Received: 4 April 2024 | Revised: 1 October 2024 | Accepted: 4 November 2024

DOI: 10.1111/psyp.14729

PSYCHOPHYSIOLOGY SPR WILEY

ORIGINAL ARTICLE

Contribution of physiological dynamics in predicting major depressive disorder severity

Esther García Pagès[1,2] | **Spyridon Kontaxis**[3] | **Sara Siddi**[4,5] | **Mar Posadas-de Miguel**[3] | **Concepción de la Cámara**[3] | **Maria Luisa Bernal**[3] | **Thais Castro Ribeiro**[1,2] | **Pablo Laguna**[2,3] | **Llorenç Badiella**[5] | **Raquel Bailón**[2,3] | **Josep Maria Haro**[4,6,7] | **Jordi Aguiló**[1,2]

[1]Department de Microelectrònica i Sistemes electrònics, Universitat Autònoma de Barcelona, Cerdanyola del Vallès, Spain

[2]Centro de Investigación Biomédica en Red de Bioingeniería, Biomateriales y Nanomedicina, Madrid, Spain

[3]Universidad de Zaragoza, Zaragoza, Spain

[4]Parc Sanitari Sant Joan de Déu, Institut de Recerca Sant Joan de Déu, Sant Boi de Llobregat, Spain

[5]Departament de Matemàtiques, Universitat Autònoma de Barcelona, Cerdanyola del Vallès, Spain

[6]Centro de Investigación Biomédica en Red de Salud Mental, Madrid, Spain

[7]Universitat de Barcelona, Barcelona, Spain

Correspondence
Esther García Pagès, Department de Microelectrònica i Sistemes electrònics, Universitat Autònoma de Barcelona, Carrer de les Sitges Engineering School, Building Q, 08193, Bellaterra, Cerdanyola del Vallès, Spain.
Email: esther.garcia.pages@uab.cat

Funding information
Instituto de Salud Carlos III, Grant/Award Number: CB06/01/0049; Agencia Estatal de Investigación, Grant/Award Number: RTI2018-096072-B-I00 and TED2021-131106B-I00; European Social Fund Plus, Grant/Award Number: T39_23R; Innovative Medicines Initiative, Grant/Award Number: 115902

Abstract

This study aimed to explore the physiological dynamics of cognitive stress in patients with Major Depressive Disorder (MDD) and design a multiparametric model for objectively measuring severity of depression. Physiological signal recordings from 40 MDD patients and 40 healthy controls were collected in a baseline stage, in a stress-inducing stage using two cognitive tests, and in the recovery period. Several features were extracted from electrocardiography, photoplethysmography, electrodermal activity, respiration, and temperature. Differences between values of these features under different conditions were used as indexes of autonomic reactivity and recovery. Finally, a linear model was designed to assess MDD severity, using the Beck Depression Inventory scores as the outcome variable. The performance of this model was assessed using the MDD condition as the response variable. General physiological hyporeactivity and poor recovery from stress predict depression severity across all physiological signals except for respiration. The model to predict depression severity included gender, body mass index, cognitive scores, and mean heart rate recovery, and achieved an accuracy of 78%, a sensitivity of 97% and a specificity of 59%. There is an observed correlation between the behavior of the autonomic nervous system, assessed through physiological signals analysis, and depression severity. Our findings demonstrated that decreased autonomic reactivity and recovery are linked with an increased level

of depression. Quantifying the stress response together with a cognitive evaluation and personalization variables may facilitate a more precise diagnosis and monitoring of depression, enabling the tailoring of therapeutic interventions to individual patient needs.

KEYWORDS

cognitive test, generalized linear models, major depressive disorder, physiological signals, stress reactivity

1 | INTRODUCTION

According to a systematic analysis by the GBD 2019 Mental Disorders Collaborators (Ferrari, 2022) there has been a significant increase in the prevalence of mental disorders, paralleling the surge in mHealth technologies (Naslund et al., 2015). The emergence of smartphones and wearable devices, which use embedded sensors to unobtrusively measure human behavior and physiological responses, has opened a world of opportunities for advancement in mental health diagnosis and patient follow-up.

Major Depressive Disorder (MDD) is one of the most prevalent mental health disorders. It affects over 7% of the adult population in Europe, and it is about 50% more common among women than men (Arias-de la Torre et al., 2021). Since patients with MDD are at increased risk of suicide, close monitoring, and follow-up by mental health workers become necessary to ensure safety and adherence to mental health treatment. Undoubtedly, the use of mHealth technologies to monitor these patients could facilitate non-intrusive yet thorough tracking, and enable timely intervention if needed. However, valid and reliable objective methods to monitor patients at risk are still lacking (Matcham et al., 2019). At present, clinical interviews are the most reliable method for determining the etiological diagnosis of MDD relapses. A considerable proportion of this data is subjective, that is, it is predicated on the patient's self-report and is subject to the clinician's interpretation. As a result, it is susceptible to a multitude of biases that have the potential to compromise the quality of healthcare provision. To mitigate these concerns, it might be advantageous to utilize the physiological data and examine its association with the severity of depression.

While extensive research on depression has been conducted, our understanding of the etiology and pathophysiology remains limited, leading to the stress-diathesis model becoming a prevalent explanation for the onset of MDD (Booij et al., 2013; Van Heeringen, 2012). This model posits that mental disorders develop due to a combination of genetic vulnerability (diathesis) and environmental stressors. In this context, individuals might have different stress reactivity patterns that lead to a predisposition to developing a disorder when exposed to significant stressors.

Stress reactivity refers to the physiological response that an individual has to stressors. Reactivity to stressors yields insights into health-related risks that cannot be gleaned from resting measurements (Kiecolt-Glaser et al., 2020; Schiweck et al., 2019). This can include digital biomarkers such as changes in heart rate, blood pressure, rapid breathing and other bodily functions that can be objectively monitored using wearables. For instance, in the NESDA cohort (Hu et al., 2016), while there were no observable differences in heart rate variability (HRV) between depressed and non-depressed individuals during resting states, hyporeactivity was prominent when a stressor was introduced. Similar results were found in Adolph et al. (2018) where high frequency-HRV (HF-HRV) reactivity, but not resting HF-HRV at baseline, was predictive of higher scores on suicidal ideation.

Back in 1987, Thorell et al. (1987) already linked hyporeactive electrodermal activity (EDA) and low stimulus-unrelated phasic activity to suicidal behavior in depressive patients. Another multi-site study, the EUDOR-A project, also found that a hyporeactive EDA response pattern was associated with a history of suicide attempts as well as with an increased risk of such behavior during a 12-month follow-up period after being tested (Carli et al., 2022). A systematic review on this topic also found that electrodermal hypoactivity seemed to be a reliable feature of depression and a valid marker of suicidal risk (Sarchiapone et al., 2018). Finally, decreased autonomic reactivity assessed by dynamic changes in the photoplethysmographic waveform was also associated with higher depression severity in a published work from our group with the same participants of the present study (Kontaxis, Gil, et al., 2021). Most studies typically explored just one digital measure, with only a handful attempting to integrate measures such as EDA with HRV (Ettore et al., 2023). Nevertheless, their sensitivity and specificity were not sufficient to support the use of these markers alone as suicide risk indicators in clinical practice.

Considering these limitations, our research aimed to enhance the understanding of MDD physiological

response and to design a multiparametric model that could lead to remote monitoring of depressed patients. To the best of our knowledge, this is the first study designed to evaluate stress reactivity through multiple physiological signals during a mild cognitive stressor, using two validated neuropsychological tests to assess cognitive executive functions as well, and psychometric tests as the gold standard for mental health evaluation. It has been established that diminished executive functioning is evident during the initial occurrence of depression (Varghese et al., 2022). This impairment endures even after remission (Rock et al., 2014) and it is exacerbated in recurrent MDD (Matcham et al., 2023). This approach allowed us to objectively measure multiple and synchronous physiological signals, providing a more accurate representation of an individual's stress response profile while assessing their executive functioning.

2 | METHODOLOGY

2.1 | Participants

Forty MDD patients and 40 healthy controls (HC) were recruited consecutively at the Hospital Clínico Lozano Blesa (Zaragoza, Spain) and the Acute Unit at the Sant Joan de Déu Numancia (Barcelona, Spain) from October 2016 until December 2019. Following the Declaration of Helsinki, all participants voluntarily signed an informed consent form validated by the ethics committee of each institution. The Clinical Research Ethics Committee of the Hospital Clínico Lozano Blesa approved the protocol the 8th of June 2016, identification number PI16-0156, while the Medicine Research Ethics Committee of the Sant Joan de Déu Research Foundation approved it on December 13, 2016, PIC-148-16.

The selection criteria for the patient group was based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (American Psychiatric Association, 2022) for non-psychotic MDD and to fall within the age range of 18–65 years. Depression severity was assessed with a score higher than 19 on the Hamilton Depression Rating Scale (HDRS) (Hamilton, 1960). Participants with other comorbidities such as cardiovascular, hypothalamic–pituitary–adrenal axis (HPA)-related disorders or neurological disorders were excluded. The use of various pharmacological treatments, including benzodiazepines, beta-blockers, tricyclic antidepressants, and anticholinergics, was collected. In the statistical analyses, antidepressant and anxiolytic drugs were grouped as categorical variables to ensure an adequate sample size in order to control for their confounding effects on the physiological markers. Other pharmacological treatments were excluded from the analysis.

For the HC group, the inclusion criteria were to be a healthy individual without comorbidities such as cardiovascular, endocrinological or neurological disorders, and whose age, sex, and body mass index (BMI) ($\pm 3\,kg/m^2$) matched those of the patient group. Any participant with psychopathology or regular consumption of psychotropic substances was excluded. Moreover, all participants were asked not to consume tobacco, alcohol, or stimulants of the nervous system (caffeine, taurine, etc.) for at least 8 h before participation in the study.

2.2 | Experimental protocol

The experimental protocol (summarized in Figure 1) had a duration of about 1 h and it consisted of a preparation period followed by physiological signals monitoring divided into three stages.

During the preparation period, a licensed psychologist asked for the sociodemographic and clinical data and conducted a psychological evaluation using the Clinical Global Impression (CGI) (Guy, 1976), the Hamilton Rating Scale for Depression (HRSD) (Hamilton, 1960), and the Beck Depression Inventory-II (Beck et al., 1996). Afterwards, the individuals were asked to fill the self-reported psychometric tests including the State–Trait Anxiety Inventory Test (Spielberg et al., 2008), the Perceived Stress Scale (PSS) (Remor, 2006), and the Visual Analogue Stress Scale (Lesage et al., 2012).

Once all the questionnaires were completed, a technician placed the medical sensors according to the Medicom system (ABP10 module of Medicom MTD, Ltd., Russia) manual. For the first stage, the baseline stage, participants

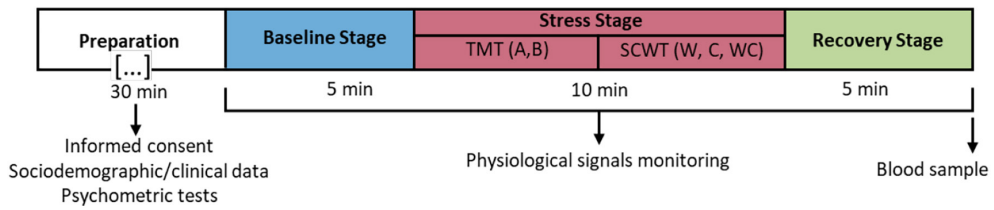

FIGURE 1 Scheme of experimental procedure.

were requested to stay quiet in a seated position. During the second stage, the stress stage, stress was induced in participants using two cognitive tests: (a) the Trail Making Test (TMT) and, (b) the Stroop Color and Word Test (SCWT), consisting of a non-verbal and verbal stressor, respectively (Scarpina & Tagini, 2017). Afterward, in the recovery stage, individuals were requested to relax and remain silent for 5 min. Finally, following the conclusion of the three session stages, blood samples were taken for biochemical analysis, aiming to explore potential physiological changes linked to depression.

2.3 | Outcome measures

2.3.1 | Psychometric tests

For the determination of mental health status and the level of depression, a psychologist administered the HRSD (Hamilton, 1960; Ma et al., 2021; Ramos-Brieva & Cordero Villafafila, 1986), the Beck Depression Inventory (BDI) (Gallagher et al., 1983; Wang & Gorenstein, 2013; Wiebe & Penley, 2005) and the CGI (Guy, 1976) through an interview. Additional psychometric tests to determine the level of stress were self-administered. The State–Trait Anxiety Inventory (STAI) (Guillén Riquelme & Buela Casal, 2011; Spielberg et al., 2008) for trait and state anxiety, the PSS (Cohen et al., 1983; Remor, 2006) to assess stressful life situations, and the Visual Analogue Scale for stress (VASS) (Lesage et al., 2012).

All the psychometric tests used are validated, well documented by the medical world, and have a contrasted and accepted Spanish language version. Psychometric tests required, on average, 15–20 min to be completed, and participants did not need any special education or training to complete them.

