

Original Paper

Smartphone Ecological Momentary Assessment and Wearable Activity Tracking in Pediatric Depression: Cohort Study

Jimena Unzueta Saavedra¹, BS; Emma A Deaso¹, BS; Margot Austin¹, BS; Laura Cadavid¹, BS; Rachel Kraff¹, BS; Emma E M Knowles^{1,2}, PhD

¹Boston Children's Hospital, Boston, MA, United States

²Harvard Medical School, Boston, MA, United States

Corresponding Author:

Emma E M Knowles, PhD
Boston Children's Hospital
300 Longwood Avenue
Boston, MA, 02445
United States
Phone: 1 6179194628
Email: Emma.Knowles@childrens.harvard.edu

Abstract

Background: Adolescent depression is a significant public health concern. The presentation of depressive symptoms varies widely among individuals, fluctuating in intensity over time. Ecological momentary assessment (EMA) offers a unique advantage by enhancing ecological validity and reducing recall bias, allowing for a more accurate and nuanced understanding of major depressive disorder (MDD) symptoms. This methodology provides valuable insights into the fluctuating nature of depression, which could inform more personalized and timely interventions.

Objective: This study aims to (1) evaluate the feasibility of collecting smartphone-based EMA data alongside activity and sleep tracking in adolescents with depression; (2) investigate the severity and variability of mood symptoms reported over time; and (3) explore the relationship between mood, activity, and sleep.

Methods: Thirty-six participants (23 with MDD, 13 unaffected controls; 75% [n=27] female, mean age 19.50 y) completed twice-daily EMA check-ins over 2 weeks, complemented by continuous activity and sleep monitoring using FitBit Charge 3 devices. The study examined feasibility, usability of the EMA app, symptom severity and variability, and relationships between mood, activity, and sleep. We applied linear mixed-effects regression to the data to examine relationships between variables.

Results: Participants completed a total of 923 unique check-ins (mean check-ins per participant=25.60). Overall compliance rates were high (91.57%), indicating the approach is highly feasible. MDD participants demonstrated greater symptom severity and variability over time compared with controls ($\beta=34.48$, $P<.001$). Individuals with MDD exhibited greater diurnal variation ($\beta=-2.54$, $P<.001$) with worse mood in the morning and worse mood than anxiety scores over time ($\beta=-6.93$, $P<.001$). Life stress was a significant predictor of more severe EMA scores ($\beta=24.50$, $P<.001$). MDD cases exhibited more inconsistent sleep patterns ($\beta=32.14$, $P<.001$), shorter total sleep times ($\beta=-94.38$, $P<.001$), and a higher frequency of naps ($\beta=14.05$, $P<.001$). MDD cases took fewer steps per day (mean 5828.64, SD 6188.85) than controls (mean 7088.47, SD 5378.18) over the course of the study, but this difference was not significant ($P=.33$), and activity levels were not significantly predictive of EMA score ($P=.75$).

Conclusions: This study demonstrates the feasibility of integrating smartphone-based EMA with wearable activity tracking in adolescents with depression. High compliance rates support the practicality of this approach, while EMA data provide valuable insights into the dynamic nature of depressive symptoms, particularly in relation to sleep and life stress. Future studies should validate these findings in larger, more diverse samples. Clinically, EMA and wearable tracking may enhance routine assessments and inform personalized interventions by capturing symptom variability and external influences in real time.

JMIR Form Res 2025;9:e66187; doi: [10.2196/66187](https://doi.org/10.2196/66187)

Keywords: depression; ecological momentary assessment; activity; sleep; mood; adolescent

Introduction

Adolescent depression is a significant mental health crisis; 14.7% of the adolescent population reports at least one major depressive episode with severe impairment [1]. This trend predates the COVID-19 pandemic [2] and has continued apace [3]. The impact of adolescent depression is severe; a depressive episode leads to immediate debilitating effects plus long-term consequences [4], for example, impaired academic performance [5] and challenges in forming interpersonal relationships [6]. Depression that begins in adolescence often follows a recurrent pattern [7], persists into adulthood, and is associated with higher levels of anxiety, substance abuse, and impaired functioning in later life [8]. Importantly, depression does not manifest uniformly across individuals; major depressive disorder (MDD) is known for its heterogeneous symptomatology [9]. The course and severity of depression can vary widely among individuals [10]. Moreover, depressive symptoms can fluctuate within the same individual over time [11,12].

Despite the fluctuating and dynamic nature of MDD, typical clinical assessment methods are cross-sectional. Most standard symptom assessments often require patients to recall how they have felt for the previous 1-2 weeks [13]. Recall of mood states in the previous weeks may be inaccurate depending on cognitive style and illness severity [14-18]. Ecological momentary assessment (EMA) bridges this gap. EMA refers to the repeated and brief assessments of the same person over multiple days. EMA enables a granular view of depressive symptoms in real-world settings, overcoming the temporal and contextual limitations of standard assessments [19]. In particular, EMA provides increased ecological validity by decreasing recall bias [20]. It also enables the measurement and modeling of within-person variability since changes can be tracked over time. Thus, this tool facilitates the assessment of symptom fluctuations, offering insights into the dynamic nature of the disorder. In addition, repeated assessment of one or more variables (eg, mood and sleep) allows for the examination of covariation between variables over time [21].

Thanks to growth in technology, EMA can be conveniently completed using the ubiquitous smartphone and can be integrated with wearables, such as activity trackers (eg, FitBit) to monitor sleep and physical activity [19]. Despite the many advantages of EMA and its potential clinical utility [22-24], literature on the topic is limited by sample type, and also by EMA method and schedule intensity [25]. The

present manuscript adds to the extant literature in a number of ways. First, most EMA mood studies have been conducted in adults [25,26], while the present study focuses on adolescents with MDD. Second, many studies that focused on adolescents used phone-based EMA [27-35] where interviewers called participants on a schedule and recorded responses, or paper-based EMA [36,37] where participants fill in paper diaries on a schedule, leaving the utility of smartphone-based EMA, which we used in the present study, still mostly unknown [25,38]. To our knowledge, only one study has measured mood over time in depressed patients using smartphone, app-based EMA, which comprised a single mood question administered over the course of 6 days [39]. Third, there is a notable gap in terms of moderate duration studies with high-frequency check-ins, the present study had 2 weeks of twice-daily check-ins. Previous studies have either been much shorter (eg, 4 days) [29] or longer studies but with less frequent check-ins (eg, 8 weeks of check-ins with 4-day intervals) [30,33]. Table 1 shows details of cited studies. More data are needed to demonstrate that momentary assessment of depressive symptoms in adolescents using smartphone-based apps is possible. Additionally, the collected data behave as expected, that is, depressed subjects show greater severity and variability of symptoms than healthy controls, and related measures covary with those symptoms.

The present study had 3 primary goals, all of which contribute to assessing the feasibility of smartphone-based EMA in adolescents with depression. First, we aimed to evaluate whether participants could be retained in the study and engage consistently with smartphone-based EMA and wearable data collection. Second, we assessed whether mood reports captured expected symptom variation, specifically whether adolescents with depression demonstrated greater severity and variability of mood and anxiety symptoms compared with healthy controls, supporting the validity and potential clinical utility of this method. Third, we examined whether participants found the app-based EMA approach acceptable by evaluating self-reported satisfaction and usability of the EMA methods. By addressing these key aspects of feasibility, this study provides critical insights into the viability of smartphone-based EMA for real-time symptom monitoring and personalized intervention development in adolescent depression. As this is a pilot study with a small sample size, the findings should be viewed as preliminary. Future studies with larger, more diverse samples will be necessary to confirm and expand upon the relationships observed here, including potential subgroup-specific patterns within the heterogeneous construct of MDD.

Table 1. Characteristics of existing studies of ecological momentary assessment (EMA) of mood and associated variables in adolescent samples with and without major depressive disorder (MDD).

Study	Sample	Age range (years)	EMA focus	EMA duration	Check-in frequency	EMA method	Actigraphy
Axelsson et al [27]	5 controls 11 MDD	10-17	Location, social context, mood, media use, future plans, significant events	5 four-day blocks (Fri-Mon)	12 calls per block	Answer-only cellular phone	No

Study	Sample	Age range (years)	EMA focus	EMA duration	Check-in frequency	EMA method	Actigraphy
Silk et al [32]	• 5 BP • 16 controls • 19 MDD	8-17	Positive and negative affect, and companions	5 four-day blocks (Fri-Mon)	12 calls per block	Answer-only cellular phone	No
Whalen et al [35]	• 23 controls • 30 MDD	7-17	Positive and negative affect, caffeine consumption	5 four-day blocks (Fri-Mon)	12 calls per block	Answer-only cellular phone	No
Mor et al [37]	• 278 healthy individuals	— ^a	Mood, stress, self-focus, social activity	3 days	6 per day	Paper diary	No
Forbes et al [29]	• 28 controls • 15 MDD	8-17	Positive and negative affect	4 days	1 per day	Cell phone	No
Silk et al [33]	• 32 controls • 47 MDD	7-17	Behavior, emotion, social context	5 four-day blocks (Fri-Mon)	12 calls per block	Answer-only cellular phone	No
Cousins et al [28]	• 23 controls • 23 anxiety • 42 MDD	8-16	Positive and negative affect	Two 4 day blocks	12 calls per block	Answer-only cellular phone	Yes
Primack et al [31]	• 60 controls • 46 MDD	7-17	Media exposures (eg, internet)	5 four-day blocks (Fri-Mon)	12 calls per block	Answer-only cellular phone	No
Forbes et al [30]	• 31 MDD and anxiety • 23 anxiety • 12 MDD	8-16	Positive and negative affect, and companions	4 days	12 calls total	Answer-only cellular phone	No
Waller et al [34]	• 31 controls • 29 MDD	11-17	Behavior, emotion, social context	3 five-day blocks (Thurs-Mon)	42 calls total	Answer-only cellular phone	No
Frost et al [36]	• 353 healthy individuals	11-18	Feelings and activities	7 days	8 per day	Paper diary	No
Bickham et al [40]	• 125 healthy individuals	12-15	Media use	Two weeks	48 assessments	Handheld computer	No
Cushing et al [41]	• 20 healthy individuals	13-18	Affect and energy	20 days	4 assessments per day	Cell-phone app on study device	No
Minich and Moreno [39]	• 253 healthy individuals	12-17	Mood	6 days	30 assessments total	Text message	No

^aNot available.

