/j.eurpsy.2019.05.007
29. Rock PL, Roiser JP, Riedel WJ, Blackwell AD. Cognitive impairment in depression: a systematic review and meta-analysis. *Psychol Med*. 2014;44(10):2029–2040. doi:10.1017/S0033291713002535
30. Kriesche D, Woll CFJ, Tschentscher N, Engel RR, Karch S. Neurocognitive deficits in depression: a systematic review of cognitive impairment in the acute and remitted state. *Eur Arch Psychiatry Clin Neurosci*. 2023;273(5):1105–1128. doi:10.1007/s00406-022-01479-5
31. Persad SM, Polivy J. Differences between depressed and nondepressed individuals in the recognition of and response to facial emotional cues. *J Abnorm Psychol*. 1993;102(3):358–368. doi:10.1037/0021-843X.102.3.358
32. Krause FC, Linardatos E, Fresco DM, Moore MT. Facial emotion recognition in major depressive disorder: a meta-analytic review. *J Affect Disord*. 2021;293:320–328. doi:10.1016/j.jad.2021.06.053
33. Nyquist AC, Luebbe AM. An emotion recognition-awareness vulnerability hypothesis for depression in adolescence: a systematic review. *Clin Child Fam Psychol Rev*. 2020;23(1):27–53. doi:10.1007/s10567-019-00302-3
34. Gonda X, Juhasz G, Laszik A, Rihmer Z, Bagdy G. Subthreshold depression is linked to the functional polymorphism of the 5HT transporter gene. *J Affect Disord*. 2005;87(2–3):291–297. doi:10.1016/j.jad.2005.05.007
35. Rodríguez MR, Nuevo R, Chatterji S, Ayuso-Mateos JL. Definitions and factors associated with subthreshold depressive conditions: a systematic review. *BMC Psychiatry*. 2012;12:181. doi:10.1186/1471-244X-12-181
36. Judd LL, Schettler PJ, Akiskal HS. The prevalence, clinical relevance, and public health significance of subthreshold depressions. *Psychiat Clin North Am*. 2002;25(4):685–+. doi:10.1016/S0193-953X(02)00026-6
37. Jinnin R, Okamoto Y, Takagaki K, et al. Detailed course of depressive symptoms and risk for developing depression in late adolescents with subthreshold depression: a cohort study. *Neuropsychiatr Dis Treat*. 2017;13:25–33. doi:10.2147/NDT.S117846