2.3.2 | Cognitive tests

For the Stress Stage, we used two tests that require cognitive control and the engagement of executive functions, such as attention, cognitive flexibility, and processing speed. Therefore, they can be stressful for individuals, especially under time pressure or when high performance is demanded.

- Trail Making Test (TMT). The TMT is a neuropsychological test of visual attention, processing speed, and task switching. It consists of two parts in which the individual is instructed to connect a set of 25 dots as quickly as possible while still maintaining accuracy. In the first part (TMTa), the individuals are required to connect, by drawing a line, consecutive numbers, while in the second one (TMTb), they are required to connect numbers and letters in an alternating progressive sequence. The TMT was scored based on the time it takes to complete the test (Bowie & Harvey, 2006).
- Stroop Color and Word Test (SCWT). This psychological test measures a person's selective attention capacity and skills, as well as their processing speed ability. The individuals are required to read, as fast as possible in a pre-established period of 45 secs, from three different lists, the names of: (a) color-words printed in black ink (STROOP_W), (b) different color patches (STROOP_C), (c) ink color of color-words printed in an inconsistent color ink (STROOP_WC, incongruent condition). To determine the pure interference score (STROOP_I), the difference between the STROOP_WC and the estimated STROOP_WC (STROOP_WC') were calculated. The higher the resulting score, the less susceptible the subject is to interference (Golden, 2020).

2.3.3 | Biochemical variables

The biochemical variables considered in the present study offer information about the HPA axis and the autonomic nervous system (ANS). The biochemical variables measured in blood samples were prolactin, interleukin 6 (IL-6), tumor necrosis factor alpha (TNF-α), and glucose. The samples were kept frozen at −20°C until processing at the Biomedical Diagnostic Center of the Hospital Clínic of Barcelona.

2.3.4 | Physiological signals

The signals chosen given their relation to the stress response were electrocardiography (ECG; sampling frequency: 1000 Hz), photoplethysmography (PPG; 250 Hz), EDA (250 Hz), skin temperature (ST; 250 Hz) and respiration (Resp; 250 Hz). All the signals were recorded using the Medicom 83 system, ABP-10 module (Medicom MTD Ltd., Russia). MATLAB software was used for the signal processing and feature extraction in 1-min intervals.

The wavelet transform-based approach described in Martínez et al. (2004) was used to detect beats in the ECG data, and the algorithm described in Mateo and Laguna (2003) was used to confirm the presence of ectopic or abnormal beats. An energy-based artifact detection was used to identify and remove PPG artifacts as described in Armañac et al. (2019). Afterwards, utilizing the algorithm given in Gil Herrando (2009), the remaining artifacted segments were eliminated according to Hjorth parameters in windows of 2 s spaced every half-second. Lastly,

the method described in Lázaro et al. (2014) was used to detect beats or pulses.

The beat occurrence times for the PPG and ECG data were used to construct the classical time and frequency domain indices. While frequency domain indices were calculated from the instantaneous heart rate signal using the classical frequency bands (Mateo & Laguna, 2003), time domain indices were calculated from the RR interval series (see Table 1). The time interval between the beats captured on the ECG and the pulse captured on the PPG was then used to estimate the pulse arrival time (PAT) (Arza et al., 2018).

Using a Fast Fourier transform to calculate the frequency corresponding to the greatest peak of the respiratory power density spectrum, the respiratory rate (RR) was identified (Lázaro et al., 2014). We used the convex optimization model (cvxEDA) for the EDA processing and decomposition (Greco et al., 2016) and for the frequency-domain analysis, we followed the approach in Posada-Quintero et al. (2016). Finally, since any artifact was identified in the temperature signal, no preprocessing was needed.

Additional information about the signal preprocessing and feature extraction was included in our earlier publications (García Pagès et al., 2023; Garzón-Rey & Aguiló Llobet, 2017). The final list of features considered in the analysis is summarized in the following table:

Finally, for all the above-mentioned features, the average for each stage was used as static feature and the intra-individual difference of a static feature between the different stages is used as a dynamic index of autonomic reactivity ($\Delta(F)$). The reactivity and recovery indices were calculated by subtracting the value of the feature at the baseline (B) and recovery (R) stages from the value at the stress (S) stage and denoted as $\Delta(F)_B^S$ and $\Delta(F)_R^S$ respectively, following the strategy previously used in Kontaxis, Gil, et al. (2021) with the same participants.

2.4 | Statistical analysis

All analyses were run in SAS 9.4 software (SAS Institute, Cary, NC, USA). To compare the distribution of categorical variables across groups, several statistical tests were applied, each chosen based on specific criteria and assumptions. The McNemar Test was used when comparing paired binary variables while when analyzing categorical variables with more than two groups, the Bowker test was employed. The Chi-squared test was used to compare proportions between groups. However, Fisher's exact test was preferred when dealing with small sample sizes. The choice between these tests is determined by the criterion that less than 25% of cells should have expected frequencies below 5, following Cochran's guidelines. The selection of the most suitable test was made on a case-by-case basis to ensure accurate analysis.

For the comparison of quantitative variables, the paired t test was chosen when comparing two normally distributed groups, and the Wilcoxon Signed-Rank Test was used in the cases where the assumptions for the Paired t test were not met. This non-parametric test does not rely on normal distribution assumptions and is robust when dealing with skewed or non-normally distributed data. The appropriateness of these tests was confirmed through the Shapiro–Wilk test. Finally, mixed effect models were used to assess confounding factors that may interfere, considering the subject as a random factor and visit, group, and their interaction as fixed effects.

Regarding the physiological features, the multivariate maximum likelihood method was used to handle missing data. To fulfill the assumption of normality required by the model generation methods, the square root or a log transformation was applied when necessary. With the resulting complete dataset, cognitive and physiological features were analyzed with linear models considering BDI and confounding factors as explanatory variables. Finally, to assess the relationship between MDD severity with respect to the outcome features and confounding factors, we used a linear model with the BDI score as the response variable. Interactions between variables were also explored. The performance of the resulting model was assessed using logistic regression with the MDD condition in this case as the response variable. The significance level was set at 0.05. Post-hoc comparisons were corrected for multiplicity using Tukey's correction.

3 | RESULTS

3.1 | Demographics and clinical characteristics

The study included 80 participants recruited, 40 candidates per site were eligible to participate.

Table 2 displays the baseline characteristics of the sample. The population studied corresponds to a group of 33 men and 47 women with an average age of 45 ± 13 years. Nearly half of the population with depression had some form of non-psychiatric chronic disease, while only 20% of the HC group did. Since the disease was managed, it did not represent an inconvenience to carry out the measurement sessions.

Despite our efforts to control for patient-control matching by BMI ($\pm 3\,kg/m^2$), patients consistently exhibited a higher BMI. This difference was statistically significant,

TABLE 1 List of extracted features from physiological signals.

Physiological signal	Extracted feature	Definition	ANS relation
Electrocardiography (ECG) or/and photoplethysmography (PPG)	mHR (bpm)	Mean heart rate	Sympathetic and parasympathetic modulation (Malik et al., 1996)
	SDNN (s)	Standard deviation of intervals between beats considered normal (NN)	Parasympathetic modulation (<5-min window) (Malik et al., 1996)
	RMSSD (s)	Root mean square of successive differences between adjacent NN intervals	Parasympathetic modulation (Malik et al., 1996)
	PHF (s^{-2})	HRV High frequency band power (0.15–0.4 Hz)	Parasympathetic modulation (respiratory sinus arrhythmia) (Malik et al., 1996)
	PLF (s^{-2})	HRV Low frequency band power (0.04–0.15 Hz)	Sympathetic and parasympathetic modulation (baroreflex) (Malik et al., 1996)
	LF_HF	Ratio of low-frequency to high-frequency power	Sympathetic and parasympathetic balance (Malik et al., 1996)
	mPAT (ms)	Mean pulse arrival time	Sympathetic and parasympathetic modulation (Deshmukh et al., 2022)
	stdPAT (ms)	Standard deviation of the pulse arrival time	
Electrodermal activity (EDA)	mTonic (μS)	Average value of the tonic component, that is, Skin Conductance Level (SCL)	Sympathetic modulation (Greco et al., 2016; Posada-Quintero et al., 2016)
	stdTonic (μS)	Standard deviation of the tonic component	
	mPhasic (μS)	Average value of the phasic component	
	stdPhasic (μS)	Standard deviation of the phasic component	
	aucPhasic (μS·s)	Area under the curve value of the phasic component normalized by the length of the session, i.e. Skin Conductance Responses (SCRs)	
	EDASymp (μS^2)	Power spectral density of EDA signal in the frequency range of 0.045–0.25 Hz	
Skin temperature (ST)	mGrad (°C)	Mean temperature gradient	Sympathetic modulation (Greaney et al., 2016)
Respiration (Resp)	RR (Hz)	Respiratory rate	Sympathetic and parasympathetic influence (Cherniack, 1990)
	Pk (%)	RR peak percentage in the power spectrum	

albeit with a mild effect size ($r = .25$). The age and gender matching between patients and controls was found to be accurate, ensuring that age-related factors did not confound our results.

Interestingly, we did not observe any differences in biochemical parameters between the two groups, with one exception: IL-6 levels were elevated in patients. This remained true even after we removed outliers and confirmed that the number of individuals with allergies was similar in both groups.

Most patients were taking anxiolytics (75%) and taking antidepressants (82%). Of the 40 patients, 26 combined both types of medication whereas only three patients did not take any medication.

In our study, we found no differences in having children, coffee consumption, or living environment between the two groups. On the other hand, we found a significant difference in the marital status between the HC and the MDD group. Among the HC, married individuals were prevalent. However, in the MDD group, there was a significant increase in the number of single or divorced individuals. Educational level emerged as a significant differentiator between the groups. A substantial 82.5% of the controls had completed university education, while only 22% of the patients had achieved this level. Most patients had either completed secondary education or discontinued education after primary school. Occupation also showed clear differences between the groups. The majority of controls were employed, while most patients were either unemployed or retired.

Finally, in terms of lifestyle habits, we found significant differences in smoking and alcohol consumption. Nearly half of the patients were regular smokers compared to only 15% of the controls. On the contrary, 62% of the controls consumed alcohol occasionally or moderately, while 60%

TABLE 2 Demographics and clinical characteristics.

	HC	MDD
BMI (m, std)**	24.78 (4.64)	27.15 (5.17)
Age (m, std)	44.33 (12.27)	45.40 (13.35)
Glucose (m, std)	85.53 (18.37)	85.37 (14.93)
Prolactin (m, std)	10.89 (5.78)	9.28 (5.65)
IL6 (m, std)**	1.79 (1.35)	2.82 (1.91)
TNFα (m, std)	3.99 (4.40)	3.36 (4.26)
Marital status (*n*, %)		
Married	26 (65)	15 (37.50)
Separated	3 (7.50)	8 (20)
Single	10 (25)	14 (35)
Widower	1 (2.50)	3 (7.50)
Gender (*n*, %)		
Female	24 (60)	23 (57.50)
Male	16 (40)	17 (42.50)
Children (*n*, %)		
No	16 (40)	19 (47.50)
Yes	24 (60)	21 (52.50)
Education (*n*, %) **		
Elementary education	2 (5)	14 (35)
High school	5 (12.50)	17 (42.50)
University	33 (82.50)	9 (22.50)
Occupation (*n*, %)**		
Employed	33 (82.50)	11 (27.50)
Retired	2 (5)	6 (15)
Student	3 (7.50)	4 (10)
Unemployed	2 (5)	19 (47.50)
Environment (*n*, %)		
Rural	5 (12.50)	10 (25)
Urban	35 (87.50)	30 (75)
Tobacco (*n*, %)*		
Non-smoker	34 (85)	21 (52.50)
Smoker	6 (15)	19 (47.50)
Coffee (*n*, %)		
No	10 (25)	9 (22.50)
Yes	30 (75)	31 (77.50)
Alcohol (*n*, %)*		
No	10 (25)	24 (60)
Occasional	22 (55)	13 (32.50)
Moderate	5 (12.50)	3 (7.50)
High	3 (7.50)	
Sports (*n*, %)*		
No	7 (17.50)	19 (47.50)
Occasional	7 (17.50)	11 (27.50)
Usual	26 (65)	10 (25)
Chronic disease (*n*, %)*		
No	32 (80)	21 (52.50)
Yes	8 (20)	19 (47.50)
Medicines (*n*, %)**,M		
No	26 (65)	5 (12.50)
Yes	14 (35)	35 (87.50)
Anxiolytic (*n*, %)**		
No	40 (100)	10 (25)
Yes		30 (75)
Antidepressive (*n*, %)**		
No	40 (100)	7 (17.50)
Yes		33 (82.50)

Note: M, benzodiazepines, beta-blockers, tricyclic antidepressants, and anticholinergics.

*$p < .05$. **$p < .005$.

of the patients did not consume alcohol at all, although this might be due to contraindications with medication. Regarding physical activity, most controls regularly engaged in sports activities, whereas only 25% of patients did so and 47% led sedentary lifestyles.