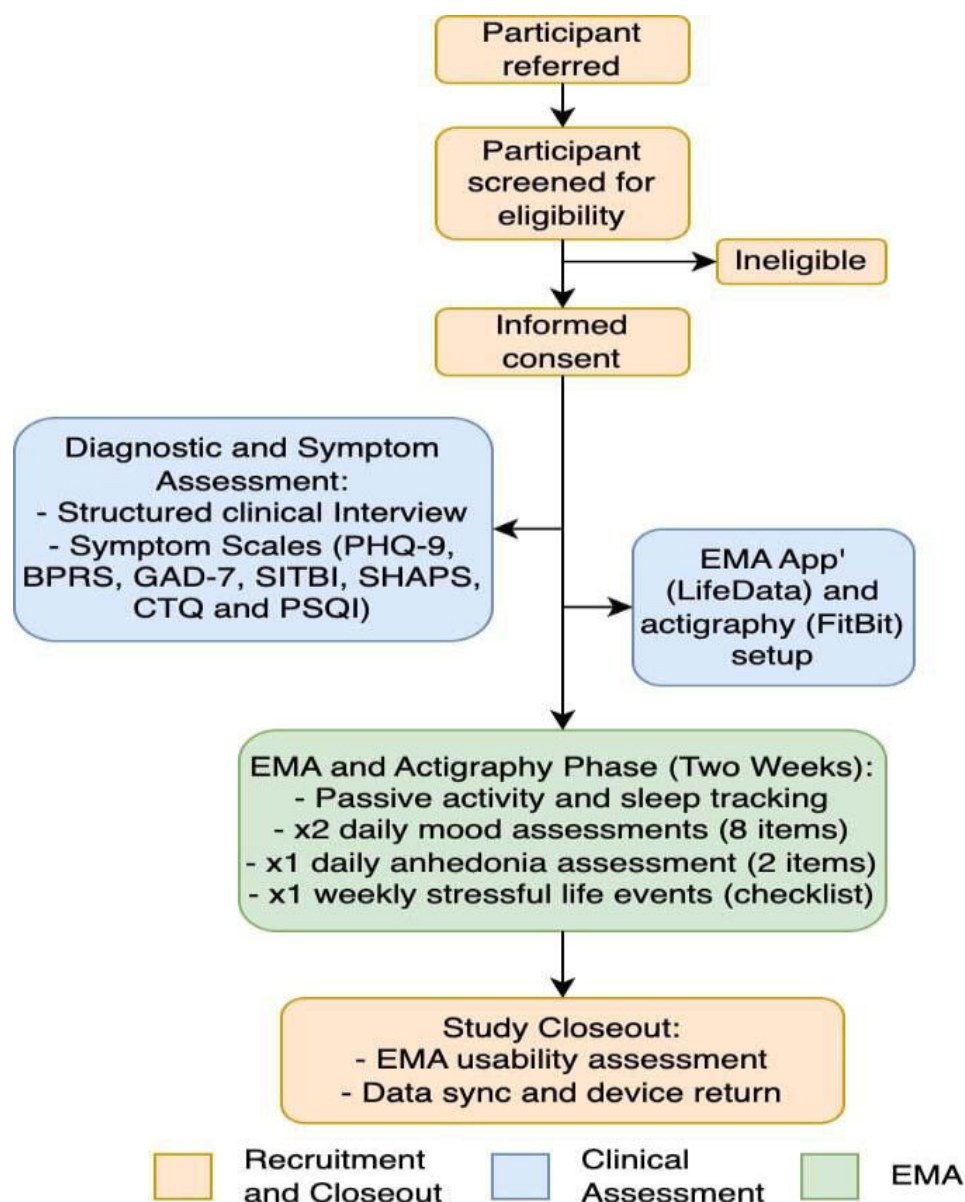
Methods

Participants completed the protocol across a mean time span of 7 weeks.

Procedures

Figure 1 depicts the study procedures. Study procedures were identical across participants regardless of case status.

Figure 1. Study protocol. After being consented, the presence (or absence) of major depressive diagnoses (and other Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition [DSM-V] diagnoses) was confirmed using a structured clinical interview. Symptom scales for depression, general psychopathology, anxiety, suicidality and self-injury, anhedonia, and trauma were administered. A smartphone app that enabled ecological momentary assessment (EMA) to take place (LifeData) was set up on the participant's phone. Each participant was given a FitBit, and the device was linked to the participant's smartphone (for data syncing). Participants completed 2 weeks of EMA and actigraphy concurrently. At the end of the 2 weeks, participants returned the FitBit device, and they completed an EMA usability assessment. BPRS: Brief Psychiatric Rating Scale; CTQ: Childhood Trauma Questionnaires; SHAPS: Snaith-Hamilton Pleasure Scale; SITBI: Self-Injurious Thoughts and Behaviors Interview.



Participant Recruitment

The sample comprised 36 individuals (75% female [$n=27$], mean age 19.50 y, SD 3.92 y, range 14-27 y) recruited from the Boston, MA, area (Table 2). The majority of the sample identified as White (67% [$n=29$]), with 2 individuals identifying as Hispanic or Latino, 4 individuals identifying as Black or African American, and 1 individual identifying as Asian. Twenty-three individuals had a MDD and were recruited from outpatient psychiatry services at Boston Children's Hospital. Providers requested consent from the participant and their family for the research team to contact them about the study, and research staff followed up to explain the study and, if appropriate, arrange for a consent

visit to take place. The control group ($N=13$) was recruited from the community via flyers and advertisements placed in Boston. Exclusion criteria for individuals with MDD included a severe neurodevelopmental disorder or other impairment that impacts the participant's ability to provide the required information for the study (eg, symptom reports), a substance or medication-induced affective disorder, an affective disorder secondary to a medical condition, and a current Axis I psychotic or bipolar disorder. Exclusion criteria for controls included a severe neurodevelopmental disorder or a current or past psychiatric diagnosis as defined by DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition) criteria.

Table 2. Baseline characteristics of the depressed and control groups.

	Depressed	Controls
Age		
Mean (SD)	17.91 (2.98)	22.31 (3.90)
Range	14-26	14-27
Sex		
Female, n (%)	15 (65)	11 (92)
Parental education, n (%)		
Graduate professional training (graduate degree)	12 (53)	— ^a
Partial college or specialized training	5 (20)	—
Standard college or university graduation	6 (27)	—
Parental occupation, n (%)		
Higher executives, proprietors of large businesses, major professionals	6 (27)	—
Administrators, lesser professionals, proprietors of medium-sized businesses	8 (36)	—
Small business owners, farm owners, managers, minor professionals	6 (27)	—
Technicians, semi-professionals, small business owners (US \$50,000-US \$70,000)	2 (10)	—
Clinical assessments		
PHQ-9 ^b , mean (SD)	9.05 (5.81)	1.69 (1.32)
GAD-7 ^c , mean (SD)	6.65 (6.34)	2.31 (2.72)
BPRS ^d , mean (SD)	32.60 (5.97)	26.20 (2.45)
PSQI ^e , mean (SD)	8.67 (3.31)	5.27 (3.29)
Medications		
Benzodiazepine	1	0
Betablocker	1	0
Birth Control	2	7
IUD ^f	0	2
Cognitive Enhancer	3	0
Anticonvulsant	2	0
Antidepressant (SSRI)	12	1
Antidepressant (NDRI)	0	1
Antidepressant (Tricyclic)	1	0
Atypical Antipsychotic	3	0
Stimulant	3	1
Depression diagnoses		
Major depressive disorder, lifetime	2	0
Major depressive disorder, recurrent	11	0
Major depressive disorder, single episode, remission	2	0
Other specified depressive disorder (insufficient symptoms), current	1	0
Past major depressive disorder	2	0
Persistent depressive disorder	5	0
Comorbidities		
Eating disorders		
Anorexia nervosa, in remission	1	0
Bulimia nervosa, in remission	1	0
Anxiety disorders		
GAD	2	0
OCD ^g	1	0
PTSD ^h	1	0
Social anxiety disorder	4	1

	Depressed	Controls
Neurodevelopmental disorders		
ADHD	6	0

^aNot applicable.
^bPHQ: patient health questionnaire.
^cGAD: generalized anxiety disorder.
^dBPRS: Brief Psychiatric Rating Scale.
^ePSQI: Pittsburgh Sleep Quality Index.
^fIUD: intrauterine device.
^gOCD: obsessive-compulsive disorder.
^hPTSD: posttraumatic stress disorder.

Ethical Considerations

This study was reviewed and approved by the Boston Children’s Hospital Institutional Review Board (BCH IRB-P00031981). The study involved the collection and analysis of data from human subjects, and all procedures were conducted in accordance with the IRB-approved protocol. All participants provided informed consent/assent using forms approved by the institutional review board at Boston Children’s Hospital. This ensured that participants were fully informed about the study’s purpose, procedures, potential risks, and benefits before agreeing to participate. To maintain participant confidentiality, all collected data were stored in secure databases and deidentified prior to analysis. All data included in the present manuscript or any supplementary materials in [Multimedia Appendix 1](#) are completely anonymous. There was no cost or fee associated with study participation. Participants were compensated for their time and effort in completing the study procedures. They were eligible to receive up to US \$144, with payments provided via ClinCard, a secure electronic payment system. Payment breakdown was as follows: US \$20 for Structured Clinical Interview for DSM-5; US \$20 for questionnaire completion; US \$14 (US \$1 per day) for FitBit wearing; US \$14 for FitBit return; US \$14 for usability questionnaire; and US \$56 (US \$2 per check-in for EMA). Participants were paid for all aspects of the study that they completed, including individual EMA check-ins. No identifiable images of individual participants are included in this manuscript or supplementary materials in [Multimedia Appendix 1](#).

Diagnostic and Symptom Assessment

The presence or absence of major depression diagnoses (and comorbid disorders) was confirmed using the Structured Clinical Interview for DSM-5 [42] (with additional modules from the Kiddie Schedule for Affective Disorders and Schizophrenia interview [43] if <18 years of age) in order to confirm past or current MDD (and other DSM diagnoses). Depressive symptoms were assessed at study intake using the patient health questionnaire (PHQ-9) [44]. Participants were also administered the Brief Psychiatric Rating Scale [45], generalized anxiety disorder (GAD-7) [46], and the Pittsburgh Sleep Quality Index [47]. Parents of probands were administered the Hollingshead Index of Socioeconomic Status [48], and 25 parents agreed to complete the scale.

