

RESEARCH

Open Access



Predicting suicidal and self-harm ideation using ecological momentary assessment: deep learning analysis in a general population sample

Heeyeon Kim^{1,2} , Seok-Jae Heo³ , Sehwan Park⁴ , Jooho Lee⁴ , Gangho Do⁴  and Jin Young Park^{1,2,5*} 

Abstract

Background Suicidal and self-harm ideation are major risk factors for suicide but are often difficult to detect, particularly in non-clinical populations. Ecological Momentary Assessment (EMA) offers a real-time, low-burden method for monitoring psychological states, yet its predictive value outside clinical settings remains unclear.

Objective To evaluate whether brief, indirect daily EMA data collected via a smartphone app can predict suicidal and self-harm ideation two weeks later in a general population sample.

Methods A total of 499 adults in Korea completed 28 days of EMA using the BIG4+ app, reporting on seven daily items related to mood, sleep, appetite, concentration, fatigue, and loneliness. Suicidal and self-harm ideation were assessed using the CESD-R at baseline, 2 weeks, and 4 weeks. A recurrent neural network with Long Short-Term Memory (LSTM) architecture was trained on two-week EMA sequences, using 10-fold cross-validation.

Results The combined model using EMA and baseline data achieved an AUC of 0.873 for suicidal ideation and 0.821 for self-harm ideation. Predictive accuracy exceeded an AUC of 0.75 by day 6. Participants with ideation consistently showed lower scores on all EMA items. The study achieved a 94% compliance rate.

Conclusions Brief, indirect EMA data can predict near-term suicidal and self-harm ideation in a general population. These findings support the feasibility of smartphone-based EMA as a scalable and non-intrusive tool for early detection of suicide risk.

Clinical trial number Not applicable.

Keywords Ecological momentary assessment, Suicidal ideation, Self-Harm, Digital mental health, Deep learning, Recurrent neural network, Community sample

*Correspondence:

Jin Young Park
empathy@yuhs.ac

¹Department of Psychiatry, Yongin Severance Hospital, Yonsei University College of Medicine, Yongin, Republic of Korea

²Institute of Behavioral Science in Medicine, Yonsei University College of Medicine, Yonsei University Health System, Seoul, Republic of Korea

³Biostatistics Collaboration Unit, Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, Republic of Korea

⁴Medical Research Team, Digital Medic co., Ltd., Seoul, Republic of Korea

⁵Center for Digital Health, Yongin Severance Hospital, Yonsei University College of Medicine, Yonsei University Health System, Yongin, Republic of Korea



© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Introduction

Suicidal ideation encompasses a spectrum of thoughts, ranging from passive desires for death to active plans for self-harm, and represents a significant global mental health issue. Passive suicidal ideation, such as wishing to die without actionable intent, often coexists with active suicidal ideation, which involves planning or intent to end one's life. The prevalence of suicidal ideation is alarmingly high, with approximately 5.8% of individuals in the general population experiencing passive suicidal ideation annually and lifetime prevalence rates reaching up to 10.6% [1]. Given the critical role of suicidal ideation in suicide prevention, mental health guidelines recommend using open-ended questions to assess suicidal thoughts. However, research indicates that many individuals feel uncomfortable disclosing their suicidal ideation [2], while others struggle to articulate their thoughts [3, 4]. This challenge is particularly problematic in suicide risk assessments, as individuals may deliberately conceal their thoughts to avoid intervention or hospitalization. Consequently, these barriers emphasize the urgent need for effective predictive tools that account for both verbalized and unexpressed suicidal ideation. As a well-established predictor of suicide attempts, suicidal ideation necessitates timely detection and targeted interventions. While traditional face-to-face assessments remain essential, non-face-to-face methods and other innovative suicide risk assessment strategies could serve as valuable complementary tools, enhancing accessibility and early detection efforts. Despite the recognized importance of suicidal prediction, progress in this area has been limited. A seminal meta-analysis by Franklin et al. revealed that our ability to predict suicidal behavior has stagnated over the past 50 years, with predictive accuracy often comparable to chance [5].