3.2 | Psychometric evaluation

The following table summarizes the results obtained from the psychometric evaluations:

In terms of mental health outcomes shown in Table 3, both the BDI and the HDRS results show that the MDD group had moderate to severe symptoms of depression whereas the HC group exhibited minimal or no symptomatology, as expected. All psychometric questionnaires pointed toward a higher severity of mental distress in the MDD group; both state and trait anxiety were around 3 times higher, and patients showed about twice the stress level compared to the HC.

This psychological distress was further corroborated by the cognitive test performance. Patients with MDD completed fewer words in the SCWT; 28 less on average for the STROOP_W part, 21 for the STROOP_C part, and a lower but still significant difference in the STROOP_WC part (14 words less). Results also showed lower resistance to interference among the MDD group, although the difference was not statistically significant ($p = .26$). Additionally, concerning the TMT, patients took almost double the time to complete both subtasks.

To mitigate the potential distortion of depression assessment due to the influence of confounding variables associated with cognitive scores, we analyzed the most

TABLE 3 Psychometric evaluation results.

	HC	MDD
CGI (*n*, %)**		
Healthy	40 (100)	
Mild		9.00 (22.50)
Moderate		21.00 (52.50)
Severe		9.00 (22.50)
Extreme		1.00 (2.50)
HDRS (m, std)**	2.13 (2.59)	22.20 (6.51)
BDI (m, std)**	3.75 (3.71)	27.48 (12.25)
PSS (m, std)**	11.15 (4.50)	26.93 (7.05)
STAIs (m, std)**	9.45 (7.28)	30.25 (10.19)
STAIt (m, std)**	12.70 (7.91)	37.33 (10.45)
VASS (m, std)**	23.06 (19.46)	44.83 (27.55)
TMTa (m, std)**	24.88 (9.41)	43.60 (22.37)
TMTb (m, std)**	58.70 (28.48)	100.73 (46.20)
STROOP_W (m, std)**	111.65 (19.32)	83.60 (20.75)
STROOP_C (m, std)**	77.10 (13.82)	56.10 (11.95)
STROOP_WC (m, std)**	49.33 (13.10)	35.13 (12.74)
STROOP_WC' (m, std)**	45.36 (7.42)	33.26 (7.11)
STROOP_I (m, std)	3.96 (8.34)	1.86 (9.82)

Note: word (W), color (C), word-color (WC), estimated word-color (WC') and interference (I).

Abbreviations: BDI, Beck Depression Inventory; CGI, Clinical Global Impression; HRSD, Hamilton Rating Scale for Depression; PSS, Perceived Stress Scale; STAI, State–Trait Anxiety Inventory for state and trait anxiety; STROOP, Stroop Color and Word Test subscores; TMT, Trail Making Test parts a and b; VASS, Visual Analogue Scale for stress.

**$p < .005$.

demanding subtasks for each test (STROOP_WC and the TMTb) controlling for age, gender, education level, and medication.

Results showed that age had a significant effect on STROOP_WC scores when controlling for BDI score ($F = 6.03$, $p = .02$), although age did not have a significant effect on TMTb scores. The interaction between age and depression was not significant in either test.

Regarding the gender effect, neither the gender itself nor the interaction with depression were independent predictors of the cognitive scores. This was also the case when analyzing the consumption of antidepressants and anxiolytics, as well as the anxiety trait measured by the STAIt score. However, the antidepressant effect on the TMTb score was close to the significance level ($F = 2.92$, $p = .09$).

Finally, education had an independent significant effect on STROOP_WC scores ($F = 4.18$, $p = .02$), whereas the interaction between education and BDI was not significant. In other words, the effect of depression severity on performance was consistent across different education levels. On the contrary, no significant effects were found for TMTb scores when controlling for depression. However, Tukey's post-hoc test showed significant differences between the mean scores in both tests showing poorer test performances at lower educational levels. For the STROOP_WC, fewer words were completed comparing Elementary Education vs. University (−20.60 words) while for the TMTb scores, it took more time to complete the task (47.53 s).

3.3 | Static levels of physiological features

Physiological data analysis showed several distinct feature reactivity patterns between MDD patients and HC. Results on the physiological features extracted during each protocol stage are presented in Figure 2 and the extended table in the supplementary material (Table A1).

From the ECG and PPG signals, results showed that the mHR increased during the stress stage and decreased during subsequent recovery in healthy individuals. However, this change was not as pronounced in patients, who also exhibited a higher heart rate during the entire session. HRV features showed similar patterns in the two groups even though all features both from time and frequency analysis showed reduced variability in patients. Particularly, the PHF feature extracted from the ECG signal showed significantly better sensibility to separate both groups compared to the PPG extracted feature, although they show a high correlation as expected.

Contrary to the mHR pattern, the mPAT decreased with stress and increased during the recovery stage in HC, while it remained constant in patients, suggesting poor cardiovascular control. Similarly, HC exhibited a greater decreasing gradient of the finger temperature during the stress stage compared to the baseline and recovery stages, while the patient group maintained a more uniform temperature. No differences were found for the face temperature.

EDA showed interesting trends as well. In healthy individuals, the EDA tonic and phasic components increased under stress conditions and decreased during subsequent recovery. However, in MDD patients, this stress state was maintained throughout the session or even continued to increase. Regarding the frequential analysis, the EDASymp feature showed similar results of sympathetic dysregulation.

Finally, the respiratory signal exhibited similar results in all participants except for the respiratory frequency during the cognitive tests, where patients showed significantly higher values, 15.5 breaths per minute compared to 13.2 for the HC group.

We examined the interactions between various physiological parameters concerning the intake of antidepressants using mixed models. Our analysis encompassed all

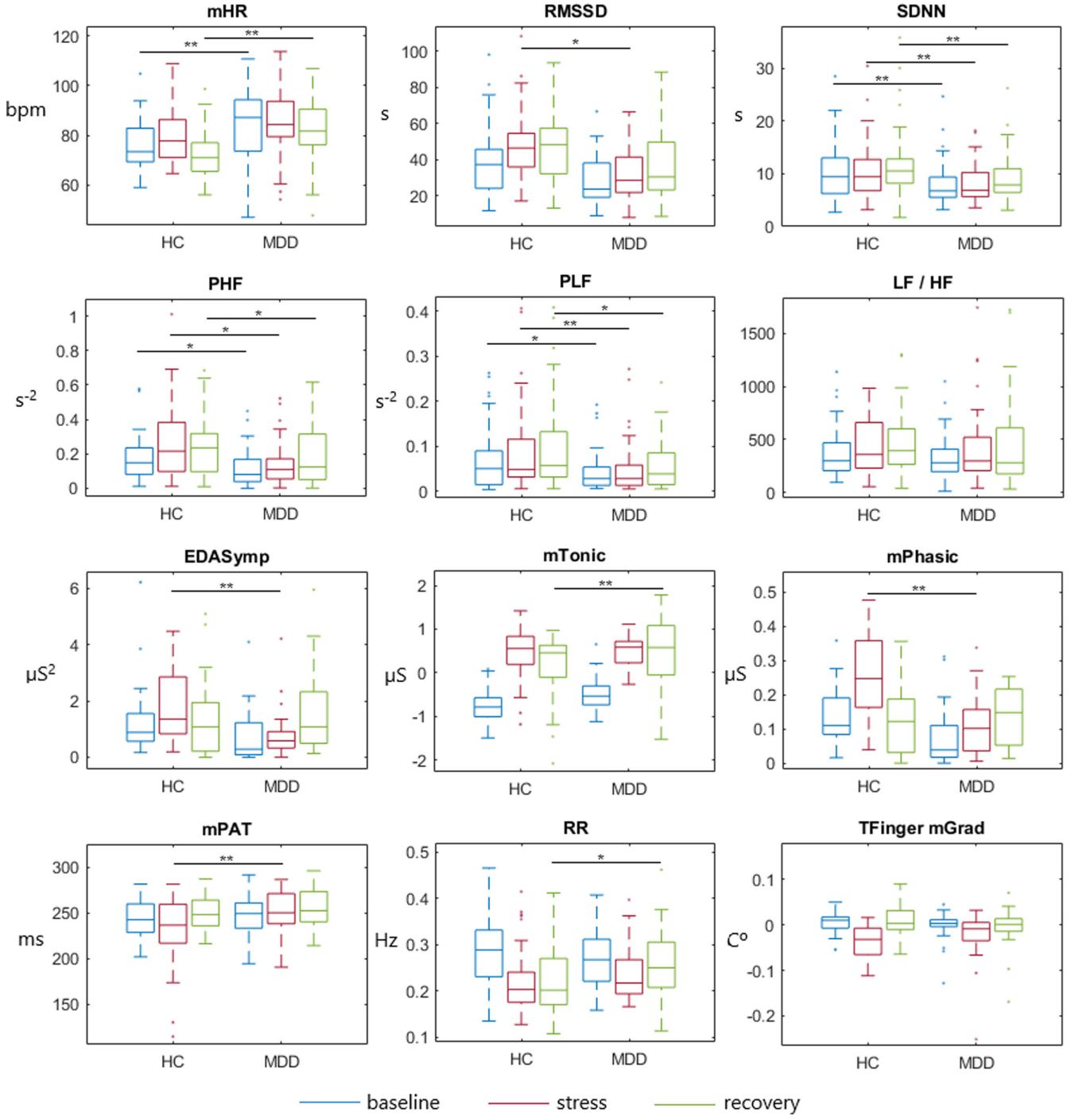

FIGURE 2 Comparison between depressive (MDD) and healthy (HC) individuals of the evolution of the features (*$p < .05$, **$p < .005$). EDASymp, EDA band power; LF/HF, PLF and PHF ratio; mHR, mean heart rate; mPAT, mean pulse arrival time; mPhasic, mean value of the EDA phasic component; mTonic, mean value of the EDA tonic component; PHF, HRV high frequency band power; PLF, HRV low frequency band power; RMSSD, root mean square of successive differences between normal to normal intervals; RR, respiratory rate; SDNN, standard deviation of normal to normal intervals; TFinger mGrad, mean finger temperature gradient.

parameters, however, our findings indicated no significant differences attributable to antidepressant consumption. Furthermore, our analysis extended to consider potential confounding factors such as tobacco use, coffee and alcohol consumption, and gender. However, none of these factors significantly impacted the physiological parameters under study.

As expected, there were significant differences between the mHR of participants who do not practice sports compared to those who regularly participate in sports.

Finally, we found that with increasing age, the SDNN is reduced in the baseline stage with a significant negative correlation ($r=-.54$) indicating less adaptability to stressors, while mPAT increased with age, albeit with a weaker correlation ($r=.22$).

3.4 | Dynamics of physiological features

To assess stress response, we employed a strategy of using the difference between the stress and the baseline stages for the reactivity feature ($\Delta(F)_B^S$), and between the stress and the recovery stages for the recovery feature ($\Delta(F)_R^S$). This approach allowed us to focus on the personal disposition that underlies individual differences in responses to stressors, a potential vulnerability factor. Table 4 shows these dynamic features and their correlation with the BDI score.

Regarding the dynamic features related to the cardiovascular system, the $\Delta(mHR)$ showed a significant negative correlation with BDI scores ($r=-.40$) both for the reactivity and recovery. This suggests that as BDI scores increase, indicating higher levels of depression, there is a decrease in heart rate dynamics. The HC group had higher $\Delta(mHR)$ compared to the MDD group. HRV features showed similar dynamics in both groups as expected from the similar patterns found in Table A1. Moreover, mHR dynamics were significantly negatively correlated with age, both reactivity and recovery values ($r=-.29$ and $r=-.25$ respectively), indicating a lower response to the cognitive stressor with age, although this association disappeared when controlling for BDI score. Lastly, the dynamics of the mPAT showed a significant positive correlation with BDI scores, indicating that higher depression levels are associated with lower vascular response to a stressor. Regarding the PAT, the recovery feature had lower interdependence with depression than the reactivity feature but was still significant. The HC group had a larger recovery, $\Delta(mPAT)_R^S$ (−20.89 ms) compared to the MDD group (−3.61 ms). In this case, the mPAT dynamics were positively correlated with age ($r=.31$ for reactivity and $r=.29$ for recovery), although they lost this association when controlling for depression.

TABLE 4 Dynamics of physiological features and BDI correlations.