Ecological Momentary Assessment (EMA)

EMA data were recorded using LifeData [49]. LifeData is a smartphone-based EMA program that runs on both Android and iPhone smartphones. Check-ins are delivered via the LifeData mobile app RealLifeExp. EMA lasted 2 weeks, during this time participants received a push notification twice per day (6 AM and 4 PM), reminding them to complete a check-in. Notifications were sent 4 times (at 1 h intervals) per check-in, allowing each participant a 4-hour window to complete the check-in (6-10 AM and 4-10 PM).

For each check-in, participants were asked the same 8 questions. Participants rated the items using a slider, which translated to a score (0-100): “I am sad.”; “I feel bothered by every little thing.”; “I have no interest in things I would usually enjoy (eg, food, TV, games, spending time with friends/family).”; “I do not have enough energy to get going.”; “I have no appetite or feel much hungrier than usual.”; “I feel physically tense and/or jittery.”; “I am nervous, anxious or on edge.”; and “I can’t stop worrying.” Items were designed to be reflective of established measures of depression (PHQ-9) [44] and anxiety (GAD-7) [46], with the adaptation that they could be administered multiple times per day (Table S1 in [Multimedia Appendix 1](#) for an explanation of each item). For a subset of statistical models (see below), the first 5 items were classed as measuring mood and the rest as measuring anxiety. Once per week, during an afternoon check-in, participants completed a checklist of stressful life events. The question reads “Did any of the following happen to you since the last time you completed one of these surveys? Check all that apply”. The participant selects as many of the following options as appropriate: “I argued with a friend or family member.”; “I was not allowed to do something I wanted to do.”; “I got a bad grade in school.”; “My parents have been arguing a lot.”; “Somebody in my family got a serious illness.”; “Someone in my family was arrested.”; “Somebody teased or threatened me.”; “I teased somebody else.”; “I did something that made me feel embarrassed.”; “Someone commented negatively on the way I look.”; “I was excluded from a group event.”; and “I got disciplined or suspended from school.”

Physical Activity (Steps) and Sleep Features

Participants were issued a FitBit Charge 3 and were instructed to wear it for the duration of the EMA portion of the study. Twenty-two individuals (11 controls (mean age 22.90, SD 3.27 y, 10 females) and 11 MDD cases (mean age 18.20, SD 2.86 y, 7 females) wore the FitBit with sufficient regularity for data to be analyzed. Data were collected on daily activity (number of steps) and sleep.

Fitbit data were obtained via the Fitbit API, which provides preprocessed JSON files reflecting proprietary algorithms for step counting and sleep [50]. For sleep, each participant had one JSON file that spanned the duration of the study, which included a sleep log with timestamps for sleep onset and offset and a breakdown of sleep stages (eg, light, rapid eye movement [REM], deep). For steps, each JSON file contained minute-level step count data linked to a timestamp. Data in the JSON files were converted to csv using a combination of standard and pandas [51] libraries in PyCharm [52] (Python version 3.9; Python Software Foundation) and rearranged for analysis using tidy [53] and dplyr [54] in R [55]. Following others' work [56] in the field, we extracted numerous sleep features from the data collected using the Fitbit device. These are related to sleep architecture, quality, and stability (Table S2 in [Multimedia Appendix 1](#) contains details or individual metrics). A nap was defined as time spent in bed lasting <180 minutes, FitBit does not calculate deep, light, and REM sleep during these shorter sleeps. Therefore, naps were included in analyses focused on total sleep time (TST) and time in bed (TIB) but were excluded from analyses focused on sleep stage. Sleep quality metrics were calculated per all logged sleeps per individual, and sleep stability metrics were derived across all logged sleeps. Steps that were recorded during a documented period of sleep were recoded to 0 as they were likely due to movement in bed rather than steps.

Usability and Tolerability Questionnaire

At the end of participation, all participants completed a short questionnaire that evaluated their experience using the EMA app (eg, if the app was easy to use; see in [Multimedia Appendix 1](#) for a list of questions). All but one of the statements were rated from 1 to 7, with 1 being strongly disagree, 4 neutral, and 7 strongly agree. Responses were recorded as disagree (1-3) and agree (5-7). Analysis included frequencies and proportions of responses per item. One item, relating to where participants used the app, included a free text response (unique responses are included in the [Multimedia Appendix 1](#)).

Statistical Analyses

All analyses were conducted in R (version 4.2.1; R Foundation for Statistical Computing) [55].

How Does Depressive Symptomatology Vary Over Time?

In line with others' EMA work [57], we opted to examine variability in depressive symptoms both visually and statistically. Time series data were examined using standardized (group-mean-centered) and raw data. Group-mean-centered plots depict the relative variability over time across the sample. Before plotting data were group-mean centered (by subtracting each participant's mean from each individual score), which results in standardization of the figure such that units are in units of deviation from each individual's average (set to 0). Per subject, raw score time-series plots depict the fine-grained differences between each individual's data. Data were plotted using ggplot2 [58].

We used several statistics to quantify variability, including the root mean square of successive differences (RMSSD) [59], coefficient of variation (CoV) [60], and intraclass correlations (ICC). The RMSSD reflects variability over time. The CoV reflects the dispersion of the data and is useful for comparing variability when means differ (eg, in cases vs controls). The ICC reflects the proportion of variance attributable to between-person variability, thus, 1-ICC indicates the proportion of variability attributable to within-person variability. Variability statistics were calculated using a combination of base R and the psych [61] package.

Predictors of Depression Severity and Variability Over Time

The EMA data were nonindependent, for example, we might expect data collected within the same person to be more strongly related than those collected from 2 different people. This structure is sometimes referred to as multilevel; observations are nested within item type (mood or anxiety), which are nested within the time of day (AM or PM) and day, which are nested within the participant. Therefore, when evaluating the impact of various predictors on EMA scores; we applied linear mixed-effects regression to the data using the package lme4 [62]. Effect sizes for fixed effects (η_p^2) were calculated using the effectsize package [63], and were interpreted as small effect: $\eta_p^2 \approx 0.01$, medium effect: $\eta_p^2 \approx 0.06$, large effect: $\eta_p^2 \approx 0.14$. Similarly, observations for sleep architecture and quality were nested within a day (where some individuals logged sleep with onset times within the same day), within the sleep stage, and within the participant. Multilevel modeling is robust to missing data, and therefore we did not impute the small amount of missing data that was present in the EMA, of 1008 expected check-ins, 923 were observed (overall compliance=92%). Prior to testing the effect of various predictors on depression, we tested an intercept-only, or unconditioned model to evaluate the need for a random effect. Significant random effects (determined using the lmerTest [64] package) were retained in subsequent models. Given the different distributions of age and sex in case and control groups, these variables were included as covariates in all models.

First, we asked whether individuals with a MDD diagnosis demonstrated greater symptom severity (indexed using EMA score) and variability than controls. We also evaluated the impact of sex and age on severity, as well as fixed effects of Time of Day and Item Type. Second, we examined whether anxiety and mood covaried over time and whether anxiety or mood at the preceding check-in significantly predicted mood at the following one. Third, we asked whether life stress was associated with increased symptom severity. Fourth, we examined whether activity levels (number of steps) were associated with MDD diagnosis or symptom severity. Fifth, we assessed whether number of minutes asleep, sleep onset, or sleep-waking was associated with MDD diagnosis or symptom severity. Sixth, we examined the results of the usability questionnaire.

Results

Overview

The total unique check-ins completed by participants=923 (mean check-ins per participant 25.60, SD 2.78). Assessments

were completed across a total of 486 unique days (mean days per participant 13.50, SD 1.25), and yielded 1.90 mean check-ins per participant per day. For the 2 required check-ins per day compliance rate of 91.57%, cases demonstrated slightly lower compliance (89.94%) than controls. Across individual EMA items, this yielded 7336 data points for regression modeling.

How Does Depressive Symptomatology Vary Over Time?

Table 3 shows the descriptive and variability statistics for the sum score of all EMA items as well as for the individual items. In terms of reported symptom severity, MDD cases scored higher than controls on all items. Examination of variability statistics (Table 3) indicates that there was considerable variability in both the sum score and the individual items. The RMSSD and CoV suggest greater variability in MDD cases than in controls. Examination of the ICC across items suggests that ~50% of the variance in item-level responses was due to within-person variance in MDD cases.

Table 3. Descriptive and variability statistics for individual ecological momentary assessment items in depressed and control groups.

Item	Group	Total responses	Non-zero responses, n (%)	Mean (SD)	Range	Skewness	RMSSD ^a	CoV ^b	ICC ^c
Anxious	Ctrl	340	44.41	9.77 (8.93)	0-90	1.57	11.93	0.91	0.46
Anxious	MDD	572	76.05	36.1 (19.91)	0-100	0.4	24.59	0.55	0.55
Appetite	Ctrl	342	17.54	4.23 (7.41)	0-80	3	9.72	1.75	0.17
Appetite	MDD	575	76	41.65 (20.38)	0-100	0.03	25.89	0.49	0.58
Bothered	Ctrl	346	30.35	4.94 (5.89)	0-80	2.27	7.53	1.19	0.35
Bothered	MDD	576	68.58	35.97 (20.07)	0-100	0.54	24.55	0.56	0.56
Energy	Ctrl	342	39.77	8.43 (9.91)	0-80	1.72	14.03	1.18	0.39
Energy	MDD	575	80.52	46.04 (22.7)	0-100	0.29	30.75	0.49	0.5
Interest	Ctrl	344	13.95	2.17 (3.11)	0-60	2.84	4.37	1.43	0.45
Interest	MDD	575	71.13	32.14 (18.38)	0-100	0.6	23.88	0.57	0.5
Sad	Ctrl	347	27.38	4.38 (5.52)	0-50	2.72	7.26	1.26	0.31
Sad	MDD	576	65.80	28.71 (17.01)	0-100	1.15	22.41	0.59	0.59
Tense	Ctrl	341	34.90	6.98 (7.82)	0-80	1.87	10.13	1.12	0.36
Tense	MDD	574	65.85	30.73 (20.22)	0-100	0.72	24.03	0.66	0.54
Worry	Ctrl	340	22.35	5.5 (5.73)	0-100	2.48	6.43	1.04	0.5
Worry	MDD	571	69	31.27 (18.05)	0-100	0.77	21.3	0.58	0.57
Sum score	Ctrl	340	65.29	46.58 (34.95)	0-400	1.64	44.39	0.75	0.51
Sum score	MDD	571	93.35	282.52 (99.06)	0-800	0.38	115.58	0.35	0.7

^aRMSSD: root mean square of successive differences.