Self-harm, defined as the deliberate injury to oneself without suicidal intent, is an escalating public health concern, particularly among adolescents and young adults. Recent studies indicate a significant rise in self-harm behaviors, with emergency department visits for self-inflicted injuries increasing among young people. Although self-harm is often used as a maladaptive strategy for emotional regulation [6], it is also closely associated with impulsivity. Many individuals who engage in self-harm report difficulty articulating the precise reasons for their actions, complicating both assessment and intervention [7]. Despite its high prevalence and clinical significance, the accurate prediction of self-harm remains challenging. A systematic review and meta-analysis examining clinicians' ability to predict future self-harm revealed low predictive accuracy, with a pooled sensitivity of approximately 31% and a positive predictive value of only 22% [8]. Traditional assessments, which often rely on retrospective self-reports, may not capture the

transient and dynamic nature of self-harm urges [9]. This suggests that current clinical assessments may not effectively identify individuals at risk, underscoring the need for improved predictive methodologies.

Recently, smartphone-based solutions have emerged to monitor suicide risk, with the aim of detecting, in real time, the potential suicidal gesture within a short period [10]. These technologies are widely available and easily leveraged to collect real-time ecological momentary assessment (EMA) data [9]. EMA involves repeated sampling of an individual's behaviors and experiences in real time, in the person's natural environment. Patients are prompted to enter information into their smartphone at specific time intervals based on the type of assessment conducted. Unlike retrospective assessments, EMA minimizes recall bias and captures temporal dynamics. By providing a more accurate and ecologically valid depiction of symptoms, EMA offers greater ecological validity by capturing risk within daily life contexts in which suicide risk fluctuates [11]. Daily monitoring of mood and other symptoms may also improve patient insight and facilitate earlier interventions [12]. The strength of EMA for suicide research lies in its ability to detect nuanced, time-sensitive changes in risk that may be missed by traditional methods.

Indeed, a growing body of literature has explored the application of EMA for predicting suicidal ideation and self-harm, primarily within clinical or high-risk populations [13, 14]. Psychiatric inpatient and outpatient samples have been widely studied using EMA to examine short-term risk patterns. For example, real-time assessments of emotional states such as sadness, tension, and boredom have been shown to predict suicidal ideation in hospitalized patients with major depressive disorder [15]. Among adolescents and young adults, EMA has revealed that self-injurious thoughts frequently co-occur with negative affect and urges for other impulsive behaviors [16]. In individuals with borderline personality disorder, heightened mood intensity captured through EMA was linked to increased suicidal thoughts and behaviors [17]. Moreover, studies in high-risk patients recently discharged after a suicide attempt have demonstrated that intensive EMA protocols are both feasible and safe [18]. These findings underscore the predictive potential of EMA in clinical settings, particularly when paired with machine learning techniques capable of detecting patterns associated with imminent risk. Building on this work, several studies have demonstrated that deep learning models—particularly Long Short-Term Memory (LSTM) networks—are well suited for capturing temporal dependencies and affective fluctuations in momentary assessments [19, 20]. Because suicidal ideation and self-harm urges fluctuate over short timescales, these sequential models provide a promising framework for

identifying dynamic risk patterns that may not be detectable with traditional statistical approaches [21]. This prior work highlights the methodological relevance of LSTM-based models for examining short-term suicide risk.

Although prior EMA research has demonstrated feasibility and strong predictive potential in clinical or high-risk samples, it remains unclear whether brief, low-burden EMA assessments can identify suicidal or self-harm ideation in the general population, where such thoughts often go unnoticed due to limited help-seeking. To address this gap, the present study examines whether simple daily EMA responses collected over two weeks can predict suicidal or self-harm ideation assessed two weeks later. We further evaluate whether EMA-derived affective patterns provide incremental predictive value beyond baseline demographic and psychosocial characteristics, and whether predictive performance improves as additional days of EMA data are incorporated. By clarifying the utility of brief EMA protocols in non-clinical samples, this study aims to inform scalable, early identification strategies for suicide prevention.

Methods

Participants and procedures

Eligible participants were aged between 18 and 65 years and resided in the Republic of Korea. Recruitment was conducted through internet-based advertisements between November 2023 and January 2024. To be eligible, participants needed to own a smartphone registered in their name, have the capability to install the BIG4+ mobile app used in this study, and run an operating system version of at least 12 for Android devices or 15.5 for iOS devices. Digital consent forms were obtained prior to participation, and installation instructions were provided by the consumer service team. Participants who encountered technical issues preventing successful app installation were deemed ineligible. Participants received modest compensation for their time and effort upon completion of the study, independent of EMA response content. The study was approved from the Public Institutional Bioethics Committee designated by the Ministry of Health and Welfare of the Korean Government (MOHW; P01-202411-01-038).