	Reactivity ($\Delta(F)_B^S$)				Recovery ($\Delta(F)_R^S$)			
	HC	MDD	BDI		HC	MDD	BDI	
F			*r*	*p*			*r*	*p*
Electrocardiography								
mHR (m, std)	4.83 (7.21)	0.45 (3.83)**	−.40	<.01	8.49 (7.81)	3.08 (3.68)**	−.40	<.01
RMSSD (m, std)	0.06 (0.32)	0.03 (0.14)	.00	.98	−0.01 (0.37)	−0.10 (0.20)	.01	.96
SDNN (m, std)	0.23 (0.31)	0.13 (0.24)	−.06	.61	−0.05 (0.30)	−0.12 (0.30)	−.05	.67
PHF (m, std)	0.11 (0.24)	0.05 (0.14)	−.17	.13	−0.01 (0.26)	−0.03 (0.20)	−.09	.42
PLF (m, std)	0.23 (0.41)	0.10 (0.20)	−.20	.07	0.01 (0.45)	−0.12 (0.31)	−.12	.31
LF_HF (m, std)	0.15 (0.55)	0.13 (0.52)	−.04	.72	0.01 (0.68)	0.02 (0.79)	.00	.97
Photoplethysmography								
mHR (m, std)	4.28 (6.40)	0.46 (3.74)**	−.35	<.01	7.60 (6.85)	3.51 (4.46)**	−.40	<.01
RMSSD (m, std)	0.11 (0.29)	0.18 (0.25)	.22	.05	0.04 (0.31)	.03 (0.32)	0.03	.78
SDNN (m, std)	0.14 (0.26)	0.16 (0.24)	.13	.25	−0.08 (0.28)	−0.09 (0.28)	.10	.37
PHF (m, std)	0.07 (0.18)	0.06 (0.17)	−.17	.14	0.00 (0.20)	0.00 (0.19)	−.06	.58
PLF (m, std)	0.10 (0.35)	0.10 (0.32)	−.10	.37	−0.09 (0.45)	−0.07 (0.41)	.06	.57
LF_HF (m, std)	0.03 (0.55)	0.00 (0.45)	−.10	.39	−0.13 (0.69)	−0.05 (0.75)	.12	.31
Pulse arrival time								
mPAT (m, std)	−14.88 (29.57)	3.88 (9.72)**	.41	<.01	−20.89 (31.54)	−3.6 (12.06)**	.32	<.01
stdPAT (m, std)	0.28 (0.47)	0.24 (0.39)	−.06	.60	0.28 (0.41)	0.22 (0.40)	−.11	.35
Electrodermal activity								
EDASymp (m, std)	0.51 (0.84)	0.26 (1.07)	−.11	.34	0.57 (1.09)	−0.37 (1.02)**	−.28	.01
mTonic (m, std)	1.05 (0.88)	1.01 (0.93)	.01	.92	0.36 (1.27)	−0.18 (1.19)*	−.25	.02
stdTonic (m, std)	0.09 (0.54)	0.36 (0.56)*	.31	.01	0.10 (0.45)	−0.06 (0.42)	.02	.89
mPhasic (m, std)	0.14 (0.11)	0.05 (0.09)*	−.27	.02	0.18 (0.20)	−0.01 (0.17)**	−.29	.01
stdPhasic (m, std)	0.14 (0.27)	0.11 (0.39)	−.07	.53	0.13 (0.35)	−0.17 (0.37)**	−.27	.02

TABLE 4 (Continued)

	Reactivity ($\Delta(F)_B^S$)				Recovery ($\Delta(F)_R^S$)			
	HC	MDD	BDI		HC	MDD	BDI	
F			*r*	*p*			*r*	*p*
aucPhasic (m, std)	7.77 (5.91)	4.33 (4.02)**	−.24	.03	6.87 (7.84)	0.42 (5.37)**	−.26	.02
Respiration								
RR (m, std)	−0.06 (0.09)	−0.03 (0.06)	.13	.24	0.00 (0.09)	−0.02 (0.08)	−.15	.20
Pk (m, std)	−0.03 (0.04)	−0.02 (0.03)	.06	.61	−0.03 (0.05)	−0.02 (0.04)	−.12	.32
Skin temperature								
TFace_mGrad (m, std)	−0.01 (0.02)	0.00 (0.01)	.06	.58	0.00 (0.02)	0.00 (0.01)	−.05	.67
TFinger_mGrad (m, std)	−0.04 (0.04)	−0.02 (0.04)*	.30558	<.01	−0.04 (0.05)	−0.02 (0.03)**	.22	.05

Note: For a description of the abbreviated features, please refer to Table 1.

Abbreviation: F: feature.

*$p < .05$. **$p < .005$.

Concerning the electrical characteristics of the skin, recovery features also showed to be more relevant than the reactivity features. Most recovery features showed significant negative correlations with depression, including $\Delta(EDASymp)_R^S$ ($r=-.28$), $\Delta(medTonic)_R^S$ ($r=-.25$), $\Delta(medPhasic)_R^S$ ($r=-.29$), $\Delta(stdPhasic)_R^S$ ($r=-.27$) and $\Delta(aucPhasic)_R^S$ ($r=-.26$). These results reveal the divergent pattern in EDA seen in Figure 2; while the HC group returned to a neutral condition during the recovery stage, resulting in a positive recovery value, the MDD group remained equal or even showed higher activation, and hence, negative values. In relation to the reactivity features, only the $\Delta(medPhasic)_B^S$ ($r=-.27$) and the $\Delta(aucPhasic)_B^S$ ($r=-.24$) exhibited significant differences between both groups although with a slightly weaker correlation to the BDI.

Finally, while the dynamics of the respiratory features showed no significant relationship with depression, the skin temperature, specifically on the finger, showed significant positive correlations with the BDI scores, $\Delta(TFinger\_medGrad)_B^S$ ($r=.31$), $\Delta(TFinger\_medGrad)_R^S$ ($r=.22$). The HC group had lower means for both reactivity and recovery, $\Delta(TFinger\_medGrad)$ (−0.04°C) compared with the MDD group (−0.02°C), indicating that a smaller response to the stressor and lower thermoregulation adaptability are related to depression.

Additionally, we performed a Wilcoxon Signed-Rank Test to evaluate differences in physiological dynamics among the MDD groups with and without suicidal ideation, indicated by a score greater than 1 on item 9 of the BDI inventory. Only five subjects presented suicidal ideation, which limits the interpretive value of the results. Although the physiological dynamics did not show statistically significant differences, the results from this small group indicated that subjects with suicidal ideation exhibited a tendency toward higher overall HR and lower HRV, as indicated by the RMSSD and SDNN parameters, suggesting reduced resilience and behavioral flexibility.

3.5 | MDD model

To obtain the Major Depressive Disorder model, different sequential selections were made in the feature selection process to simplify further analyses. First, from all the dynamic features showing statistically significant linear dependencies, only one was selected. From this set of features, only those having a significant correlation with the BDI score were selected in a second step.

For the generalized linear modeling, cognitive and physiological variables were included in the final set of features. The model was adjusted by age, gender and BMI, and interactions were also considered. A backward stepwise approach was employed to select the predictor variables, utilizing random cross-validation with five partitions. It should be noted that the depression severity scores were transformed using the square root (log transformation was avoided due to zeros in the BDI scores) to increase the linear relationship with the autonomic reactivity indices.

The final regression models for Major Depressive Disorder (MDDm) considering the gender were the following:

$$\text{Women} \rightarrow \sqrt{\text{BDI}} = 6.59 + 0.05 \cdot \text{BMI} - 0.12 \cdot \text{STROOP\_WC}' - 0.03 \cdot \Delta(mHR)_R^S$$

$$\text{Men} \rightarrow \sqrt{\text{BDI}} = 12.95 - 0.18 \cdot \text{BMI} - 0.12 \cdot \text{STROOP\_WC}' - 1.04 \cdot \Delta(mHR)_R^S$$

At first glance, the MDDm seems to score higher depression severity in men than women, as can be seen by the intercepts in our models (12.95 and 6.59 respectively). However, upon closer inspection, both $\Delta(mHR)_R^S$ and BMI exert a multiplying factor that reduces the final score to a larger extent in men compared to women.

The coefficient of determination of the model was $r^2 = .45$, meaning that 45% of the variance of the depression severity score is explained by the predictors in the model. According to the standardized estimates of the derived model, gender is the most relevant factor in predicting BDI score. The second factor observed is the BMI. Within the demographic under study, it appears that men with higher depression tended to lose weight whereas it had the opposite effect on women. The cognitive component (STROOP_WC'), which is observed to be equivalent across both genders, significantly contributes to the model, independent of sociodemographic variables such as gender and BMI. Lastly, an elevated reactivity of mHR extracted from the ECG notably diminishes the model's score, with a more pronounced effect observed in men as compared to women. Even though age was included in the analysis, it was discarded during the selection process.

Figure 3a shows a scatter plot of the predicted MDDm values versus the actual BDI scores while Figure 3b displays a graph with the residuals of the regression model, which are the differences between the actual values and the predicted values. The graph provides a visual representation of how well the regression model fits the data, without any specific trend that may influence the results of the prediction. Finally, Figure 3c represents a graph with the ROC curves comparison of the BDI, the final MDD model, and the model without the physiological data, that is, $\Delta(mHR)_R^S$ to analyze its added value (marked as SCWT, green line on Figure 3c). The graph is a visual representation of the performance of these methods in detecting depression. While both proposed models showed a high area under the curve, 0.91 for the MDD model with the mHR dynamics, and 0.90 for the SCWT model without physiological data, correlation analysis indicated that the MDDm correlated better than the SCWT model with the BDI score ($r = .71$ and $r = .61$ respectively). The same MDD model, replacing the $\Delta(mHR)_R^S$ from the PPG instead of the ECG, gave similar results, with an area under the curve of 0.91. Finally, according to Youden's J statistic, we found an accuracy of 78%, a sensitivity of 97%, and a specificity of 59%.

4 | DISCUSSION

While Major Depressive Disorder is an increasingly prevalent problem worldwide, the lack of objective tools for diagnosing and monitoring this condition hinders early detection, treatment research, and the assessment of disease progression. This study aimed to design and validate a multiparametric model to assess depression severity, using the analysis of several physiological signals to evaluate differences in autonomic reactivity and recovery to a mild cognitive stressor and including relevant clinical data. From a practical point of view, this tool was designed to help clinicians find a more accurate diagnosis and follow-up therapeutic interventions while facilitating a common ground among professionals, thereby significantly improving patient care and outcomes.

4.1 | Cognitive stress protocol in MDD

Currently, the stress-diathesis theory constitutes the most empirically robust etiological model of various mental health disorders, including depression (Booij et al., 2013; Van Heeringen, 2012). The significant individual differences in the risk for disease following stress exposure highlight the moderating influence of inherent psychophysiological vulnerabilities.

The protocol employed in this study holds significant relevance due to its ability to assess stress dynamics in MDD patients and compare it to paired HC. Monitoring individuals during a baseline state provides a baseline measure of physiological signals, but at the same time, this measure can be altered by many confounding factors such as menstrual cycle phase, age, diet, physical activity, and sleep (Bari et al., 2020; Tada et al., 2017). However, the introduction of neuropsychological tests to induce stress followed by a recovery period allows for a dynamic assessment of an individual's physiological response to stressors and it offers insights into the specific vulnerability trait rather than their physical health transient condition (Bamert & Inauen, 2022). While there are discrepancies on whether reactivity or recovery features are better indicative of mental health disorders, our results are similar to the results in Kontaxis, Gil, et al. (2021) and Schiweck et al. (2022) which showed that changes in autonomic reactivity from the baseline to the stress stage have lower predictive power compared to changes from stress to recovery. These results suggest that the recovery may amplify differences in ANS regulation between individuals with MDD and HC. Focusing on the recovery period, we can obtain measurements that are less likely to be influenced by transient factors such as anxiety or nervousness related to being in a clinical setting.

Certainly, when executing protocols that involve monitoring physiological signals in response to cognitive stressors in future investigations, a multitude of factors warrant careful consideration. Empirical evidence suggests that determinants such as age, gender, BMI, educational attainment, medication regimen, and presence of comorbid conditions can modulate the outcomes, thereby potentially confounding the primary objective of the study (Bjelland et al., 2008; Nuño et al., 2021).

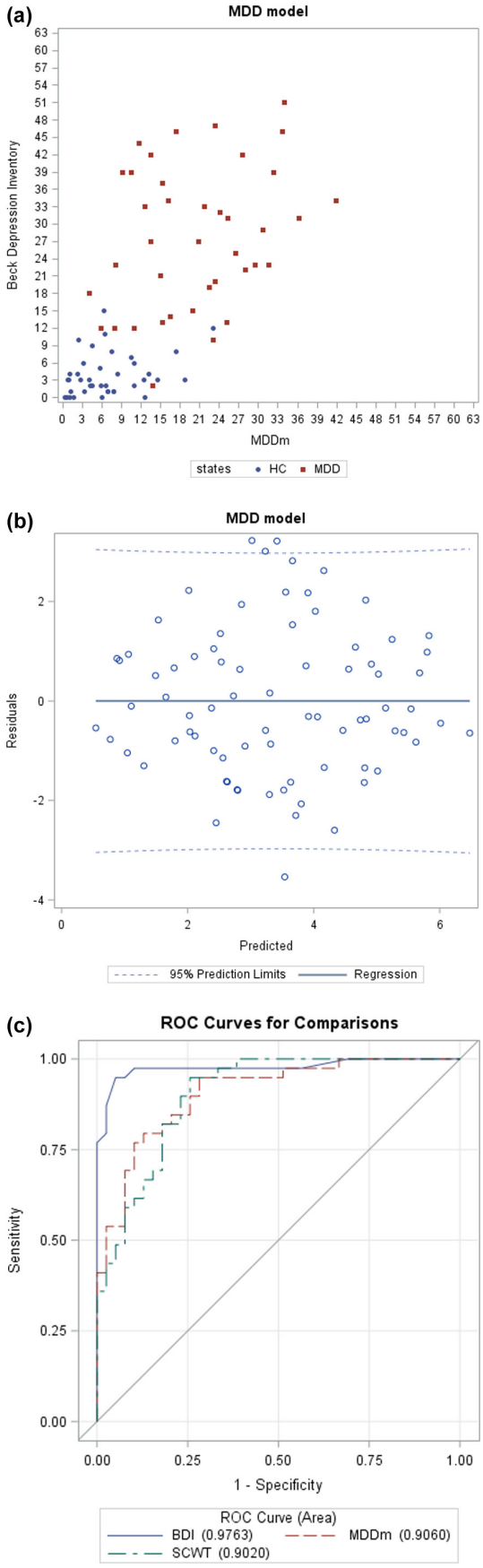

FIGURE 3 (a) MDD model performance. (b) MDDm residuals plot. (c) Accuracy assessment of the BDI, MDDm, and SCWT models.