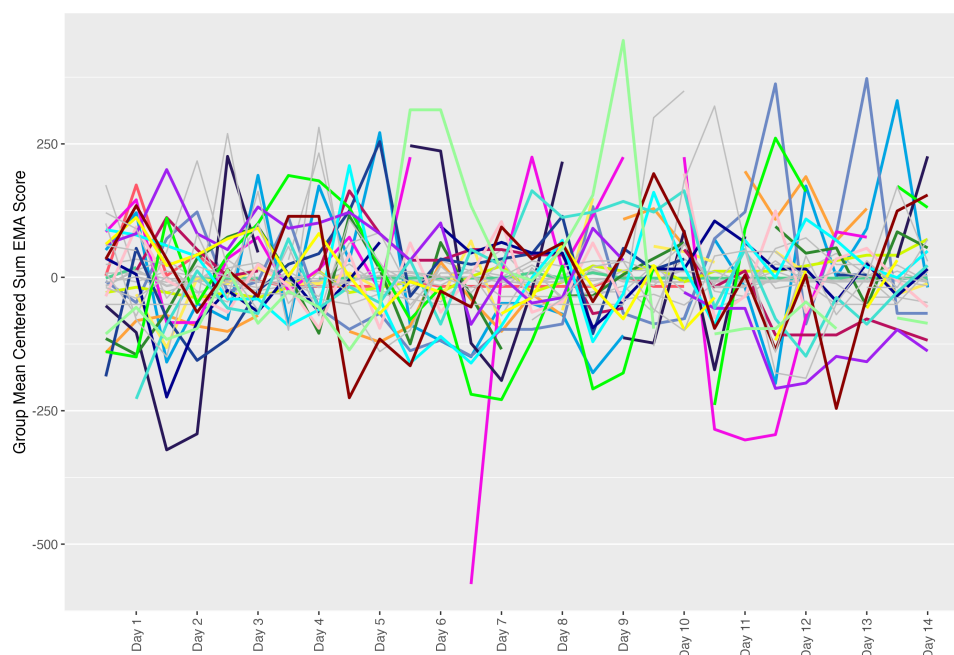
^bCoV: coefficient of variation.

^cICC: intraclass correlation.

Figure 2 shows the group-mean centered time series plot of the sum score. At the group level, scores fluctuated in a saw-tooth pattern, with repeated rises and falls but no clear linear trend of improvement or worsening over time. This pattern reflects short-term variability in EMA data and highlights overall fluctuations in symptom severity across the

sample. Item-level data followed a similar pattern (Figure S1 in Multimedia Appendix 1); in MDD cases, variability in symptom reporting did not appear to vary as a function of item. Controls appear to report a greater degree of variability to the irritability and anhedonic items (Bothered and Interest) than they did on mood or anxiety symptoms.

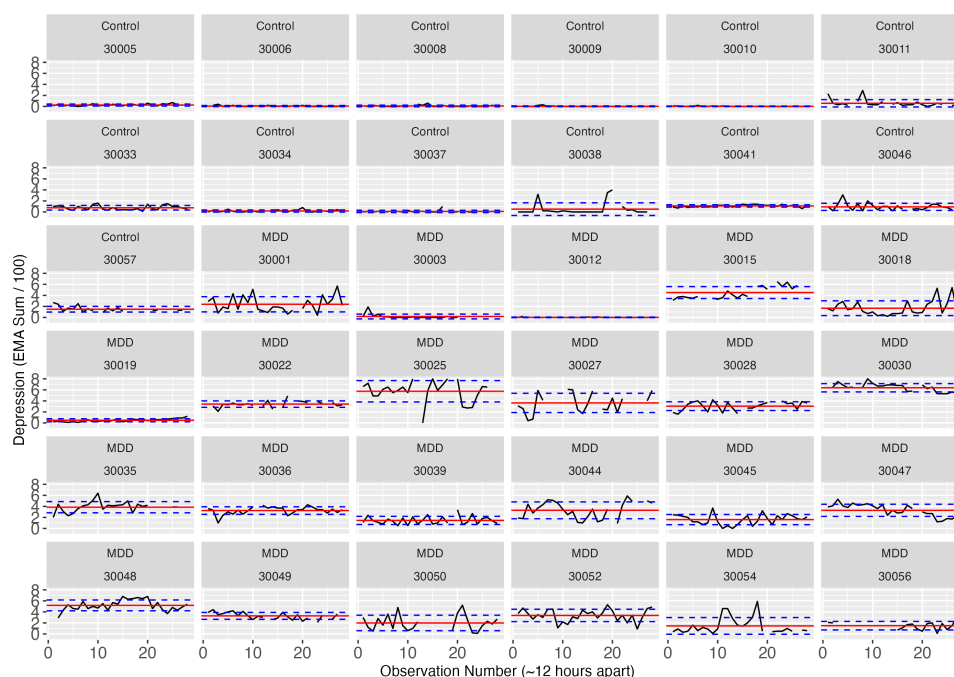
Figure 2. Time-series plot of the group-mean centered ecological momentary assessment (EMA) data (sum score across all 8 items). Colored lines represent individuals with a major depressive disorder (MDD) diagnosis, gray lines represent controls.



At the individual level, we observe heterogeneity (Figure 3) where some individuals with a MDD diagnosis demonstrate worsening throughout the course of the study (eg, 30,001, 30,015, 30,028, 30,030, 30,044, and 30,048), and others appear to improve (eg, 30,035, 30,036, and 30,047). Interestingly, not all controls demonstrated floor effects. While controls reported more zero responses than cases (Table 3) some endorsed minimal symptoms (eg, 30,011, 30,033, 30,038, 30,046, and 30,057) throughout

the course of the study; admittedly these symptoms would likely amount to sub-threshold symptoms for a MDD diagnosis but this suggests that healthy controls can demonstrate variation in depressive symptomatology. Three MDD cases (30,003, 30,012, and 30,019) demonstrated minimal symptoms; diagnoses for these individuals were in line with low symptoms (Other Specified Depressive Disorder [Insufficient Symptoms], Recurrent MDD, and Past MDD Single Episode, respectively).

Figure 3. Time series plots of the raw ecological momentary assessment (EMA) data (sum score across all 8 items). The black line represents the reported score, dashed lines are controls, and solid lines are those with an major depressive disorder (MDD) diagnosis. Mean score (per individual) is represented by the red horizontal line, and blue dashed lines represent SD of 1.



Predictors of Depression Severity and Variability Over Time

The unconditioned model indicated significant between-subject nested effects of Participant, Day, and Item Type but not Time of Day (Table S3 in [Multimedia Appendix 1](#)). All significant random effects were retained in subsequent modeling.

Effect of MDD on EMA Score Severity Over Time

We examined the impact of MDD diagnosis on summed EMA score using a linear mixed model ($R^2_{\text{marginal}}=0.26$, $R^2_{\text{conditional}}=0.65$) with fixed effects of MDD (plus age, sex, item type, and time of day). The model indicated significant main effects of MDD ($\eta^2_p=0.41$, $\beta=34.48$, 95% CI 20.22 to 48.74, $P<.001$), sex ($\eta^2_p=0.16$, $\beta=16.81$, 95% CI 3.55 to 30.07, $P=.01$), and a small but significant effects of item type ($\eta^2_p=0.01$, $\beta=-1.69$, 95% CI -2.27 to -0.55, $P=.01$) and time of day ($\eta^2_p=0.002$, $\beta=-1.41$, 95% CI -2.27 to -0.55, $P<.001$). The effect of age ($P=.85$) was not significant.

We fit an additional linear mixed model ($R^2_{\text{marginal}}=0.27$, $R^2_{\text{conditional}}=0.65$) with MDD*time of day and MDD*item type interactions to establish whether there was greater diurnal variation in mood in cases than controls and whether there was a greater effect of item type in cases than in controls. There was a significant MDD*time of day interaction ($\eta^2_p=0.001$, $\beta=-2.54$, 95% CI -4.32 to -0.76, $P<.001$; Figure S2 in [Multimedia Appendix 1](#)) where cases exhibited diurnal mood variation (higher scores in the morning) but controls did not, η^2_p indicates that the effect was minimal. There was also a significant MDD*item type interaction ($\eta^2_p=0.001$, $\beta=-6.93$, 95% CI -9.64 to -4.21, $P<.001$; Figure S3 in [Multimedia Appendix 1](#)) where cases demonstrated worse mood than anxiety scores and controls demonstrated the opposite effect; however, η^2_p indicates that the effect was minimal.

PHQ-9 score at baseline ($\eta^2_p=0.58$, $\beta=2.74$, 95% CI 1.92 to 3.57, $P<.001$) significantly predicted EMA scores over the course of the study, suggesting that the items we selected for EMA have good convergent validity with an established measure of depression.

Cross-Sectional Covariation of Mood and Anxiety Over Time

We examined cross-sectional relationships between mood and anxiety over time using a linear mixed model ($R^2_{\text{marginal}}=0.33$, $R^2_{\text{conditional}}=0.87$) of summed EMA mood items with fixed effects of summed EMA anxiety items and MDD (plus age and sex). This model indicated significant covariation between mood and anxiety over time (Figure S4 in [Multimedia Appendix 1](#) shows raw data per participant), main effects of anxiety ($\eta^2_p=0.25$, $\beta=0.64$, 95% CI 0.56 to 0.71, $P<.001$) and MDD ($\eta^2_p=0.45$, $\beta=132.14$, 95% CI 81.24 to 183.05,

$P<.001$) were significant, but neither age ($P=.22$) nor sex were ($P=.39$).

Impact of Preceding Anxiety Score on Mood

We examined whether prior anxiety ratings predicted later mood scores using a linear mixed model with the same predictors as above, plus anxiety score from the preceding check-in (where $t=\text{time}$, T-1) as a fixed effect ($R^2_{\text{marginal}}=0.63$, $R^2_{\text{conditional}}=0.89$). The preceding anxiety score was a nominally significant fixed effect, but the size of the effect was small ($\eta^2_p=0.01$, $\beta=0.11$, 95% CI 0.04 to 0.19, $P<.001$).