Measures

Demographics and suicidal/self-harm ideation

Demographic information collected included ages, sex, education level, and history of psychiatric diagnoses. Additionally, participants provided subjective assessments of presence of a close confidant, family relationships, number of social groups with perceived belonging, and socioeconomic status.

Suicidal and self-harm ideation at baseline were assessed using the Korean version of the Center for Epidemiologic Studies Depression Scale – Revised (K-CESD-R). The presence of suicidal ideation was determined using item 14 (“I wished I were dead”), while self-harm ideation was assessed with item 15 (“I wanted to hurt myself”). Responses were rated on a 0 to 4 scale, where 0 indicated “not at all” and 1 to 4 reflected increasing frequencies from “1 day” to “nearly every day for 2 weeks.” Responses of 0 were categorized as absence of ideation, while scores of 1 or higher were classified as presence of suicidal or self-harm ideation.

The K-CESD-R was administered at baseline and repeated at two-week intervals, resulting in three time-points (baseline, week 2, and week 4). To align with the prediction framework, the 28-day EMA dataset was split into two 14-day segments, with each segment used to predict ideation at the corresponding K-CESD-R time-point (day 14 and day 28). As a result, each participant contributed two time-aligned observations, yielding a total of 998 observations from 499 participants. These observations were categorized based on the presence or absence of suicidal and self-harm ideation and used for further group comparisons.

Ecological momentary assessment (EMA)

Active EMA data were collected using the BIG4+ mobile application, an open-source platform available for Android and iOS. The app administered brief daily self-reports on mental health indicators. All participants responded to four core items (overall mood, appetite, perceived sleep quality, and overall daily experience), and additionally selected three optional items (ability to concentrate, perceived fatigue, and feelings of loneliness), which were fixed for this study. Each item assessed the participant’s state over the past 24 h and was rated on a 7-point Likert scale from 1 (“very bad”) to 7 (“very good”). Higher scores indicated more favorable subjective states. Participants also reported their perceived sleep duration (in hours) for the previous night.

Participants were prompted at a self-selected time each day to complete the EMA. Data were collected for 28 consecutive days. The daily sampling design aimed to capture within-person fluctuations while minimizing recall bias through real-time assessment in participants’ natural environments.

The present study was not designed to model fine-grained within-day affective variability. Rather, it aimed to examine whether low-burden, once-daily EMA assessments could capture cumulative affective patterns associated with suicidal or self-harm ideation at a two-week follow-up. Given the community-based, non-clinical sample and the emphasis on feasibility and sustained

engagement, a once-daily EMA design was selected to balance informational value with participant burden.

Statistical analysis

Continuous variables were reported as means \pm standard deviation (SDs), while categorical variables were expressed as counts and percentages. To compare the differences between two groups based on the presence or absence of suicidal or self-harm ideation, independent two sample t-tests were used for continuous variables and chi-squared tests for categorical variables. These analyses were conducted using Python version 3.8 (Python Software Foundation, Wilmington) and R software (version 4.1.1; R Foundation for Statistical Computing, Vienna, Austria). Statistical significance was set at a p-value of less than 0.05.

Predictive model performance was evaluated using the area under the receiver operating characteristic curve (AUC), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). AUC reflected overall discriminative ability, while sensitivity and specificity indicated the correct identification of individuals with and without suicidal or self-harm ideation. PPV and NPV represented the probabilities that predicted high- or low-risk classifications corresponded to observed ideation status at follow-up.

Overall, missing EMA responses accounted for approximately 6% of all scheduled EMA assessments.

Modeling approach and predictive performance

To evaluate whether daily EMA data could predict near-future ideation, we trained a recurrent neural network using Long Short-Term Memory (LSTM) architecture, chosen for its ability to model time-dependent sequences and learn long-term dependencies in sequential data. We constructed LSTM cells corresponding to 14 days, which receive input from EMA and demographic information at each time point. Each LSTM cell combines information about the current time with information received from the previous LSTM cell and outputs a prediction for the outcome.