In summary, by providing an objective measure of stress reactivity and recovery, this protocol could contribute to a more standardized and objective assessment of MDD, ensuring that all participants experience the same conditions, thereby reducing potential confounding variables.

4.2 | Predictive outcomes

The design of our study was developed to capture an objective and comprehensive evaluation of cognitive executive function and physiological stress reactivity and recovery. For this reason, the study was divided into two main predictive outcome groups: the neuropsychological assessment and the stress dynamics assessment.

4.2.1 | Neuropsychological assessment

Cognitive impairment has been reliably associated with MDD with effect sizes ranging from $d = 0.32–0.97$ (Snyder, 2013), not only in the acute phase of illness, but some reports indicate that this impairment might be long-lasting despite symptom reduction and recovery (Hammar & Årdal, 2009).

In the present study, the neuropsychological assessment was evaluated using two cognitive tests, mainly analyzing two specific components of executive function. First, we examined the ability to inhibit habitual behaviors (prepotent or automatic responses), using the Stroop Color and Word Test, and second, the Trail Making Test to assess the attention to switch between task goals.

Results from Table 3 show that patients with MDD completed fewer words in all the SCWT subscores and lasted longer to complete the TMT showing a deficit in processing speed and both inhibiting and switching components of executive function. These results are also in line with a systematic review (Nuño et al., 2021) that summarized evidence of worse overall performance on the SCWT and TMT among people with MDD, supporting a deficit in both cognitive effort and processing speed and suggesting that depression is not only characterized by psychomotor slowing but also involves a specific deficit in executive function. However, the results of the analysis were inconclusive due to sample heterogeneity regarding concomitant variables such as the subtype of depression, the age, the number of lifetime episodes of depression, and current medication. Therefore, we incorporated several gathered concomitant variables into our analysis, including the educational level, age, gender, and current medication status of the participants.

Regarding the educational level, our results suggest that education had a significant effect on the STROOP_WC

score, both on its own and when controlling for depression severity (BDI score) and the interaction between BDI and education. However, it is important to note that the effect of BDI on test performance did not significantly differ across levels of education. The TMT scores have been associated with education level as well as age and sex (Llinàs-Reglà et al., 2017). However, they do not report information about mental health. In our study, the analysis between education level, gender, and age did not show a significant relation with the TMT when controlling for depression, although age had a significant effect on STROOP_WC scores indicating that the age of the individual significantly predicted worse inhibition of automatic responses.

Finally, concerning antidepressive intake, the cognitive deficits observed in depressive patients are not solely attributed to the side effects of medication. This was supported by numerous studies that have identified significant executive function deficits in individuals with depression who are not on medication (Hinkelmann et al., 2009; Jia et al., 2019). Our study results also reinforce these findings since the effect of antidepressants on the cognitive test scores was not significant when controlled by depression severity, although the antidepressant effect on the ability to switch between task goals (TMT score) was closer to the significance level. However, it should be noted that we only had information on the number of antidepressant and anxiolytic medications, not the dosage.

In conclusion, these findings provide further evidence of the cognitive executive deficits associated with MDD. It also highlights the influence of factors such as education level and age on these cognitive abilities. However, more research is needed to fully understand the complex interplay between these variables.

4.2.2 | Stress dynamics assessment

Regarding the physiological assessment, another key area of the stress response, two systems are particularly significant within the physiological domain: the activity of the sympathetic nervous system, and the production of cortisol from the adrenal cortex, which is driven by the activity of the HPA axis (Gellman, 2020). Several meta-analyses conclude that cortisol activity in MDD is characterized by blunted stress reactivity (Burke et al., 2005; Zorn et al., 2017) but in this study, we investigated the stress reactivity mediated by sympathetic activation, leveraging the non-invasiveness of monitoring the ECG, PPG, EDA, Resp, and ST signals.

Our findings revealed several distinct reactivity patterns between MDD patients and HC. From the ECG and PPG signals, it was observed that the mHR increased during the stress stage and decreased during recovery in healthy individuals and that these dynamics were negatively correlated with age. However, this change was not as pronounced in the individuals with MDD, as seen in Figure 2, who also exhibited a higher mHR during the entire session. Several studies found similar trends (Brugnera et al., 2019; Carroll et al., 2007; Salomon et al., 2009; Schiweck et al., 2022; Siddi et al., 2023) but the negative correlation between depressive symptoms and HR reactivity remained significant even after accounting for potential confounding factors such as gender, age, occupational status, physical activity, smoking, task performance, and the use of antidepressant and antihypertensive medication. For these reasons, they discarded excessive reactivity as a mediator between depression and cardiovascular disease outcomes. However, on the other hand, impaired mHR recovery has been associated with an increased risk of cardiovascular disease (Panaite et al., 2015; Steptoe & Marmot, 2006) and it is also connected to several processes that are commonly present in depression, including anxiety, worry, and persistent negative thinking (Brosschot et al., 2006; Glynn et al., 2002).

Ultra-short-term classical HRV features (i.e., HRV analysis windows <5 min) showed similar patterns in the two groups, although overall reduced variability was observed in patients. Despite the fact that current literature is far from uniform regarding the validity of ultra-short-term recordings of HRV (Baek et al., 2015; Shaffer et al., 2020), the observed trend is in line with the review in Schiweck et al. (2019) where most studies also found lower mHR and PHF reactivity in individuals with depression compared with healthy controls. On the other hand, the importance of taking into account respiratory information into HRV spectral analysis when assessing stress response has been highlighted (Varon et al., 2019). Indices based on linear and nonlinear cardiorespiratory interactions published in previous studies utilizing the same sample have shown to provide better discrimination potential as MDD biomarkers than classical HRV features (Kontaxis et al., 2018; Kontaxis, Lazaro, et al., 2021). Discussion of the analysis and interpretation of HRV in the presence of confounding factors such as respiration is beyond the scope of this study but the interested reader is referred to other publications elaborating on the subject (Sornmo et al., 2024). Lastly, it is worth mentioning that the PPG features show slightly lower performance than their ECG counterparts when comparing groups, although they have a high correlation, making them appropriate for wearable applications. However, it must be pointed out that the PPG waveform may vary due to age, posture, ambient temperature, relaxation, and acclimatization (Allen, 2007; Allen et al., 2020), so future analysis on wearable signal processing and mental health prediction using this signal should consider these variables.

Contrary to the mHR pattern, the mPAT decreased with stress and increased during the recovery stage in controls, while it remained constant in patients, reinforcing low cardiovascular reactivity. This blunted response to stress could be indicative of allostatic overload or motivational dysregulations (Brugnera et al., 2019; Salomon et al., 2013; Schiweck et al., 2022), or it could also be that greater aortic stiffness is the origin of the hyporeactivity. According to Onete et al. (2018), greater aortic stiffness measured as the carotid to femoral pulse wave velocity was associated with MDD and the presence of depressive symptoms in men aged 60 years or younger, and to a lesser extent in women of the same age. However, this association disappeared in men and women older than 60 years, probably overshadowed in later life, where other factors, such as cardiovascular disease and atherosclerosis, could have a bigger impact on the development of depression than aortic stiffness. In any case, both hypotheses may not be mutually exclusive, and a combination of factors likely contributed to the observed results in MDD patients and support the contention that blunted reactivity is an indicator of risk for poor psychological and physical health. Given the cross-sectional design of the current study, drawing robust conclusions on this matter is challenging.

Blunted peripheral temperature reactivity was also a hallmark of MDD. HC exhibited a greater decreasing gradient of the finger temperature during the stress compared to the baseline and recovery stages, while the patient group maintained a more uniform temperature. Limited literature evaluates the temperature dynamics in response to a stressor and its relation to MDD. We only found one study that supports that depression reactivity was negatively correlated with mHR and ST (Lin et al., 2011). Sarlon et al. (2021) evaluated ST in MDD patients, using a recall of unpleasant stressful experiences of a medium intensity as a stressor, and they also found no between-stage differences, although they could not relate the lack of reactivity to depression due to the lack of HC group comparison. Lastly, skin temperature and sleep time-related features were the most significant features of a machine-learning model using real-life wearable data (Tazawa et al., 2020).

In contrast to the previous dynamic patterns, the EDA in MDD patients increased during the stress stage in coherence with the healthy participants, though to a lesser extent, but this activation was maintained throughout the session or even continued to increase during the recovery stage. EDA has been found to vary linearly with emotional arousal and can be used not only to classify different emotional states such as anger and fear but also to detect stress levels while performing a task (Boucsein, 2012). Several systematic reviews (Ettore et al., 2023; Sarchiapone et al., 2018) discuss the literature up to date related to EDA hyporeactivity and its relevance as a biomarker for depression, recurring episodes (Carli et al., 2022), and suicidal behavior (Litwińska-Bołtuć et al., 2021; Thorell, 1987) although few explore the recovery dynamics. Our findings show that EDA recovery features were negatively associated with depression indicating that even though patients' reaction is weaker, they might require more time to return to their baseline levels. In Kim et al. (2018, 2019)), recovery features were also predictive of depression even though their focus on the features was more as input data for a support vector machine model.

Finally, the respiratory signal was the least informative data in predicting depression. Features exhibited similar results in all participants except for the respiratory frequency during the cognitive tests, where patients showed significantly higher values. However, results should be analyzed carefully since the SCWT uses speech which can alter respiration patterns. In Khandoker et al. (2018), MDD patients who had suicidal ideation were observed to have a faster breathing rate and a reduced amplitude of respiratory sinus arrhythmia compared to those patients without suicidal ideation. In any case, the measurements were taken in a baseline situation without a stressor agent, and in a subsequent study with an expanded database (Zitouni et al., 2022), they found higher model accuracies with cardiac features instead of respiratory features.

The study also examined several confounding factors under study such as antidepressant intake, tobacco use, coffee, and alcohol consumption. Despite these considerations, the results demonstrated no significant impact of these variables on the physiological features under investigation. Notably, BMI and gender were identified as influencing factors for HR and EDA features, corroborating previous literature (Aldosky, 2019; Herhaus et al., 2023; Quer et al., 2020). Given their reported association with depression, these factors were integrated into our model. It is recommended that future studies with larger sample sizes account for these potential confounders.

Overall, these findings indicate that a greater physiological flexibility to cope with the cognitive stressor is related to better mental health. Therefore, stress dynamic response provides valuable insights into the physiological correlates of depression and highlights potential markers for its objective assessment.

4.3 | Non-invasive MDD model

To conclude the discussion, the main aim of our study was to find a comprehensive model for depression severity measurement that could be later used as an objective reference for clinicians offering synthesized information on both the response corresponding to the cognition as well as of the somatic response to a stressful stimulus.

Our protocol included the synchronous physiological monitoring of reactivity and recovery to extract several non-invasive stress biomarkers that are altered in depression. However, after an exhaustive literature analysis and results comparison (please refer to Sections 3.4 and 4.1), the recovery features were selected for modeling purposes. In particular, the stepwise backward selection process selected heart rate as the most relevant for the task at hand, and accordingly, it is the most prevalent in literature (Brugnera et al., 2019; Carroll et al., 2007; Salomon et al., 2009; Schiweck et al., 2019, 2022). It is pertinent to highlight that the EDA features and their interplay with gender were on the verge of attaining significance for their incorporation into the final model. Nonetheless, within the confines of our reduced sample size, they did not contribute substantially to the enhancement of the model's predictive precision.

The resulting model (Table 5) showed that BMI appears to have a protective influence in men, while it intensifies the severity of depression in women. Similar results were published in a systematic review (Jung et al., 2017), where the pooled odds ratio showed a protective effect of being overweight in men, whereas it showed an increased likelihood of depression in women. These findings could also indicate different coping mechanisms for depressed men and women, the former being appetite loss while the latter would be overeating or inactivity as a result of depression leading to weight gain. Despite our intention to assemble a group with comparable BMI, we observed a modest relation with depression that could be even more noticeable in a sample selected at random. Therefore, adding these demographic or medical history variables as context enhances the personalization of the model to determine a more accurate mental state of the individual (Hatton et al., 2019; Rahmani et al., 2022).