Impact of Preceding Mood on Subsequent Mood Score

We examined whether prior mood score predicted later mood score using a linear mixed model ($R^2_{\text{marginal}}=0.46$, $R^2_{\text{conditional}}=0.88$) with T-1, T-2, and T-3 mood scores as a fixed effect in addition to MDD (plus age and sex). The immediately preceding mood score (T-1; $\eta^2_p=0.02$, $\beta=0.11$, 95% CI 0.04 to 0.18, $P<.001$) was a significant predictor of subsequent mood score, as was mood score that was separated by one check-in (T-2; $\eta^2_p=0.03$, $\beta=0.16$, 95% CI 0.09 to 0.23, $P<.001$) and 2 (T-3; $\eta^2_p=0.02$, $\beta=0.11$, 95% CI 0.04 to 0.18, $P<.001$) check-ins. MDD ($\eta^2_p=0.53$) remained a significant predictor of mood score, age ($P=.66$), and sex ($P=.55$) were not significant predictors in this model.

Effect of Life Stress on EMA Score Severity

Both MDD cases and controls reported life stress (Table S4 in [Multimedia Appendix 1](#)). Most stressful life events in control are related to arguments with friends or family and embarrassment. These stressors were also common in MDD cases, but they also endorsed stress related to instances of not being allowed to do what they wanted, being teased (and teasing others), or having negative comments made about their appearance. MDD cases reported 91 stressful life events while controls reported 35. Because controls were older than MDD cases and because the type of events included in the checklist were focused on younger individuals, differences in the frequency of life stress are likely not meaningful. A linear model comparing the number of stressful life events between groups while covarying for age and sex indicated a nonsignificant effect of group ($P=.43$). Notwithstanding this confound, looking at the entire sample EMA scores were higher in individuals that reported life stress, this was true in both controls and MDD cases (Figure S5 in [Multimedia Appendix 1](#)). We examined cross-sectional and lagged relationships between self-reported stress and summed EMA scores across each week (stress was reported once per week) using linear mixed models, including MDD, age, and sex as predictors. Concurrent self-reported stress, occurring in the same week, significantly predicted EMA scores ($\eta^2_p=0.16$, $\beta=24.50$, 95% CI 6.07 to 42.93, $P<.001$), whereas lagged stress (T-1; $\eta^2_p=0.07$, $\beta=6.15$, 95% CI -1.52 to 13.82, $P=.28$) did not, suggesting an immediate rather than cumulative impact of stress on mood and anxiety. MDD ($\eta^2_p=0.16$, $P<.001$) and sex

($\eta_p^2=0.14$, $P<.001$) remained significant predictors, but age ($P=.48$) was not a significant predictor.

Relationship Between MDD and Mood and Activity Levels

Number of steps taken per day (and across the course of the study) ranged widely between participants (Figure S6 in [Multimedia Appendix 1](#)). MDD cases took fewer steps per day (mean 5828.64, SD 6188.85) than controls (mean 7088.47, SD 5378.18) over the course of the study. However, MDD was not a significant predictor of steps taken ($\eta_p^2=0.18$, $P=.33$). We also examined cross-sectional relationships between EMA summed score and steps taken

each day over time using a linear mixed model, steps were not a significant predictor of EMA score ($\eta_p^2=-0.01$, $P=.75$). Thus, activity, indexed using number of steps taken per day, and mood were not significantly related in the present study.

Relationship Between MDD and Mood and Sleep

In both cases and controls, sleep onset was most commonly between 8 PM and 4 AM (Figure S7 in [Multimedia Appendix 1](#)), and sleep offset was most commonly between 5 AM and 11 AM (Figure S8 in [Multimedia Appendix 1](#)), though there was more spread in sleep offset hours in cases. Descriptive statistics of sleep metrics are shown in [Table 4](#).

Table 4. Descriptive statistics of sleep metrics in depressed and control groups.

	Control	MDD ^a
Nights, n	160	232
Individuals, n	11	11
Total sleep time (TST), mean (SD)	388.62 (108.56)	277.19 (163.31)
Time in bed (TIB), mean (SD)	439.84 (124.20)	315.11 (186.48)
Time awake in bed (TAB), mean (SD)	51.12 (20.95)	37.41 (27.57)
Deep sleep		
Mean (SD)	79.24 (23.69)	66.47 (27.90)
Percentage	17	15
Light sleep		
Mean (SD)	264.05 (55.20)	250.49 (82.77)
Percentage	57	58
REM ^b sleep		
Mean (SD)	94.80 (34.44)	83.43 (39.12)
Percentage	20	19
Efficiency (TST/TIB)	0.89	0.89
5 min wakes	390	370
Hypersomnia	2	5
Number of naps	13	88
Coefficient of variation (CoV) of TST	27.93	58.92
Mode onset hour	23:00	23:00
SD onset time (minutes from 12 PM)	123.54	202.99
Mode offset hour	7:00	06:00
SD offset time (minutes from 12 PM)	113.48	166.14

^aMDD: major depressive disorder

^bREM: rapid eye movement.

In terms of sleep staging, an inspection of [Table 4](#) shows that the amount of deep, light, and REM sleep between cases and controls did not differ, nor did sleep efficiency (how much TIB was spent asleep). However, there were differences in (1) TST, (2) number of naps, and (3) variability in sleep length. We tested differences between cases and controls using linear mixed models, and age and sex were included as covariates in all models. Cases spent significantly less time asleep (TST; $\eta_p^2=0.20$, $\beta=-94.38$, 95% CI -178.62 to -10.14 , $P<.001$) and TIB ($\eta_p^2=0.18$, $\beta=-103.39$, 95% CI -199.85 to

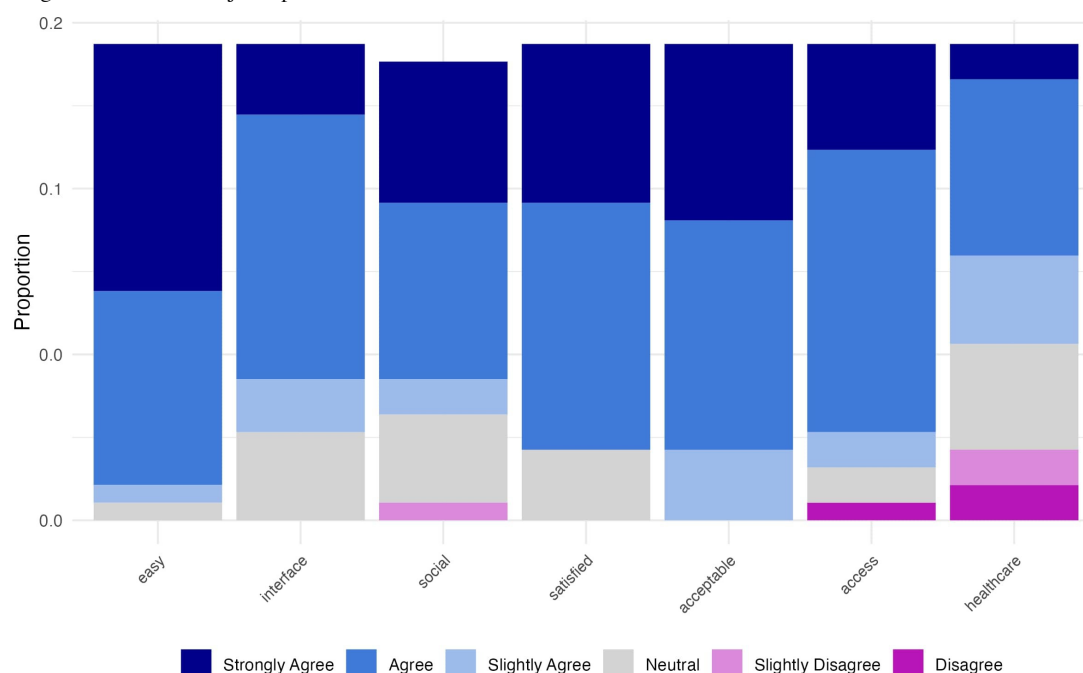
-6.94 , $P<.001$). Cases took significantly more naps ($\beta=14.05$, 95% CI 2.32 to 85.02, $P<.001$). Cases had significantly more variability in their TST (the length of each sleep was more inconsistent than that taken by controls) as indexed by the CoV of TST ($\eta_p^2=0.53$, $\beta=32.14$, 95% CI 16.04 to 48.23, $P<.001$). We examined the relationship between TST and EMA data using a linear mixed model. The effect of TST on EMA score the following day was not significant ($P=.69$).

Did Participants Find Using the App to Report on Their Mood Acceptable?

Figure 4 shows the results of the usability questionnaire. Inspection of Figure 4 shows that the majority agreed that the EMA app was easy to use (96%) and that they were comfortable using the app in social settings (77%); participants reported using the app in a wide variety of places (Table S5

in [Multimedia Appendix 1](#) contains a complete list of unique responses). A total of 100% of individuals found the app an acceptable way to report on their mood, 89% of them said they felt the app improved their access to research participation, and 63% of individuals agreed that they would like to use an app like this to communicate with their health care provider.

Figure 4. Usability data regarding the ecological momentary assessment (EMA) app. The majority of participants agreed that the app was “easy” to use, that they liked the “interface” of the app, that they felt comfortable using the app in “social” settings, that they were “satisfied” with the app, and that the app provided them with an “acceptable” way to report on their mood. A small minority of participants disagreed that the app increased their “access” to research participation (N=1) and that they would like to use a similar app to communicate with their “healthcare” provider (N=4), but the majority were in agreement. MDD: major depressive disorder.



Discussion

The present study comprehensively assessed the feasibility of smartphone-based EMA in adolescents with depression by (1) evaluating participant retention and engagement, (2) examining whether mood reports captured expected symptom variation, and (3) determining the acceptability of the app-based approach. Together, these findings demonstrate that smartphone-based EMA is a feasible, viable, and effective method for real-time symptom monitoring in adolescents with depression.