Missing EMA values were imputed using linear interpolation, given the low proportion of missing EMA responses. Model performance was evaluated using 10-fold cross-validation. To assess the incremental contribution of EMA data, we compared models trained on: (1) baseline data only (2), EMA data only (3), basic EMA items (4 core variables), and (4) combined baseline and full EMA data. We also examined how predictive performance evolved as additional days of EMA data were incorporated over the 14-day window, by tracking changes in AUC at each step.

Results

Baseline characteristics

Among the 499 participants (mean age 32.79 years, SD 11.12; 342 [68.5%] female), 57 individuals (11.4%) reported a history of psychiatric diagnoses. Over the 28-day EMA period, a total of 13,140 daily EMA responses were collected, yielding a high compliance rate of 94.0%.

EMA data were split into two 14-day segments per participant, resulting in 998 total observations. Among these, 146 were classified as indicating suicidal ideation (mean age 31.90, SD 10.72; 76.7% female), and 75 as indicating self-harm ideation (mean age 30.55, SD 9.48; 77.3% female). Compared to those without ideation, the suicidal ideation group included a significantly higher proportion of females, while the self-harm ideation group was significantly younger on average.

Significant group differences were also found in key psychosocial variables, including the presence of a close confidant, quality of family relationships, number of social groups with perceived belonging, education level (middle school graduate, high school graduate, some college or 2-year degree, university graduate, and graduate-level education), and perceived socioeconomic status. The prevalence of psychiatric diagnoses was highest in the self-harm ideation group (40.0%), followed by the suicidal ideation group (26.0%).

Regarding temporal patterns of ideation, 66.4% of participants who reported suicidal ideation at the two-week follow-up had also reported it at baseline, while 33.6% developed suicidal ideation during the two-week interval. Conversely, 6.1% of those who initially reported suicidal ideation at baseline no longer did so at follow-up. For self-harm ideation, 62.7% of follow-up reports reflected persistence from baseline, while 37.3% represented newly developed ideation. Only 2.9% of individuals showed remission of self-harm ideation over the same interval. Full demographic details are presented in Table 1.

Comparison of two-week EMA data by group

Participants with suicidal or self-harm ideation consistently reported significantly lower scores across all seven EMA items—overall mood, appetite, perceived sleep quality, overall daily experience, ability to concentrate, perceived fatigue, and feelings of loneliness—compared to those without ideation. The group with self-harm ideation exhibited the lowest scores across all domains, including subjective sleep duration, suggesting more profound impairments in daily functioning and well-being (Table 2).

Prediction model performance

Table 3 presents the performance metrics for models predicting suicidal or self-harm ideation at the two-week

Table 1 Demographic characteristics and baseline presence of suicidal and self-harm ideation for 499 participants