The cognitive flexibility component (STROOP_WC'), which is observed to be equivalent across both genders, significantly contributes to the model, independent of sociodemographic variables such as gender and BMI.

Lastly, a large mHR recovery ($\Delta(mHR)_R^S$), as a result of a good recuperation after an elevated reactivity, notably diminishes the model's score (i.e., less depression severity), with a more pronounced effect observed in men as compared to women. Similarly, Carroll et al. (2007) also found that men tended to be more reactive than women. Furthermore, this feature, extracted from the PPG signal, showed similar model performance and, given its computational simplicity, it could be easily measured using wearable technology.

There are several updated literature reviews on the state of the art of multimodal physiological models for the detection and classification of depression severity (Ahmed et al., 2022; Zitouni et al., 2022). The main physiological signals used are the ECG (or PPG), EDA, and electroencephalography (EEG). While several EEG models achieve high accuracies above 90%, they involve a small number of individuals (Zitouni et al., 2022). Moreover, this technology requires special equipment preparation and it is more invasive than the others, making them less promising and/or practical for a broader application.

Zitouni et al. (2022) developed a model to classify MDD using a multi-modal approach combining ECG, PPG, and Resp data achieving 96.6% accuracy. In machine learning models, the goal is prediction, and it is assumed that there is sufficient data to fit models with high complexity and many parameters (Rajput et al., 2023; Wilson Van Voorhis & Morgan, 2007). However, most of the published models assessing depression severity use machine learning algorithms without appropriate sample sizes which may lead to a lack of reproducibility of their results in other populations. In our approach using statistical models, the goal was to understand the causes of the observed changes and provide measures of variable importance, which can be useful in understanding the relationships between variables.

Ultimately, the scientific achievement of this research contributes to the objective assessment of MDD, and its potential for replication in a mHealth environment thanks to wearable technology and smartphones reinforces its relevance in the field (Bajaj et al., 2013; Tazawa et al., 2020; Yang & Kershaw, 2022).

4.4 | Limitations

The main limitation of the current study is the small sample size. Moreover, validation in independent datasets or analysis of depression subtypes were not performed. However, to mitigate this point, five-fold cross-validation was used to reduce possible overfitting.

Another limitation of our study is the use of two artificial cognitive stressors during the experimental setting which may challenge the replication of the results or comparison with other studies. Moreover, the differences between these two tests were not explored as they were out of the scope of this study. In any case, the results should be cautiously interpreted, and they may vary in a real-life setting with other cognitive stressors (e.g., work deadlines, performance evaluations, and university exams).

Additionally, to assess the evolution of the stress dynamics during the cognitive tests, ultra-short-term analysis for the physiological features was a trade-off between usability and reliability and should also be more carefully studied. Moreover, linear and nonlinear cardiorespiratory

TABLE 5 MDD model parameter estimates.

						Cross validation estimates				
Parameter	**Est**	**Std Est**	**SE**	***t* value**	**Pr > |*t*|**	**1**	**2**	**3**	**4**	**5**
Intercept	12.95	0.00	2.02	6.42	<0.0001	13.31	13.12	11.20	13.27	14.08
Gender female	−6.36	−1.51	2.14	−2.98	<0.01	−6.42	−5.82	−5.06	−7.12	−7.75
Gender male	0.00	0.00				0.00	0.00	0.00	0.00	0.00
BMI	−0.18	−0.43	0.06	−2.72	<0.01	−0.18	−0.18	−0.12	−0.20	−0.22
$\Delta(\text{mHR})_R^S$	−1.04	−0.50	0.34	−3.06	<0.01	−1.09	−1.12	−0.75	−1.05	−1.26
BMI*female	0.23	1.46	0.08	2.84	<0.01	0.23	0.21	0.19	0.26	0.28
BMI*male	0.00	0.00				0.00	0.00	0.00	0.00	0.00
STROOP_WC'	−0.12	−0.54	0.02	−5.25	<0.0001	−0.12	−0.12	−0.11	−0.11	−0.12
$\Delta(\text{mHR})_R^S$*female	1.01	0.39	0.40	2.53	0.01	0.91	1.04	0.72	1.37	1.24
$\Delta(\text{mHR})_R^S$*male	0.00	0.00				0.00	0.00	0.00	0.00	0.00

Abbreviations: $\Delta(\text{mHR})_R^S$, recovery dynamic feature of the mean heart rate; Est, estimate; SE, standard error; Std Est, standardized estimate; STROOP_WC', estimated stroop word-color cognitive subtask.

interactions inclusion into the analysis could greatly improve the reliability of the HRV measurement.

Finally, only linear relations between the predictor variables and the MDD reference (i.e., BDI scores) were analyzed in this approach, although including non-linear relations could contribute to a more accurate measure of the stress reactivity and MDD severity.

5 | CONCLUSION AND FUTURE DIRECTIONS

In conclusion, this study offers a comprehensive approach to understanding MDD by assessing cognition and stress dynamics. Moreover, we propose an objective tool to assess depression severity using non-invasive physiological and cognitive measures and also including biometric and demographic variables to increase personalization. This tool, in combination with traditional methods, provides a more holistic view of a person's mental health. It can be administered repeatedly reflecting improvement or worsening of depression in response to treatment, and it could give a common ground for healthcare providers.

Future research should continue to refine and validate this protocol for broader application, with larger and diverse samples, investigating the variability among depression subtypes, and including prospective studies to differentiate whether stress dynamics dysfunction is a consequence or correlate of depression or a marker of vulnerability to future depression.

AUTHOR CONTRIBUTIONS

Esther García Pagès: Conceptualization; data curation; formal analysis; investigation; methodology; project administration; validation; visualization; writing – original draft; writing – review and editing. **Spyridon Kontaxis:** Investigation; resources; software; writing – review and editing. **Sara Siddi:** Conceptualization; investigation; methodology; project administration; resources; writing – review and editing. **Mar Posadas-de Miguel:** Investigation; methodology; project administration; writing – review and editing. **Concepción de la Cámara:** Funding acquisition; resources; supervision; writing – review and editing. **Maria Luisa Bernal:** Investigation; methodology; writing – review and editing. **Thais Castro Ribeiro:** Investigation; visualization; writing – review and editing. **Pablo Laguna:** Conceptualization; funding acquisition; project administration; resources; writing – review and editing. **Llorenç Badiella:** Data curation; supervision; validation; writing – review and editing. **Raquel Bailón:** Conceptualization; funding acquisition; project administration; resources; supervision; writing – review and editing. **Josep Maria Haro:** Funding acquisition; project administration; resources; supervision; writing – review and editing. **Jordi Aguiló:** Conceptualization; funding acquisition; project administration; resources; supervision; writing – review and editing.

FUNDING INFORMATION

This research was supported by CIBER—Consorcio Centro de Investigación Biomédica en Red- (CB06/01/0049), Instituto de Salud Carlos III and the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 115902 and the European Union's Horizon 2020 research and innovation program, www.imi.europa.eu. This work was also funded by TED2021-131106B-I00 (Ministerio de Ciencia e Innovación and European Social Fund), Spain, and by European Social Fund and

Aragón Government, Spain through BSICoS group, Spain T39_23R. Llorenç Badiella received funding from the Spanish Ministry of Science, Innovation and Universities (grant RTI2018-096072-B-I00).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data available on request due to privacy/ethical restrictions.

ORCID

Esther García Pagès https://orcid.org/0000-0002-9225-845X
Thais Castro Ribeiro https://orcid.org/0000-0003-2536-7194

REFERENCES

Adolph, D., Teismann, T., Forkmann, T., Wannemüller, A., & Margraf, J. (2018). High frequency heart rate variability: Evidence for a transdiagnostic association with suicide ideation. *Biological Psychology*, *138*, 165–171. https://doi.org/10.1016/J.BIOPSYCHO.2018.09.006

Ahmed, T., Qassem, M., & Kyriacou, P. A. (2022). Physiological monitoring of stress and major depression: A review of the current monitoring techniques and considerations for the future. *Biomedical Signal Processing and Control*, *75*, 103591. https://doi.org/10.1016/J.BSPC.2022.103591

Aldosky, H. Y. Y. (2019). Impact of obesity and gender differences on electrodermal activities. *General Physiology and Biophysics*, *38*(6), 513–518. https://doi.org/10.4149/GPB_2019036

Allen, J. (2007). Photoplethysmography and its application in clinical physiological measurement. *Physiological Measurement*, *28*(3), R1–R39. https://doi.org/10.1088/0967-3334/28/3/R01

Allen, J., O'Sullivan, J., Stansby, G., & Murray, A. (2020). Age-related changes in pulse risetime measured by multi-site photoplethysmography. *Physiological Measurement*, *41*(7), 74001. https://doi.org/10.1088/1361-6579/AB9B67

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Publishing, Inc. https://doi.org/10.1176/APPI.BOOKS.9780890425787

Arias-de la Torre, J., Vilagut, G., Ronaldson, A., Serrano-Blanco, A., Martín, V., Peters, M., Valderas, J. M., Dregan, A., & Alonso, J. (2021). Prevalence and variability of current depressive disorder in 27 European countries: A population-based study. *The Lancet. Public Health*, *6*(10), e729–e738. https://doi.org/10.1016/S2468-2667(21)00047-5

Armañac, P., Kontaxis, S., Lazaro, J., Laguna, P., Bailón, R., & Gil, E. (2019). Cardiovascular changes induced by acute emotional stress estimated from the pulse transit time difference. *2019 Computing in Cardiology Conference (CinC)*, *46*, 45. https://doi.org/10.22489/CINC.2019.383

Arza, A., Garzón-Rey, J. M., Lázaro, J., Gil, E., Lopez-Anton, R., de la Camara, C., Laguna, P., Bailon, R., & Aguiló, J. (2018). Measuring acute stress response through physiological signals: Towards a quantitative assessment of stress. *Medical & Biological Engineering & Computing*, *1–17*, 271–287. https://doi.org/10.1007/s11517-018-1879-z

Baek, H. J., Cho, C. H., Cho, J., & Woo, J. M. (2015). Reliability of ultra-short-term analysis as a surrogate of standard 5-min analysis of heart rate variability. *Telemedicine Journal and E-Health: The Official Journal of the American Telemedicine Association*, *21*(5), 404–414. https://doi.org/10.1089/TMJ.2014.0104

Bajaj, J. S., Thacker, L. R., Heuman, D. M., Fuchs, M., Sterling, R. K., Sanyal, A. J., Puri, P., Siddiqui, M. S., Stravitz, R. T., Bouneva, I., Luketic, V., Noble, N., White, M. B., Monteith, P., Unser, A., & Wade, J. B. (2013). The Stroop smartphone application is a short and valid method to screen for minimal hepatic encephalopathy. *Hepatology (Baltimore, Md.)*, *58*(3), 1122–1132. https://doi.org/10.1002/HEP.26309

Bamert, M., & Inauen, J. (2022). Physiological stress reactivity and recovery: Some laboratory results transfer to daily life. *Frontiers in Psychology*, *13*, 943065. https://doi.org/10.3389/FPSYG.2022.943065

Bari, D. S., Yacoob Aldosky, H. Y., & Martinsen, Ø. G. (2020). Simultaneous measurement of electrodermal activity components correlated with age-related differences. *Journal of Biological Physics*, *46*(2), 177–188. https://doi.org/10.1007/S10867-020-09547-4

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. F. (1996). Comparison of Beck depression inventories -IA and -II in psychiatric outpatients. *Journal of Personality Assessment*, *67*(3), 588–597. https://doi.org/10.1207/S15327752JPA6703_13

Bjelland, I., Krokstad, S., Mykletun, A., Dahl, A. A., Tell, G. S., & Tambs, K. (2008). Does a higher educational level protect against anxiety and depression? The HUNT study. *Social Science & Medicine*, *66*(6), 1334–1345. https://doi.org/10.1016/J.SOCSCIMED.2007.12.019

Booij, L., Wang, D., Lévesque, M. L., Tremblay, R. E., & Szyf, M. (2013). Looking beyond the DNA sequence: The relevance of DNA methylation processes for the stress-diathesis model of depression. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *368*(1615), 20120251. https://doi.org/10.1098/RSTB.2012.0251

Boucsein, W. (2012). *Electrodermal activity*. Springer US. https://doi.org/10.1007/978-1-4614-1126-0

Bowie, C. R., & Harvey, P. D. (2006). Administration and interpretation of the trail making test. *Nature Protocols*, *1*(5), 2277–2281. https://doi.org/10.1038/NPROT.2006.390

Brosschot, J. F., Gerin, W., & Thayer, J. F. (2006). The perseverative cognition hypothesis: A review of worry, prolonged stress-related physiological activation, and health. *Journal of Psychosomatic Research*, *60*(2), 113–124. https://doi.org/10.1016/J.JPSYCHORES.2005.06.074

Brugnera, A., Zarbo, C., Tarvainen, M. P., Carlucci, S., Tasca, G. A., Adorni, R., Auteri, A., & Compare, A. (2019). Higher levels of depressive symptoms are associated with increased resting-state heart rate variability and blunted reactivity to a laboratory stress task among healthy adults. *Applied Psychophysiology and Biofeedback*, *44*(3), 221–234. https://doi.org/10.1007/S10484-019-09437-Z

Burke, H. M., Davis, M. C., Otte, C., & Mohr, D. C. (2005). Depression and cortisol responses to psychological stress: A meta-analysis. *Psychoneuroendocrinology*, *30*(9), 846–856. https://doi.org/10.1016/j.psyneuen.2005.02.010

Carli, V., Hadlaczky, G., Petros, N. G., Iosue, M., Zeppegno, P., Gramaglia, C., Amore, M., Baca-Garcia, E., Batra, A., Cosman, D., Courtet, P., Di Sciascio, G., Ekstrand, J., Galfalvy, H., Gusmão, R., Jesus, C., Heitor, M. J., Constante, M., Rad, P. M., … Sarchiapone, M. (2022). A naturalistic, European multicenter clinical study of Electrodermal reactivity and suicide risk among patients with depression. *Frontiers in Psychiatry*, *12*, 765128. https://doi.org/10.3389/FPSYT.2021.765128/FULL

Carroll, D., Phillips, A. C., Hunt, K., & Der, G. (2007). Symptoms of depression and cardiovascular reactions to acute psychological stress: Evidence from a population study. *Biological Psychology*, *75*(1), 68–74. https://doi.org/10.1016/J.BIOPSYCHO.2006.12.002

Cherniack, N. S. (1990). The central nervous system and respiratory muscle coordination. *Chest*, *97*(3 SUPPL), 52S–57S. https://doi.org/10.1378/chest.97.3_Supplement.52S-a

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, *24*(4), 385–396.