Engagement with EMA was high, with exceptional compliance rates across participants. This suggests that adolescents were able and willing to complete frequent mood assessments over an extended period, reinforcing the feasibility of implementing this method in both research and clinical settings. The usability data indicate that participants found the app easy and satisfying to use, that they used the app wherever they were, including in social settings, and that most participants would be amenable to using such an app in a clinical setting. These responses indicate that adolescents find smartphone-based apps for mood reporting convenient

and further suggest that approaches using them, whether that be in the context of research or in the clinic, are likely to be well tolerated.

Beyond high compliance and self-reported acceptability, feasibility was further supported by the clinical distinctions captured in the collected data. Adolescents with depression reported significantly more severe and variable symptoms than healthy controls, reinforcing EMA's ability to differentiate between clinical and nonclinical populations. This supports the discriminant validity of our approach, further reinforcing its utility and feasibility as a real-time assessment tool. The EMA scores were significantly predicted by baseline PHQ-9 scores, supporting its criterion validity against an established depression measure. While this alignment was expected given that EMA items were adapted from the PHQ-9 and GAD-7, it is notable that these adaptations retained the ability of those items to reflect depression severity while also being suitable for repeated, momentary assessment. Furthermore, individual-level trajectories revealed that while some participants experienced worsening symptoms, others improved. These findings underscore EMA's ability to capture dynamic mood changes [9,22-24,65] that would be difficult to assess with retrospective measures

[14-19]. Notably, some healthy controls also reported transient depressive symptoms, underscoring the importance of including control groups in EMA studies and highlighting the potential for identifying individuals with emerging depressive episodes. Interestingly, increased life stress was associated with greater mood and anxiety symptom severity across both groups. The effect of stress on EMA scores was observed concurrently, but not lagged, indicating an immediate rather than cumulative effect of stress on mood and anxiety symptoms.

Activity data revealed that MDD cases spent less time asleep and in bed but took more naps than controls, suggesting that their sleep was more fragmented and variable than in controls. These findings are consistent with previous research showing an association between sleep disturbances and depression and anxiety [66,67], particularly short sleep duration and symptoms of depressed mood [68] and increased napping in adults [69,70] and adolescents [71] with depression. Although we did not find a significant relationship between daily mood variations and the previous night's sleep or daily physical activity levels, our study supports the notion that sleep disturbances are a critical component of adolescent depression [72].

The study should also be viewed in the context of some important limitations. First, our sample size limits the generalizability of our results. The sample size is small, and so future replication of our main findings is key. Second, the racial homogeneity and relatively high Socioeconomic Status of our sample, with 67% of participants identifying as White, may limit the generalizability of our findings to more diverse populations. Third, there is an age disparity between the studied groups, with the controls being significantly older on average than the MDD cases. This could have introduced an age-related bias, particularly for age-specific assessments like the life stress questionnaire. Fourth, while fitbit is a cost-effective way to collect sleep and activity information, it may overestimate sleep time (particularly sleep staging estimates [73,74] and activity levels [75] compared with gold standard methods (eg, polysomnography for sleep). We did not find differences between cases and controls in

terms of number of steps taken or in terms of sleep staging; these might be true findings, or the lack of difference might be clouded by the measurement error associated with Fitbit devices. Future research should aim to include a larger and more age-matched sample using alternate devices to mitigate these limitations.

When considering the clinical applications of EMA, it is important to acknowledge that this study was conducted as a paid research project. As a result, the high compliance rates observed may have been influenced by financial compensation, which does not necessarily translate to similar real-world clinical settings. Additionally, the implementation of EMA in clinical practice presents several challenges that must be addressed. These include logistical barriers, such as ensuring access to the necessary technology, as well as concerns about privacy and data confidentiality, particularly when working with vulnerable populations like children.

The findings of this study have strong implications for both research and, potentially, clinical practice. From a research perspective, this study is one of the first to explore the integration of smartphone-based EMA and wearable activity tracking in adolescents with depression. The findings highlight the potential of these technologies to capture real-time data and provide valuable insights into the dynamic nature of depressive symptoms. Future studies should explore the use of EMA and activity-tracking methods in a larger and more diverse population to validate our findings and ultimately expand this novel method to clinical practices. In terms of clinical practice, the demonstrated feasibility of EMA and activity tracking in adolescents suggests that these methods could be integrated into routine clinical assessments to provide a more comprehensive and dynamic picture of symptomatology. Moreover, the development of personalized intervention strategies that take into account individual patterns of symptom fluctuations could lead to more effective management of adolescent depression [57]. A more comprehensive understanding of the external factors influencing symptom variability would provide insight into this relationship, particularly by looking at external factors such as environmental stressors and social interactions.

Acknowledgments

This work was supported by a Tommy Fuss Independent Investigator Award (grant Number: 92188)

Data Availability

The data supporting the findings of this study are available upon request. Interested researchers may contact the corresponding author at emma.knowles@childrens.harvard.edu for access.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Additional methodological details and supplementary materials related to the main analyses.

[\[DOCX File \(Microsoft Word File\), 2172 KB-Multimedia Appendix 1\]](#)

References

1. Major depression. National Institute of Mental Health (NIMH). URL: <https://www.nimh.nih.gov/health/statistics/major-depression> [Accessed 2023-06-14]

2. Twenge JM, Cooper AB, Joiner TE, Duffy ME, Binau SG. Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005-2017. *J Abnorm Psychol.* Apr 2019;128(3):185-199. [doi: [10.1037/abn0000410](https://doi.org/10.1037/abn0000410)] [Medline: [30869927](https://pubmed.ncbi.nlm.nih.gov/30869927/)]
3. Data and statistics on children's mental health. CDC: Centers for Disease Control and Prevention. URL: <https://www.cdc.gov/children-mental-health/data-research/index.html> [Accessed 2024-07-17]
4. Clayborne ZM, Varin M, Colman I. Systematic review and meta-analysis: adolescent depression and long-term psychosocial outcomes. *J Am Acad Child Adolesc Psychiatry.* Jan 2019;58(1):72-79. [doi: [10.1016/j.jaac.2018.07.896](https://doi.org/10.1016/j.jaac.2018.07.896)] [Medline: [30577941](https://pubmed.ncbi.nlm.nih.gov/30577941/)]
5. Fletcher JM. Adolescent depression and educational attainment: results using sibling fixed effects. *Health Econ.* Jul 2010;19(7):855-871. [doi: [10.1002/hec.1526](https://doi.org/10.1002/hec.1526)] [Medline: [19582699](https://pubmed.ncbi.nlm.nih.gov/19582699/)]
6. Rao U, Hammen C, Daley SE. Continuity of depression during the transition to adulthood: a 5-year longitudinal study of young women. *J Am Acad Child Adolesc Psychiatry.* Jul 1999;38(7):908-915. [doi: [10.1097/00004583-199907000-00022](https://doi.org/10.1097/00004583-199907000-00022)] [Medline: [10405510](https://pubmed.ncbi.nlm.nih.gov/10405510/)]
7. Wilson S, Dumornay NM. Rising rates of adolescent depression in the United States: challenges and opportunities in the 2020s. *J Adolesc Health.* Mar 2022;70(3):354-355. [doi: [10.1016/j.jadohealth.2021.12.003](https://doi.org/10.1016/j.jadohealth.2021.12.003)] [Medline: [35183317](https://pubmed.ncbi.nlm.nih.gov/35183317/)]
8. Copeland WE, Alaie I, Jonsson U, Shanahan L. Associations of childhood and adolescent depression with adult psychiatric and functional outcomes. *J Am Acad Child Adolesc Psychiatry.* May 2021;60(5):604-611. [doi: [10.1016/j.jaac.2020.07.895](https://doi.org/10.1016/j.jaac.2020.07.895)] [Medline: [32758528](https://pubmed.ncbi.nlm.nih.gov/32758528/)]
9. Fried EI. Moving forward: how depression heterogeneity hinders progress in treatment and research. *Expert Rev Neurother.* May 2017;17(5):423-425. [doi: [10.1080/14737175.2017.1307737](https://doi.org/10.1080/14737175.2017.1307737)] [Medline: [28293960](https://pubmed.ncbi.nlm.nih.gov/28293960/)]
10. Schoevers RA, van Borkulo CD, Lamers F, et al. Affect fluctuations examined with ecological momentary assessment in patients with current or remitted depression and anxiety disorders. *Psychol Med.* Aug 2021;51(11):1906-1915. [doi: [10.1017/S0033291720000689](https://doi.org/10.1017/S0033291720000689)] [Medline: [32234092](https://pubmed.ncbi.nlm.nih.gov/32234092/)]
11. Clasen PC, Fisher AJ, Beevers CG. Mood-reactive self-esteem and depression vulnerability: person-specific symptom dynamics via smart phone assessment. *PLoS ONE.* 2015;10(7):e0129774. [doi: [10.1371/journal.pone.0129774](https://doi.org/10.1371/journal.pone.0129774)] [Medline: [26131724](https://pubmed.ncbi.nlm.nih.gov/26131724/)]
12. Hosenfeld B, Bos EH, Wardenaar KJ, et al. Major depressive disorder as a nonlinear dynamic system: bimodality in the frequency distribution of depressive symptoms over time. *BMC Psychiatry.* Sep 18, 2015;15:222. [doi: [10.1186/s12888-015-0596-5](https://doi.org/10.1186/s12888-015-0596-5)] [Medline: [26385384](https://pubmed.ncbi.nlm.nih.gov/26385384/)]
13. Ben-Zeev D, Young MA. Accuracy of hospitalized depressed patients' and healthy controls' retrospective symptom reports: an experience sampling study. *J Nerv Ment Dis.* Apr 2010;198(4):280-285. [doi: [10.1097/NMD.0b013e3181d6141f](https://doi.org/10.1097/NMD.0b013e3181d6141f)] [Medline: [20386257](https://pubmed.ncbi.nlm.nih.gov/20386257/)]
14. Ben-Zeev D, Young MA, Madsen JW. Retrospective recall of affect in clinically depressed individuals and controls. *Cogn Emot.* Aug 2009;23(5):1021-1040. [doi: [10.1080/02699930802607937](https://doi.org/10.1080/02699930802607937)]
15. Ruhe HG, Mocking RJT, Figueroa CA, et al. Emotional biases and recurrence in major depressive disorder. results of 2.5 years follow-up of drug-free cohort vulnerable for recurrence. *Front Psychiatry.* 2019;10:145. [doi: [10.3389/fpsy.2019.00145](https://doi.org/10.3389/fpsy.2019.00145)] [Medline: [30984039](https://pubmed.ncbi.nlm.nih.gov/30984039/)]
16. Bianchi R, Laurent E, Schonfeld IS, Bietti LM, Mayor E. Memory bias toward emotional information in burnout and depression. *J Health Psychol.* Sep 2020;25(10-11):1567-1575. [doi: [10.1177/1359105318765621](https://doi.org/10.1177/1359105318765621)] [Medline: [29600730](https://pubmed.ncbi.nlm.nih.gov/29600730/)]
17. Marsh L, Edgington T, Conway MA, Loveday C. Positivity bias in past and future episodic thinking: relationship with anxiety, depression, and retrieval-induced forgetting. *Q J Exp Psychol (Hove).* Mar 2019;72(3):508-522. [doi: [10.1177/1747021818758620](https://doi.org/10.1177/1747021818758620)] [Medline: [29364056](https://pubmed.ncbi.nlm.nih.gov/29364056/)]
18. Urban EJ, Charles ST, Levine LJ, Almeida DM. Depression history and memory bias for specific daily emotions. *PLoS ONE.* 2018;13(9):e0203574. [doi: [10.1371/journal.pone.0203574](https://doi.org/10.1371/journal.pone.0203574)] [Medline: [30192853](https://pubmed.ncbi.nlm.nih.gov/30192853/)]
19. Russell MA, Gajos JM. Annual research review: ecological momentary assessment studies in child psychology and psychiatry. *J Child Psychol Psychiatry.* Mar 2020;61(3):376-394. [doi: [10.1111/jcpp.13204](https://doi.org/10.1111/jcpp.13204)] [Medline: [31997358](https://pubmed.ncbi.nlm.nih.gov/31997358/)]
20. Armey MF, Schatten HT, Haradthvala N, Miller IW. Ecological momentary assessment (EMA) of depression-related phenomena. *Curr Opin Psychol.* Aug 2015;4:21-25. [doi: [10.1016/j.copsyc.2015.01.002](https://doi.org/10.1016/j.copsyc.2015.01.002)]
21. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annu Rev Clin Psychol.* 2008;4(1):1-32. [doi: [10.1146/annurev.clinpsy.3.022806.091415](https://doi.org/10.1146/annurev.clinpsy.3.022806.091415)] [Medline: [18509902](https://pubmed.ncbi.nlm.nih.gov/18509902/)]
22. Price M, Yuen EK, Goetter EM, et al. mHealth: a mechanism to deliver more accessible, more effective mental health care. *Clin Psychol Psychother.* 2014;21(5):427-436. [doi: [10.1002/cpp.1855](https://doi.org/10.1002/cpp.1855)] [Medline: [23918764](https://pubmed.ncbi.nlm.nih.gov/23918764/)]
23. Beames JR, Kikas K, Werner-Seidler A. Prevention and early intervention of depression in young people: an integrated narrative review of affective awareness and ecological momentary assessment. *BMC Psychol.* Aug 16, 2021;9(1):113. [doi: [10.1186/s40359-021-00614-6](https://doi.org/10.1186/s40359-021-00614-6)] [Medline: [34392830](https://pubmed.ncbi.nlm.nih.gov/34392830/)]