38. Zhu Y, Li C, Xie W, Zhong B, Wu Y, Blumenthal JA. Trajectories of depressive symptoms and subsequent cognitive decline in older adults: a pooled analysis of two longitudinal cohorts. *Age Ageing*. 2022;51(1). doi:10.1093/ageing/afab191
39. Dintica CS, Habes M, Schreiner PJ, Launer LJ, Yaffe K. Trajectories in depressive symptoms and midlife brain health. *Transl Psychiatry*. 2024;14(1):169. doi:10.1038/s41398-024-02883-2
40. Lazarus RS. From psychological stress to the emotions: a history of changing outlooks. *Annu Rev Psychol*. 1993;44:1–21. doi:10.1146/annurev.ps.44.020193.000245
41. Shinohara S, Nakamura M, Omiya Y, et al. Depressive mood assessment method based on emotion level derived from voice: comparison of voice features of individuals with major depressive disorders and healthy controls. *Int J Environ Res Public Health*. 2021;18(10):5435. doi:10.3390/ijerph18105435
42. Shinohara S, Toda H, Nakamura M, et al. Evaluation of emotional arousal level and depression severity using voice-derived sound pressure change acceleration. *Sci Rep*. 2021;11(1):13615. doi:10.1038/s41598-021-92982-7
43. Uruguchi T, Omiya Y, Takano T, et al. Evaluation of smart devices in the IoT segment using medical healthcare technology MIMOSYS 2019. 12th Asian Control Conference (ASCC); 2019: 91–96.
44. van Dijk W, Huizink AC, Oosterman M, Lemmers-Jansen ILJ, de Vente W. Validation of photoplethysmography using a mobile phone application for the assessment of heart rate variability in the context of heart rate variability-biofeedback. *Psychosom Med*. 2023;85(7):568–576. doi:10.1097/PSY.0000000000001236
45. Electrophysiology TFOTESOCTNA. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task force of the european society of cardiology and the north american society of pacing and electrophysiology. *Circulation*. 1996;93(5):1043–1065. doi:10.1161/01.CIR.93.5.1043
46. Ekman P. An argument for basic emotions. *Cogn Emot*. 1992;6(3–4):169–200. doi:10.1080/02699939208411068
47. Griffiths S, Jarrold C, Penton-Voak IS, Woods AT, Skinner AL, Munafò MR. Impaired recognition of basic emotions from facial expressions in young people with autism spectrum disorder: assessing the importance of expression intensity. *J Autism Dev Disord*. 2019;49(7):2768–2778. doi:10.1007/s10803-017-3091-7
48. Wagner HL. On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior*. 1993;17(1):3–28. doi:10.1007/BF00987006
49. Ito T, Yokokawa K, Yahata N, Isato A, Suhara T, Yamada M. Neural basis of negativity bias in the perception of ambiguous facial expression. *Sci Rep*. 2017;7(1):420. doi:10.1038/s41598-017-00502-3
50. Snyder HR. Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: a meta-analysis and review. *Psychol Bull*. 2013;139(1):81–132. doi:10.1037/a0028727
51. Sahakian BJ, Owen AM. Computerized assessment in neuropsychiatry using CANTAB: discussion paper. *J R Soc Med*. 1992;85(7):399–402.
52. Masuda Y, Okada G, Takamura M, et al. White matter abnormalities and cognitive function in euthymic patients with bipolar disorder and major depressive disorder. *Brain Behav*. 2020;10(12):e01868. doi:10.1002/brb3.1868
53. Kojima M, Furukawa TA, Takahashi H, Kawai M, Nagaya T, Tokudome S. Cross-cultural validation of the beck depression inventory-II in Japan. *Psychiatry Res*. 2002;110(3):291–299. doi:10.1016/S0165-1781(02)00106-3
54. Matsushita M, Adachi H, Arakida M, et al. Presenteeism in college students: reliability and validity of the presenteeism scale for students. *Qual Life Res*. 2011;20(3):439–446. doi:10.1007/s11136-010-9763-9
55. Jaeger J. Digit symbol substitution test: the case for sensitivity over specificity in neuropsychological testing. *J Clin Psychopharmacol*. 2018;38(5):513–519. doi:10.1097/JCP.0000000000000941
56. Miller AH, Raison CL. The role of inflammation in depression: from evolutionary imperative to modern treatment target. *Nat Rev Immunol*. 2016;16(1):22–34. doi:10.1038/nri.2015.5
57. Souetre E, Salvati E, Wehr TA, Sack DA, Krebs B, Darcourt G. Twenty-four-hour profiles of body temperature and plasma TSH in bipolar patients during depression and during remission and in normal control subjects. *Am J Psychiatry*. 1988;145(9):1133–1137.
58. Montano N, Ruscone TG, Porta A, Lombardi F, Pagani M, Malliani A. Power spectrum analysis of heart rate variability to assess the changes in sympathovagal balance during graded orthostatic tilt. *Circulation*. 1994;90(4):1826–1831. doi:10.1161/01.CIR.90.4.1826
59. Hayano J, Kiyono K, Struzik ZR, et al. Increased non-gaussianity of heart rate variability predicts cardiac mortality after an acute myocardial infarction. *Front Physiol*. 2011;2:65. doi:10.3389/fphys.2011.00065
60. Tan Y, Zhou M, Wang J, et al. Heart rate variability in subthreshold depression and major depressive disorder. *J Affect Disord*. 2025;373:306–313. doi:10.1016/j.jad.2025.01.003
61. Brown L, Karmakar C, Gray R, Jindal R, Lim T, Bryant C. Heart rate variability alterations in late life depression: a meta-analysis. *J Affect Disord*. 2018;235:456–466. doi:10.1016/j.jad.2018.04.071
62. Harmer CJ, O’Sullivan U, Favaron E, et al. Effect of acute antidepressant administration on negative affective bias in depressed patients. *Am J Psychiatry*. 2009;166(10):1178–1184. doi:10.1176/appi.ajp.2009.09020149
63. Bone JK, Lewis G, Button KS, et al. Variation in recognition of happy and sad facial expressions and self-reported depressive symptom severity: a prospective cohort study. *J Affect Disord*. 2019;257:461–469. doi:10.1016/j.jad.2019.06.025
64. Chiu I, Piguet O, Diehl-Schmid J, et al. Facial emotion recognition performance differentiates between behavioral variant frontotemporal dementia and major depressive disorder. *J Clin Psychiatry*. 2018;79(1):16m11342. doi:10.4088/JCP.16m11342
65. Munkler P, Rothkirch M, Dalati Y, Schmack K, Sterzer P. Biased recognition of facial affect in patients with major depressive disorder reflects clinical state. *PLoS One*. 2015;10(6):e0129863. doi:10.1371/journal.pone.0129863
66. Austin MP, Ross M, Murray C, O’Carroll RE, Ebmeier KP, Goodwin GM. Cognitive function in major depression. *J Affect Disord*. 1992;25(1):21–29. doi:10.1016/0165-0327(92)90089-0

Neuropsychiatric Disease and Treatment

Dovepress
Taylor & Francis Group

Publish your work in this journal

Neuropsychiatric Disease and Treatment is an international, peer-reviewed journal of clinical therapeutics and pharmacology focusing on concise rapid reporting of clinical or pre-clinical studies on a range of neuropsychiatric and neurological disorders. This journal is indexed on PubMed Central, the 'PsycINFO' database and CAS, and is the official journal of The International Neuropsychiatric Association (INA). The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/neuropsychiatric-disease-and-treatment-journal>