Deshmukh, C., Jindal, G., Bagal, U., & Nagare, G. (2022). Pulse arrival time: Measurement and clinical applications. *MGM Journal of Medical Sciences*, *9*(1), 103–111. https://doi.org/10.4103/MGMJ.MGMJ_23_22

Ettore, E., Müller, P., Hinze, J., Benoit, M., Giordana, B., Postin, D., Lecomte, A., Lindsay, H., Robert, P., & König, A. (2023). Digital phenotyping for differential diagnosis of major depressive episode: Narrative review. *JMIR Mental Health*, *10*, e37225. https://doi.org/10.2196/37225

Ferrari, A. (2022). Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: A systematic analysis for the global burden of disease study 2019. *The Lancet Psychiatry*, *9*(2), 137–150. https://doi.org/10.1016/S2215-0366(21)00395-3

Gallagher, D., Breckenridge, J., Steinmetz, J., & Thompson, L. (1983). The Beck depression inventory and research diagnostic criteria: Congruence in an older population. *Journal of Consulting and Clinical Psychology*, *51*(6), 945–946. https://doi.org/10.1037/0022-006X.51.6.945

García Pagès, E., Arza, A., Lazaro, J., Puig, C., Castro, T., Ottaviano, M., Arredondo, M. T., Bernal, M. L., López-Antón, R., Cámara, C. D., Gil, E., Laguna, P., Bailón, R., Aguiló, J., & Garzón-Rey, J. M. (2023). Psychosomatic response to acute emotional stress in healthy students. *Frontiers in Physiology*, *13*. https://doi.org/10.3389/FPHYS.2022.960118

Garzón-Rey, J. M., & Aguiló Llobet, J. (2017). *Psychosomatic approach to stress measurement*. Autonomous University of Barcelona.

Gellman, M. D. (2020). Behavioral medicine. *Encyclopedia of Behavioral Medicine*, 223–226. https://doi.org/10.1007/978-3-030-39903-0_1660

Gil Herrando, E. (2009). Análisis multimodal de las señales fotopletismográfica de pulso y electrocardiográfica para el diagnóstico del síndrome de apnea obstructiva durante el sueño en niños. Instituto de Investigación en Ingeniería de Aragón, I3A.

Glynn, L. M., Christenfeld, N., & Gerin, W. (2002). The role of rumination in recovery from reactivity: Cardiovascular consequences of emotional states. *Psychosomatic Medicine*, *64*(5), 714–726. https://doi.org/10.1097/01.PSY.0000031574.42041.23

Golden, C. J. (2020). *STROOP. Test de Colores y Palabras – Edición Revisada (B. Ruiz-Fernández, T. Luque y F. Sánchez-Sánchez, adaptadores)*. TEA Ediciones.

Greaney, J. L., Kenney, W. L., & Alexander, L. M. (2016). Sympathetic regulation during thermal stress in human aging and disease. *Autonomic Neuroscience: Basic & Clinical*, *196*, 81–90. https://doi.org/10.1016/J.AUTNEU.2015.11.002

Greco, A., Valenza, G., Lanata, A., Scilingo, E. P., & Citi, L. (2016). cvxEDA: A convex optimization approach to Electrodermal activity processing. *IEEE Transactions on Bio-Medical Engineering*, *63*(4), 797–804. https://doi.org/10.1109/TBME.2015.2474131

Guillén Riquelme, A., & Buela Casal, G. (2011). Actualización psicométrica y funcionamiento diferencial de los ítems en el State Trait Anxiety Inventory (STAI). *Psicothema*, *23*(3), 510–515. https://dialnet.unirioja.es/servlet/articulo?codigo=3687191&info=resumen&idioma=SPA

Guy, W. (1976). *ECDEU assessment manual for psychopharmacology*. Department of Health, Education, and Welfare.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, *23*(1), 56–62. https://doi.org/10.1136/JNNP.23.1.56

Hammar, Å., & Årdal, G. (2009). Cognitive functioning in major depression – A summary. *Frontiers in Human Neuroscience*, *3*(SEP), 26. https://doi.org/10.3389/NEURO.09.026.2009

Hatton, C. M., Paton, L. W., McMillan, D., Cussens, J., Gilbody, S., & Tiffin, P. A. (2019). Predicting persistent depressive symptoms in older adults: A machine learning approach to personalised mental healthcare. *Journal of Affective Disorders*, *246*, 857–860. https://doi.org/10.1016/J.JAD.2018.12.095

Herhaus, B., Ghassabei, S., & Petrowski, K. (2023). Obesity: Heart rate variability during standardized psychosocial stress induction. *Biological Psychology*, *178*, 108509. https://doi.org/10.1016/J.BIOPSYCHO.2023.108509

Hinkelmann, K., Moritz, S., Botzenhardt, J., Riedesel, K., Wiedemann, K., Kellner, M., & Otte, C. (2009). Cognitive impairment in major depression: Association with salivary cortisol. *Biological Psychiatry*, *66*(9), 879–885. https://doi.org/10.1016/J.BIOPSYCH.2009.06.023

Hu, M. X., Lamers, F., De Geus, E. J. C., & Penninx, B. W. J. H. (2016). Differential autonomic nervous system reactivity in depression and anxiety during stress depending on type of stressor. *Psychosomatic Medicine*, *78*(5), 562–572. https://doi.org/10.1097/PSY.0000000000000313

Jia, Q. F., Chen, P., Zhu, H. L., Chen, S. S., Gu, X. C., Yin, X. Y., Wu, Y. H., Yin, G. Z., & Hui, L. (2019). Cognitive impairments in first-episode drug-Naïve versus medicated depressive patients: RBANS in a Chinese population. *The Psychiatric Quarterly*, *90*(3), 471–480. https://doi.org/10.1007/S11126-019-09641-4

Jung, S. J., Woo, H. T., Cho, S., Park, K., Jeong, S., Lee, Y. J., Kang, D., & Shin, A. (2017). Association between body size, weight change and depression: Systematic review and meta-analysis. *The British Journal of Psychiatry: The Journal of Mental Science*, *211*(1), 14–21. https://doi.org/10.1192/BJP.BP.116.186726

Khandoker, A. H., Widatalla, N., Jelinek, H., Niizeki, K., & Hadjileontiadis, L. (2018). Incoherent synchronization between resting state respiratory sinus arrhythmia and respiratory movement in depressed patients with suicidal ideation. *Computing in Cardiology*, *45*, 1–4. https://doi.org/10.22489/CINC.2018.173

Kiecolt-Glaser, J. K., Renna, M. E., Shrout, M. R., & Madison, A. A. (2020). Stress reactivity: What pushes us higher, faster, and longer – And why it matters. *Current Directions in Psychological*

Science, *29*(5), 492–498. https://doi.org/10.1177/0963721420949521

Kim, A. Y., Jang, E. H., Choi, K. W., Jeon, H. J., Byun, S., Sim, J. Y., Choi, J. H., & Yu, H. Y. (2019). Skin conductance responses in major depressive disorder (MDD) under mental arithmetic stress. *PLoS One*, *14*(4), 1–13. https://doi.org/10.1371/JOURNAL.PONE.0213140

Kim, A. Y., Jang, E. H., Kim, S., Choi, K. W., Jeon, H. J., Yu, H. Y., & Byun, S. (2018). Automatic detection of major depressive disorder using electrodermal activity. *Scientific Reports*, *8*(1), 17030. https://doi.org/10.1038/S41598-018-35147-3

Kontaxis, S., Gil, E., Marozas, V., Lazaro, J., Garcia, E., Posadas-De Miguel, M., Siddi, S., Bernal, M. L., Aguilo, J., Haro, J. M., De La Camara, C., Laguna, P., & Bailon, R. (2021). Photoplethysmographic waveform analysis for autonomic reactivity assessment in depression. *IEEE Transactions on Bio-Medical Engineering*, *68*(4), 1273–1281. https://doi.org/10.1109/TBME.2020.3025908

Kontaxis, S., Lazaro, J., Gil, E., Laguna, P., & Bailon, R. (2021). The added value of nonlinear cardiorespiratory coupling indices in the assessment of depression. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, *2021*, 5473–5476. https://doi.org/10.1109/EMBC46164.2021.9631096

Kontaxis, S., Orini, M., Gil, E., De Miguel, M. P., Bernal, M. L., Aguilo, J., De La Camara, C., Laguna, P., & Bailon, R. (2018). Heart rate variability analysis guided by respiration in major depressive disorder. *Computing in Cardiology*, *45*, 1. https://doi.org/10.22489/CINC.2018.319

Lázaro, J., Gil, E., Vergara, J. M., & Laguna, P. (2014). Pulse rate variability analysis for discrimination of sleep-apnea-related decreases in the amplitude fluctuations of pulse photoplethysmographic signal in children. *IEEE Journal of Biomedical and Health Informatics*, *18*(1), 240–246. https://doi.org/10.1109/JBHI.2013.2267096

Lesage, F. X., Berjot, S., & Deschamps, F. (2012). Clinical stress assessment using a visual analogue scale. *Occupational Medicine (Oxford, England)*, *62*(8), 600–605. https://doi.org/10.1093/OCCMED/KQS140

Lin, H. P., Lin, H. Y., Lin, W. L., & Huang, A. C. W. (2011). Effects of stress, depression, and their interaction on heart rate, skin conductance, finger temperature, and respiratory rate: Sympathetic-parasympathetic hypothesis of stress and depression. *Journal of Clinical Psychology*, *67*(10), 1080–1091. https://doi.org/10.1002/JCLP.20833

Litwińska-Bołtuć, M., Święcicki, Ł., Spreco, A., & Timpka, T. (2021). Clinical effectiveness of the electrodermal orienting reactivity test for evaluating relapse and recurrence risk in patients hospitalized for depression. *BMC Psychiatry*, *21*(1), 88. https://doi.org/10.1186/S12888-021-03088-3

Llinàs-Reglà, J., Vilalta-Franch, J., López-Pousa, S., Calvó-Perxas, L., Torrents Rodas, D., & Garre-Olmo, J. (2017). The trail making test. *Assessment*, *24*(2), 183–196. https://doi.org/10.1177/1073191115602552

Ma, S., Yang, J., Yang, B., Kang, L., Wang, P., Zhang, N., Wang, W., Zong, X., Wang, Y., Bai, H., Guo, Q., Yao, L., Fang, L., & Liu, Z. (2021). The patient health Questionnaire-9 vs. the Hamilton rating scale for depression in assessing major depressive disorder. *Frontiers in Psychiatry*, *12*, 747139. https://doi.org/10.3389/FPSYT.2021.747139/BIBTEX

Malik, M., Camm, A. J., Bigger, J. T., Breithardt, G., Cerutti, S., Cohen, R. J., Coumel, P., Fallen, E. L., Kennedy, H. L., Kleiger, R. E., Lombardi, F., Malliani, A., Moss, A. J., Rottman, J. N., Schmidt, G., Schwartz, P. J., & Singer, D. H. (1996). Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. In. *European Heart Journal*, *17*(3), 354–381. https://doi.org/10.1093/oxfordjournals.eurheartj.a014868

Martínez, J. P., Almeida, R., Olmos, S., Rocha, A. P., & Laguna, P. (2004). A wavelet-based ECG delineator: Evaluation on standard databases. *IEEE Transactions on Bio-Medical Engineering*, *51*(4), 570–581. https://doi.org/10.1109/TBME.2003.821031