24. Schueller SM, Aguilera A, Mohr DC. Ecological momentary interventions for depression and anxiety. *Depress Anxiety*. Jun 2017;34(6):540-545. [doi: [10.1002/da.22649](https://doi.org/10.1002/da.22649)] [Medline: [28494123](https://pubmed.ncbi.nlm.nih.gov/28494123/)]
25. Baltasar-Tello I, Miguélez-Fernández C, Peñuelas-Calvo I, Carballo JJ. Ecological momentary assessment and mood disorders in children and adolescents: a systematic review. *Curr Psychiatry Rep*. Aug 1, 2018;20(8):66. [doi: [10.1007/s11920-018-0913-z](https://doi.org/10.1007/s11920-018-0913-z)] [Medline: [30069650](https://pubmed.ncbi.nlm.nih.gov/30069650/)]
26. Yim SJ, Lui LMW, Lee Y, et al. The utility of smartphone-based, ecological momentary assessment for depressive symptoms. *J Affect Disord*. Sep 1, 2020;274:602-609. [doi: [10.1016/j.jad.2020.05.116](https://doi.org/10.1016/j.jad.2020.05.116)] [Medline: [32663993](https://pubmed.ncbi.nlm.nih.gov/32663993/)]
27. Axelson DA, Bertocci MA, Lewin DS, et al. Measuring mood and complex behavior in natural environments: use of ecological momentary assessment in pediatric affective disorders. *J Child Adolesc Psychopharmacol*. 2003;13(3):253-266. [doi: [10.1089/104454603322572589](https://doi.org/10.1089/104454603322572589)] [Medline: [14642013](https://pubmed.ncbi.nlm.nih.gov/14642013/)]
28. Cousins JC, Whalen DJ, Dahl RE, et al. The bidirectional association between daytime affect and nighttime sleep in youth with anxiety and depression. *J Pediatr Psychol*. Oct 2011;36(9):969-979. [doi: [10.1093/jpepsy/jsr036](https://doi.org/10.1093/jpepsy/jsr036)] [Medline: [21795377](https://pubmed.ncbi.nlm.nih.gov/21795377/)]
29. Forbes EE, Hariri AR, Martin SL, et al. Altered striatal activation predicting real-world positive affect in adolescent major depressive disorder. *AJP*. Jan 2009;166(1):64-73. [doi: [10.1176/appi.ajp.2008.07081336](https://doi.org/10.1176/appi.ajp.2008.07081336)]
30. Forbes EE, Stepp SD, Dahl RE, et al. Real-world affect and social context as predictors of treatment response in child and adolescent depression and anxiety: an ecological momentary assessment study. *J Child Adolesc Psychopharmacol*. Feb 2012;22(1):37-47. [doi: [10.1089/cap.2011.0085](https://doi.org/10.1089/cap.2011.0085)] [Medline: [22339611](https://pubmed.ncbi.nlm.nih.gov/22339611/)]
31. Primack BA, Silk JS, DeLozier CR, et al. Using ecological momentary assessment to determine media use by individuals with and without major depressive disorder. *Arch Pediatr Adolesc Med*. Apr 2011;165(4):360-365. [doi: [10.1001/archpediatrics.2011.27](https://doi.org/10.1001/archpediatrics.2011.27)] [Medline: [21464384](https://pubmed.ncbi.nlm.nih.gov/21464384/)]
32. Silk JS, Dahl RE, Ryan ND, et al. Pupillary reactivity to emotional information in child and adolescent depression: links to clinical and ecological measures. *AJP*. Dec 2007;164(12):1873-1880. [doi: [10.1176/appi.ajp.2007.06111816](https://doi.org/10.1176/appi.ajp.2007.06111816)]
33. Silk JS, Forbes EE, Whalen DJ, et al. Daily emotional dynamics in depressed youth: a cell phone ecological momentary assessment study. *J Exp Child Psychol*. Oct 2011;110(2):241-257. [doi: [10.1016/j.jecp.2010.10.007](https://doi.org/10.1016/j.jecp.2010.10.007)] [Medline: [21112595](https://pubmed.ncbi.nlm.nih.gov/21112595/)]
34. Waller JM, Silk JS, Stone LB, Dahl RE. Co-rumination and co-problem solving in the daily lives of adolescents with major depressive disorder. *J Am Acad Child Adolesc Psychiatry*. Aug 2014;53(8):869-878. [doi: [10.1016/j.jaac.2014.05.004](https://doi.org/10.1016/j.jaac.2014.05.004)] [Medline: [25062594](https://pubmed.ncbi.nlm.nih.gov/25062594/)]
35. Whalen DJ, Silk JS, Semel M, et al. Caffeine consumption, sleep, and affect in the natural environments of depressed youth and healthy controls. *J Pediatr Psychol*. May 2008;33(4):358-367. [doi: [10.1093/jpepsy/jsm086](https://doi.org/10.1093/jpepsy/jsm086)] [Medline: [17947257](https://pubmed.ncbi.nlm.nih.gov/17947257/)]
36. Frost A, Hoyt LT, Chung AL, Adam EK. Daily life with depressive symptoms: gender differences in adolescents' everyday emotional experiences. *J Adolesc*. Aug 2015;43:132-141. [doi: [10.1016/j.adolescence.2015.06.001](https://doi.org/10.1016/j.adolescence.2015.06.001)] [Medline: [26107805](https://pubmed.ncbi.nlm.nih.gov/26107805/)]
37. Mor N, Doane LD, Adam EK, et al. Within-person variations in self-focused attention and negative affect in depression and anxiety: a diary study. *Cogn Emot*. Jan 2010;24(1):48-62. [doi: [10.1080/02699930802499715](https://doi.org/10.1080/02699930802499715)]
38. Sequeira L, Battaglia M, Perrotta S, Merikangas K, Strauss J. Digital phenotyping with mobile and wearable devices: advanced symptom measurement in child and adolescent depression. *J Am Acad Child Adolesc Psychiatry*. Sep 2019;58(9):841-845. [doi: [10.1016/j.jaac.2019.04.011](https://doi.org/10.1016/j.jaac.2019.04.011)] [Medline: [31445619](https://pubmed.ncbi.nlm.nih.gov/31445619/)]
39. Minich M, Moreno M. Real-world adolescent smartphone use is associated with improvements in mood: an ecological momentary assessment study. *PLoS ONE*. 2024;19(5):e0298422. [doi: [10.1371/journal.pone.0298422](https://doi.org/10.1371/journal.pone.0298422)] [Medline: [38809836](https://pubmed.ncbi.nlm.nih.gov/38809836/)]
40. Bickham DS, Hsuen Y, Rich M. Media use and depression: exposure, household rules, and symptoms among young adolescents in the USA. *Int J Public Health*. Feb 2015;60(2):147-155. [doi: [10.1007/s00038-014-0647-6](https://doi.org/10.1007/s00038-014-0647-6)] [Medline: [25586816](https://pubmed.ncbi.nlm.nih.gov/25586816/)]
41. Cushing CC, Marker AM, Bejarano CM, Crick CJ, Huffhines LP. Latent variable mixture modeling of ecological momentary assessment data: implications for screening and adolescent mood profiles. *J Child Fam Stud*. Jun 2017;26(6):1565-1572. [doi: [10.1007/s10826-017-0689-5](https://doi.org/10.1007/s10826-017-0689-5)]
42. First MB. Structured clinical interview for the DSM (SCID). In: Cautin RL, Lilienfeld SO, editors. *The Encyclopedia of Clinical Psychology*. Wiley-Blackwell; 2015:1-6. [doi: [10.1002/9781118625392](https://doi.org/10.1002/9781118625392)]
43. Geller B, Zimmerman B, Williams M, et al. Reliability of the Washington University in St. Louis kiddie schedule for affective disorders and schizophrenia (WASH-U-KSADS) mania and rapid cycling sections. *J Am Acad Child Adolesc Psychiatry*. Apr 2001;40(4):450-455. [doi: [10.1097/00004583-200104000-00014](https://doi.org/10.1097/00004583-200104000-00014)] [Medline: [11314571](https://pubmed.ncbi.nlm.nih.gov/11314571/)]
44. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. Sep 2001;16(9):606-613. [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]