	Overall (n = 499, obser- vations = 998)	Suicide ide- ation present (observa- tions = 146)	No sui- cidal ideation (observa- tions = 852)	<i>p</i>	Self-harm ideation pres- ent (observa- tions = 75)	No self-harm ideation (observa- tions = 923)	<i>p</i>
Age	32.79 ± 11.12	31.90 ± 10.72	33.09 ± 11.17	0.220	30.55 ± 9.48	33.10 ± 11.22	0.029
Gender				0.027			0.115
Male	157 (31.5%)	34 (23.3%)	280 (32.9%)		17 (22.7%)	297 (32.2%)	
Female	342 (68.5%)	112 (76.7%)	572 (67.1%)		58 (77.3%)	626 (67.8%)	
Presence of a close confidant				< 0.001			0.045
Yes	417 (83.6%)	106 (72.6%)	728 (85.4%)		56 (74.7%)	778 (84.3%)	
No	82 (16.4%)	40 (27.4%)	124 (14.6%)		19 (25.3%)	145 (15.7%)	
Quality of family relationships				< 0.001			< 0.001
Very poor	4 (0.8%)	4 (2.7%)	4 (0.5%)		4 (5.3%)	4 (0.4%)	
Poor	8 (1.6%)	1 (0.7%)	15 (1.8%)		1 (1.3%)	15 (1.6%)	
Average	131 (26.3%)	66 (45.2%)	196 (23.0%)		27 (36.0%)	235 (25.5%)	
Good	220 (44.1%)	57 (39.0%)	383 (45.0%)		30 (40.0%)	410 (44.4%)	
Very good	136 (27.3%)	18 (12.3%)	254 (29.8%)		13 (17.3%)	259 (28.1%)	
Number of groups with sense of belonging				< 0.001			0.054
None	100 (20.0%)	36 (24.7%)	164 (19.2%)		21 (28.0%)	179 (19.4%)	
1	127 (25.5%)	52 (35.6%)	202 (23.7%)		25 (33.3%)	229 (24.8%)	
2	162 (32.5%)	36 (24.7%)	288 (33.8%)		18 (24.0%)	306 (33.2%)	
3	59 (11.8%)	5 (3.4%)	113 (13.3%)		4 (5.3%)	114 (12.4%)	
4	51 (10.2%)	17 (11.6%)	85 (10.0%)		7 (9.3%)	95 (10.3%)	
Education				0.004			0.002
Middle school graduate	3 (0.6%)	4 (2.7%)	2 (0.2%)		3 (4.0%)	3 (0.3%)	
High school graduate	137 (27.5%)	46 (31.5%)	228 (26.8%)		22 (29.3%)	252 (27.3%)	
Some college or 2-year degree	68 (13.6%)	17 (11.6%)	119 (14.0%)		9 (12.0%)	127 (13.8%)	
University graduate	266 (53.3%)	74 (50.7%)	458 (53.8%)		39 (52.0%)	493 (53.4%)	
Graduate-level education	25 (5.0%)	5 (3.4%)	45 (5.3%)		2 (2.7%)	48 (5.2%)	
Perceived socioeconomic status				< 0.001			0.001
Very low	15 (3.0%)	13 (8.9%)	17 (2.0%)		6 (8.0%)	24 (2.6%)	
Low	136 (27.3%)	50 (34.2%)	222 (26.1%)		30 (40.0%)	242 (26.2%)	
Middle	305 (61.1%)	79 (54.1%)	531 (62.3%)		38 (50.7%)	572 (62.0%)	
High	41 (8.2%)	4 (2.7%)	78 (9.2%)		1 (1.3%)	81 (8.8%)	
Very high	2 (0.2%)	0 (0.0%)	4 (0.5%)		0 (0.0%)	4 (0.4%)	
History of psychiatric diagnosis				< 0.001			< 0.001
Yes	57 (11.4%)	38 (26.0%)	76 (8.9%)		30 (40.0%)	84 (9.1%)	
No	442 (88.6%)	108 (74.0%)	776 (91.1%)		45 (60.0%)	839 (90.9%)	
Baseline suicidal ideation				< 0.001			< 0.001
Yes	149 (14.9%)	97 (66.4%)	52 (6.1%)		54 (72.0%)	95 (10.3%)	
No	849 (85.1%)	49 (33.6%)	800 (93.9%)		21 (28.0%)	828 (89.7%)	
Baseline self-harm ideation				< 0.001			< 0.001
Yes	74 (7.4%)	51 (34.9%)	23 (2.7%)		47 (62.7%)	27 (2.9%)	
No	924 (92.6%)	95 (65.1%)	829 (97.3%)		28 (37.3%)	896 (97.1%)	

follow-up. The models were evaluated based on different sets of predictors, including baseline characteristics (Table 1), the basic four EMA items (mood, appetite, sleep, and general condition), the full set of EMA items, and a combined model incorporating both baseline data and the full set of EMA items.

For suicidal ideation, the combined model achieved the highest accuracy with an AUC of 0.873 (95% CI: 0.854–0.892), significantly outperforming the baseline-only

model (AUC = 0.808, $p < 0.001$). Models using only EMA (AUC = 0.805, 95% CI: 0.770–0.840) or basic EMA items (AUC = 0.784, 95% CI: 0.740–0.828) showed comparable but slightly lower performance. The combined model demonstrated a sensitivity of 0.50 and a specificity of 0.96, indicating a balance between identifying individuals with suicidal ideation and minimizing false-positive classifications. The PPV was 0.797, suggesting that individuals classified as high risk were highly likely to report

Table 2 Two-week EMA data by group based on the presence or absence of suicidal or self-harm ideation