Matcham, F., Barattieri di San Pietro, C., Bulgari, V., de Girolamo, G., Dobson, R., Eriksson, H., Folarin, A. A., Haro, J. M., Kerz, M., Lamers, F., Li, Q., Manyakov, N. V., Mohr, D. C., Myin-Germeys, I., Narayan, V., Bwjh, P., Ranjan, Y., Rashid, Z., Rintala, A., … RADAR-CNS consortium. (2019). Remote assessment of disease and relapse in major depressive disorder (RADAR-MDD): A multi-centre prospective cohort study protocol. *BMC Psychiatry*, *19*(1), 72. https://doi.org/10.1186/s12888-019-2049-z

Matcham, F., Simblett, S. K., Leightley, D., Dalby, M., Siddi, S., Haro, J. M., Lamers, F., Penninx, B. W. H. J., Bruce, S., Nica, R., Zormpas, S., Gilpin, G., White, K. M., Oetzmann, C., Annas, P., Brasen, J. C., Narayan, V. A., Hotopf, M., & Wykes, T. (2023). The association between persistent cognitive difficulties and depression and functional outcomes in people with major depressive disorder. *Psychological Medicine*, *53*(13), 6334–6344. https://doi.org/10.1017/S0033291722003671

Mateo, J., & Laguna, P. (2003). Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal. *IEEE Transactions on Bio-Medical Engineering*, *50*(3), 334–343. https://doi.org/10.1109/TBME.2003.808831

Naslund, J. A., Marsch, L. A., McHugo, G. J., & Bartels, S. J. (2015). Emerging mHealth and eHealth interventions for serious mental illness: A review of the literature. *Journal of Mental Health (Abingdon, England)*, *24*(5), 321–332. https://doi.org/10.3109/09638237.2015.1019054

Nuño, L., Gómez-Benito, J., Carmona, V. R., & Pino, O. (2021). A systematic review of executive function and information processing speed in major depression disorder. *Brain Sciences*, *11*(2), 1–18. https://doi.org/10.3390/BRAINSCI11020147

Onete, V., Henry, R. M., Sep, S. J. S., Koster, A., Van Der Kallen, C. J., Dagnelie, P. C., Schaper, N., Köhler, S., Reesink, K., Stehouwer, C. D. A., & Schram, M. T. (2018). Arterial stiffness is associated with depression in middle-aged men — The Maastricht study. *Journal of Psychiatry & Neuroscience: JPN*, *43*(2), 111–119. https://doi.org/10.1503/JPN.160246

Panaite, V., Salomon, K., Jin, A., & Rottenberg, J. (2015). Cardiovascular recovery from psychological and physiological challenge and risk for adverse cardiovascular outcomes and all-cause mortality. *Psychosomatic Medicine*, *77*(3), 215–226. https://doi.org/10.1097/PSY.0000000000000171

Posada-Quintero, H. F., Florian, J. P., Orjuela-Cañón, A. D., Aljama-Corrales, T., Charleston-Villalobos, S., & Chon, K. H. (2016). Power spectral density analysis of Electrodermal activity for sympathetic function assessment. *Annals of Biomedical*

Engineering, *44*(10), 3124–3135. https://doi.org/10.1007/s10439-016-1606-6

Quer, G., Gouda, P., Galarnyk, M., Topol, E. J., & Steinhubl, S. R. (2020). Inter- and intraindividual variability in daily resting heart rate and its associations with age, sex, sleep, BMI, and time of year: Retrospective, longitudinal cohort study of 92,457 adults. *PLoS One*, *15*(2), e0227709. https://doi.org/10.1371/JOURNAL.PONE.0227709

Rahmani, A. M., Lai, J., Jafarlou, S., Azimi, I., Yunusova, A., Rivera, A. P., Labbaf, S., Anzanpour, A., Dutt, N., Jain, R., & Borelli, J. L. (2022). Personal mental health navigator: Harnessing the power of data, personal models, and health cybernetics to promote psychological well-being. *Frontiers in Digital Health*, *4*, 933587. https://doi.org/10.3389/FDGTH.2022.933587

Rajput, D., Wang, W. J., & Chen, C. C. (2023). Evaluation of a decided sample size in machine learning applications. *BMC Bioinformatics*, *24*(1), 1–17. https://doi.org/10.1186/S12859-023-05156-9/FIGURES/5

Ramos-Brieva, J. A., & Cordero Villafafila, A. (1986). The relationship between validity and reliability of the Spanish version of the Hamilton rating scale for depression. *Actas Luso-Espanolas de Neurologia, Psiquiatria y Ciencias Afines*, *14*(4), 335–338.

Remor, E. (2006). Psychometric properties of a European Spanish version of the perceived stress scale (PSS). *The Spanish Journal of Psychology*, *9*(1), 86–93. https://doi.org/10.1017/S1138741600006004

Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. D. (2014). Cognitive impairment in depression: A systematic review and meta-analysis. *Psychological Medicine*, *44*(10), 2029–2040. https://doi.org/10.1017/S0033291713002535

Salomon, K., Bylsma, L. M., White, K. E., Panaite, V., & Rottenberg, J. (2013). Is blunted cardiovascular reactivity in depression mood-state dependent? A comparison of major depressive disorder remitted depression and healthy controls. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *90*(1), 50–57. https://doi.org/10.1016/J.IJPSYCHO.2013.05.018

Salomon, K., Clift, A., Karlsdóttir, M., & Rottenberg, J. (2009). Major depressive disorder is associated with attenuated cardiovascular reactivity and impaired recovery among those free of cardiovascular disease. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association*, *28*(2), 157–165. https://doi.org/10.1037/A0013001

Sarchiapone, M., Gramaglia, C., Iosue, M., Carli, V., Mandelli, L., Serretti, A., Marangon, D., & Zeppegno, P. (2018). The association between electrodermal activity (EDA), depression and suicidal behaviour: A systematic review and narrative synthesis. *BMC Psychiatry*, *18*(1), 22. https://doi.org/10.1186/S12888-017-1551-4

Sarlon, J., Staniloiu, A., & Kordon, A. (2021). Heart rate variability changes in patients with major depressive disorder: Related to confounding factors, not to symptom severity? *Frontiers in Neuroscience*, *15*, 675624. https://doi.org/10.3389/FNINS.2021.675624/BIBTEX

Scarpina, F., & Tagini, S. (2017). The Stroop color and word test. *Frontiers in Psychology*, *8*(APR), 557. https://doi.org/10.3389/FPSYG.2017.00557

Schiweck, C., Gholamrezaei, A., Hellyn, M., Vaessen, T., Vrieze, E., & Claes, S. (2022). Exhausted heart rate responses to repeated psychological stress in women with major depressive disorder. *Frontiers in Psychiatry*, *13*, 869608. https://doi.org/10.3389/FPSYT.2022.869608/BIBTEX

Schiweck, C., Piette, D., Berckmans, D., Claes, S., & Vrieze, E. (2019). Heart rate and high frequency heart rate variability during stress as biomarker for clinical depression. A systematic review. *Psychological Medicine*, *49*(2), 200–211. https://doi.org/10.1017/S0033291718001988

Shaffer, F., Meehan, Z. M., & Zerr, C. L. (2020). A critical review of ultra-short-term heart rate variability norms research. *Frontiers in Neuroscience*, *14*. https://doi.org/10.3389/FNINS.2020.594880

Siddi, S., Bailon, R., Giné-Vázquez, I., Matcham, F., Lamers, F., Kontaxis, S., Laporta, E., Garcia, E., Lombardini, F., Annas, P., Hotopf, M., Penninx, B. W. J. H., Ivan, A., White, K. M., Difrancesco, S., Locatelli, P., Aguiló, J., Peñarrubia-Maria, M. T., Narayan, V. A., … Haro, J. M. (2023). The usability of daytime and night-time heart rate dynamics as digital biomarkers of depression severity. *Psychological Medicine*, *1–12*, 3249–3260. https://doi.org/10.1017/S0033291723001034

Snyder, H. R. (2013). Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: A meta-analysis and review. *Psychological Bulletin*, *139*(1), 81–132. https://doi.org/10.1037/A0028727

Sornmo, L., Bailon, R., & Laguna, P. (2024). Spectral analysis of heart rate variability in time-varying conditions and in the presence of confounding factors. *IEEE Reviews in Biomedical Engineering*, *17*, 322–341. https://doi.org/10.1109/RBME.2022.3220636

Spielberg, C. D., Gorsuch, R. L., & Lushene, R. E. (2008). *STAI: Cuestionario de ansiedad estado-rasgo* (7th ed.). TEA.

Steptoe, A., & Marmot, M. (2006). Psychosocial, hemostatic, and inflammatory correlates of delayed poststress blood pressure recovery. *Psychosomatic Medicine*, *68*(4), 531–537. https://doi.org/10.1097/01.PSY.0000227751.82103.65

Tada, Y., Yoshizaki, T., Tomata, Y., Yokoyama, Y., Sunami, A., Hida, A., & Kawano, Y. (2017). The impact of menstrual cycle phases on cardiac autonomic nervous system activity: An observational study considering lifestyle (diet, physical activity, and sleep) among female college students. *Journal of Nutritional Science and Vitaminology*, *63*(4), 249–255. https://doi.org/10.3177/JNSV.63.249

Tazawa, Y., Liang, K. c., Yoshimura, M., Kitazawa, M., Kaise, Y., Takamiya, A., Kishi, A., Horigome, T., Mitsukura, Y., Mimura, M., & Kishimoto, T. (2020). Evaluating depression with multimodal wristband-type wearable device: Screening and assessing patient severity utilizing machine-learning. *Heliyon*, *6*(2), e03274. https://doi.org/10.1016/J.HELIYON.2020.E03274

Thorell, L. H., Kjellman, B. F., & d'Elia, G. (1987). Electrodermal activity in relation to diagnostic subgroups and symptoms of depressive patients. *Acta Psychiatrica Scandinavica*, *76*(6), 693–701. https://doi.org/10.1111/j.1600-0447.1987.tb02941.x

Van Heeringen, K. (2012). Stress–diathesis model of suicidal behavior. *The Neurobiological Basis of Suicide*, 113–123. https://doi.org/10.1201/b12215

Varghese, S., Frey, B. N., Schneider, M. A., Kapczinski, F., & de Azevedo Cardoso, T. (2022). Functional and cognitive impairment in the first episode of depression: A systematic review. *Acta Psychiatrica Scandinavica*, *145*(2), 156–185. https://doi.org/10.1111/acps.13385

Varon, C., Lazaro, J., Bolea, J., Hernando, A., Aguilo, J., Gil, E., Van Huffel, S., & Bailon, R. (2019). Unconstrained estimation of HRV indices after removing respiratory influences from heart rate. *IEEE Journal of Biomedical and Health Informatics*, *23*(6), 2386–2397. https://doi.org/10.1109/JBHI.2018.2884644

Wang, Y. P., & Gorenstein, C. (2013). Psychometric properties of the Beck depression inventory-II: A comprehensive review. *Revista Brasileira de Psiquiatria (Sao Paulo, Brazil: 1999)*, *35*(4), 416–431. https://doi.org/10.1590/1516-4446-2012-1048

Wiebe, J. S., & Penley, J. A. (2005). A psychometric comparison of the Beck depression inventory – II in English and Spanish. *Psychological Assessment*, *17*(4), 481–485. https://doi.org/10.1037/1040-3590.17.4.481

Wilson Van Voorhis, C. R., & Morgan, B. L. (2007). Understanding power and rules of thumb for determining sample sizes. *Tutorials in Quantitative Methods for Psychology*, *3*, 43–50. https://doi.org/10.20982/tqmp.03.2.p043

Yang, J., & Kershaw, K. N. (2022). Feasibility of using ecological momentary assessment and continuous heart rate monitoring to measure stress reactivity in natural settings. *PLoS One*, *17*(3), e0264200. https://doi.org/10.1371/JOURNAL.PONE.0264200

Zitouni, M. S., Lih Oh, S., Vicnesh, J., Khandoker, A., & Acharya, U. R. (2022). Automated recognition of major depressive disorder from cardiovascular and respiratory physiological signals. *Frontiers in Psychiatry*, *13*, 970993. https://doi.org/10.3389/FPSYT.2022.970993/BIBTEX

Zorn, J. V., Schür, R. R., Boks, M. P., Kahn, R. S., Joëls, M., & Vinkers, C. H. (2017). Cortisol stress reactivity across psychiatric disorders: A systematic review and meta-analysis. *Psychoneuroendocrinology*, *77*, 25–36. https://doi.org/10.1016/j.psyneuen.2016.11.036

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Data S1: Supplementary data.

How to cite this article: Pagès, E. G., Kontaxis, S., Siddi, S., Miguel, M.-d., de la Cámara, C., Bernal, M. L., Ribeiro, T. C., Laguna, P., Badiella, L., Bailón, R., Haro, J. M., & Aguiló, J. (2024). Contribution of physiological dynamics in predicting major depressive disorder severity. *Psychophysiology*, *00*, e14729. https://doi.org/10.1111/psyp.14729