45. Overall JE, Gorham DR. The brief psychiatric rating scale. *Psychol Rep.* Jun 1962;10(3):799-812. [doi: [10.2466/pr0.1962.10.3.799](https://doi.org/10.2466/pr0.1962.10.3.799)]
46. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med.* May 22, 2006;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]
47. Buysse DJ, Reynolds CF III, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry Res.* May 1989;28(2):193-213. [doi: [10.1016/0165-1781\(89\)90047-4](https://doi.org/10.1016/0165-1781(89)90047-4)] [Medline: [2748771](https://pubmed.ncbi.nlm.nih.gov/2748771/)]
48. Hollingshead A. Four factor index of social status. Yale University; 1975.
49. LifeData: experience sampling app for research. LifeData Corp. URL: <https://www.lifedatacorp.com> [Accessed 2024-07-17]
50. Fitbit web API data dictionary. fitbit. 2023. URL: <https://enterprise.fitbit.com/wp-content/uploads/Fitbit-Web-API-Data-Dictionary-Downloadable-Version-2023.pdf> [Accessed 2025-02-04]
51. Pandas. The pandas development team. Zenodo; 2020.
52. PyCharm. JetBrains. 2024. URL: <https://www.jetbrains.com/pycharm/download/?section=windows> [Accessed 2025-04-30]
53. Wickham H, Vaughan D, Girlich M. Tidy messy data. tidy; 2023.
54. Wickham H, François R, Henry L, Vaughan D. A grammar of data manipulation. dplyr. 2023. URL: <https://dplyr.tidyverse.org/> [Accessed 2025-04-30]
55. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; 2017.
56. Zhang Y, Folarin AA, Sun S, et al. Relationship between major depression symptom severity and sleep collected using a wristband wearable device: multicenter longitudinal observational study. *JMIR Mhealth Uhealth.* Apr 12, 2021;9(4):e24604. [doi: [10.2196/24604](https://doi.org/10.2196/24604)] [Medline: [33843591](https://pubmed.ncbi.nlm.nih.gov/33843591/)]
57. Kleiman EM, Turner BJ, Fedor S, Beale EE, Huffman JC, Nock MK. Examination of real-time fluctuations in suicidal ideation and its risk factors: results from two ecological momentary assessment studies. *J Abnorm Psychol.* Aug 2017;126(6):726-738. [doi: [10.1037/abn0000273](https://doi.org/10.1037/abn0000273)] [Medline: [28481571](https://pubmed.ncbi.nlm.nih.gov/28481571/)]
58. Wickham H. Ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York; 2016. [doi: [10.1007/978-3-319-24277-4_9](https://doi.org/10.1007/978-3-319-24277-4_9)]
59. von Neumann J, Kent RH, Bellinson HR, Hart BI. The mean square successive difference. *Ann Math Statist.* Jun 1941;12(2):153-162. [doi: [10.1214/aoms/1177731746](https://doi.org/10.1214/aoms/1177731746)]
60. Brown CE. Coefficient of variation. In: *Applied Multivariate Statistics in Geohydrology and Related Sciences*. Springer; 1998:155-157. [doi: [10.1007/978-3-642-80328-4](https://doi.org/10.1007/978-3-642-80328-4)]
61. Revelle W. psych: procedures for psychological, psychometric, and personality research. R package version 1.0–95. Psychology (Irvine). 2013.
62. Bates D, Maechler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw.* 2015;67:1-48. [doi: [10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01)]
63. Ben-Shachar M, Lüdtke D, Makowski D. Effectsize: estimation of effect size indices and standardized parameters. *JOSS.* 2020;5(56):2815. [doi: [10.21105/joss.02815](https://doi.org/10.21105/joss.02815)]
64. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest package: tests in linear mixed effects models. *J Stat Soft.* 2017;82(13). [doi: [10.18637/jss.v082.i13](https://doi.org/10.18637/jss.v082.i13)]
65. Fried EI, Nesse RM. Depression is not a consistent syndrome: an investigation of unique symptom patterns in the STAR*D study. *J Affect Disord.* Feb 1, 2015;172:96-102. [doi: [10.1016/j.jad.2014.10.010](https://doi.org/10.1016/j.jad.2014.10.010)] [Medline: [25451401](https://pubmed.ncbi.nlm.nih.gov/25451401/)]
66. Alvaro PK, Roberts RM, Harris JK. A systematic review assessing bidirectionality between sleep disturbances, anxiety, and depression. *Sleep.* Jul 1, 2013;36(7):1059-1068. [doi: [10.5665/sleep.2810](https://doi.org/10.5665/sleep.2810)] [Medline: [23814343](https://pubmed.ncbi.nlm.nih.gov/23814343/)]
67. Gradisar M, Kahn M, Micic G, et al. Sleep's role in the development and resolution of adolescent depression. *Nat Rev Psychol.* 2022;1(9):512-523. [doi: [10.1038/s44159-022-00074-8](https://doi.org/10.1038/s44159-022-00074-8)] [Medline: [35754789](https://pubmed.ncbi.nlm.nih.gov/35754789/)]
68. Roberts RE, Duong HT. The prospective association between sleep deprivation and depression among adolescents. *Sleep.* Feb 1, 2014;37(2):239-244. [doi: [10.5665/sleep.3388](https://doi.org/10.5665/sleep.3388)] [Medline: [24497652](https://pubmed.ncbi.nlm.nih.gov/24497652/)]
69. Li L, Zhang Q, Zhu L, et al. Daytime naps and depression risk: a meta-analysis of observational studies. *Front Psychol.* 2022;13:1051128. [doi: [10.3389/fpsyg.2022.1051128](https://doi.org/10.3389/fpsyg.2022.1051128)]
70. Foley DJ, Vitiello MV, Bliwise DL, Ancoli-Israel S, Monjan AA, Walsh JK. Frequent napping is associated with excessive daytime sleepiness, depression, pain, and nocturia in older adults: findings from the national sleep foundation “2003 Sleep in America” poll. *Am J Geriatr Psychiatry.* Apr 2007;15(4):344-350. [doi: [10.1097/01.JGP.0000249385.50101.67](https://doi.org/10.1097/01.JGP.0000249385.50101.67)] [Medline: [17384317](https://pubmed.ncbi.nlm.nih.gov/17384317/)]

71. Murray CB, Murphy LK, Palermo TM, Clarke GM. Pain and sleep-wake disturbances in adolescents with depressive disorders. *J Clin Child Adolesc Psychol*. 2012;41(4):482-490. [doi: [10.1080/15374416.2012.658613](https://doi.org/10.1080/15374416.2012.658613)] [Medline: [22420746](https://pubmed.ncbi.nlm.nih.gov/22420746/)]
72. Daly M. Prevalence of depression among adolescents in the U.S. from 2009 to 2019: analysis of trends by sex, race/ethnicity, and income. *J Adolesc Health*. Mar 2022;70(3):496-499. [doi: [10.1016/j.jadohealth.2021.08.026](https://doi.org/10.1016/j.jadohealth.2021.08.026)] [Medline: [34663534](https://pubmed.ncbi.nlm.nih.gov/34663534/)]
73. Haghayegh S, Khoshnevis S, Smolensky MH, Diller KR, Castriotta RJ. Accuracy of wristband Fitbit models in assessing sleep: systematic review and meta-analysis. *J Med Internet Res*. Nov 28, 2019;21(11):e16273. [doi: [10.2196/16273](https://doi.org/10.2196/16273)] [Medline: [31778122](https://pubmed.ncbi.nlm.nih.gov/31778122/)]
74. Renerts K, Doheny EP, Baumann C, et al. F63 validation of fitbit charge 4 for sleep monitoring in participants with huntington's disease. Presented at: EHDN 2022 Plenary Meeting; Sep 2022; Bologna, Italy. [doi: [10.1136/jnnp-2022-ehdn.154](https://doi.org/10.1136/jnnp-2022-ehdn.154)]
75. Feehan LM, Goldman J, Sayre EC, et al. Accuracy of Fitbit devices: systematic review and narrative syntheses of quantitative data. *JMIR Mhealth Uhealth*. Aug 9, 2018;6(8):e10527. [doi: [10.2196/10527](https://doi.org/10.2196/10527)] [Medline: [30093371](https://pubmed.ncbi.nlm.nih.gov/30093371/)]

Abbreviations

CoV: coefficient of variation
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
EMA: ecological momentary assessment
GAD: generalized anxiety disorder
ICC: intraclass correlation
MDD: major depressive disorder
PHQ: patient health questionnaire
REM: rapid eye movement
RMSSD: root mean square of successive differences
TIB: time in bed
TST: total sleep time

Edited by Amaryllis Mavragani; peer-reviewed by David Benrimoh; submitted 09.09.2024; final revised version received 10.02.2025; accepted 11.02.2025; published 11.06.2025

Please cite as:

Unzueta Saavedra J, Deaso EA, Austin M, Cadavid L, Kraff R, Knowles EEM
Smartphone Ecological Momentary Assessment and Wearable Activity Tracking in Pediatric Depression: Cohort Study
JMIR Form Res 2025;9:e66187
 URL: <https://formative.jmir.org/2025/1/e66187>
 doi: [10.2196/66187](https://doi.org/10.2196/66187)

©Jimena Unzueta Saavedra, Emma A Deaso, Margot Austin, Laura Cadavid, Rachel Kraff, Emma E M Knowles. Originally published in *JMIR Formative Research* (<https://formative.jmir.org>), 11.06.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Formative Research*, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.