	Overall (<i>n</i> = 499, obser- vations = 998)	Suicide ideation present (observa- tions = 146)	No sui- cidal ideation (observa- tions = 852)	<i>p</i>	Self-harm ideation present (observations = 75)	No self-harm ideation (obser- vations = 923)	<i>p</i>
Overall mood	5.19 ± 1.18	4.10 ± 1.19	5.38 ± 1.07	< 0.001	3.94 ± 1.47	5.29 ± 1.09	< 0.001
Appetite	5.17 ± 1.17	4.22 ± 1.20	5.34 ± 1.08	< 0.001	3.96 ± 1.42	5.27 ± 1.09	< 0.001
Perceived sleep quality	4.95 ± 1.14	4.02 ± 1.12	5.11 ± 1.07	< 0.001	3.84 ± 1.32	5.04 ± 1.08	< 0.001
Sleep duration, hours	7.02 ± 0.93	6.65 ± 1.05	7.08 ± 0.90	< 0.001	6.46 ± 1.18	7.06 ± 0.89	< 0.001
Overall daily experience	5.13 ± 1.14	4.13 ± 1.14	5.30 ± 1.05	< 0.001	3.95 ± 1.38	5.22 ± 1.06	< 0.001
Ability to concentrate	4.97 ± 1.21	3.94 ± 1.12	5.15 ± 1.14	< 0.001	3.78 ± 1.38	5.07 ± 1.14	< 0.001
Perceived fatigue	5.08 ± 1.38	3.98 ± 1.25	5.26 ± 1.32	< 0.001	3.79 ± 1.51	5.18 ± 1.32	< 0.001
Feelings of loneliness	5.40 ± 1.46	4.40 ± 1.48	5.58 ± 1.39	< 0.001	4.20 ± 1.69	5.50 ± 1.39	< 0.001

Table 3 Performance of prediction models for suicidal and Self-Harm ideation at 2-week follow-up

Suicidal ideation	AUC	Sensitivity	Specificity	PPV	NPV	AUC comparison
Baseline	0.808	0.65	0.942	0.659	0.939	reference
EMA (basic 4)	0.784	0.325	0.968	0.852	0.666	0.336
EMA	0.805	0.358	0.962	0.811	0.730	0.898
Baseline + EMA	0.873	0.500	0.96	0.797	0.854	< 0.001
Self-harm ideation	AUC	Sensitivity	Specificity	PPV	NPV	AUC comparison
Baseline	0.788	0.668	0.970	0.601	0.973	reference
EMA (basic 4)	0.780	0.409	0.955	0.692	0.786	0.802
EMA	0.770	0.380	0.960	0.730	0.770	0.676
Baseline + EMA	0.821	0.714	0.980	0.740	0.952	0.162

AUC: area under the curve; PPV: positive predictive value; NPV: negative predictive value; EMA: ecological momentary assessment

suicidal ideation at follow-up, while the NPV of 0.854 indicated reliable identification of individuals at low risk.

For self-harm ideation, the combined model again yielded the highest AUC (0.821, 95% CI: 0.799–0.843), modestly exceeding the baseline-only model (AUC = 0.788), although this difference was not statistically significant ($p = 0.162$). EMA-only and basic EMA models performed similarly (AUCs = 0.770 and 0.780, respectively). The combined model showed a sensitivity of 0.714 and a specificity of 0.980, with a PPV of 0.740 and an NPV of 0.952. These metrics indicate particularly strong performance in ruling out individuals unlikely to report self-harm ideation at follow-up, which is relevant for screening in low-prevalence community samples.

These results underscore the added value of EMA data in predicting suicidal ideation and, to a lesser extent, self-harm ideation. Integration of EMA with baseline characteristics notably improved model performance for suicidal ideation.

Cumulative impact of daily EMA data on predictive accuracy

Figure 1 depicts the progression of AUC values for suicidal ideation as additional days of EMA data were incorporated. Using the basic EMA set (Fig. 1A), AUC values surpassed 0.75 by day 6. With the full EMA set (Fig. 1B), AUC exceeded 0.75 by day 9 and continued to improve, reaching above 0.80 by day 14. This suggests

that incorporating a more comprehensive EMA dataset enhances predictive accuracy over time, with the full EMA set achieving superior performance compared to the basic EMA set.

For self-harm ideation (Fig. 2), the basic EMA model (Fig. 2A) maintained AUC values above 0.75 starting on day 7, peaking at day 11 before slightly declining. The full EMA model (Fig. 2B) showed more variable performance, with fluctuations in AUC across the 14-day period, though generally maintaining predictive capability above the acceptable threshold.

These cumulative analyses suggest that predictive accuracy for suicidal ideation improves steadily with longer EMA input, while predictions for self-harm ideation remain more variable over time.

Discussion

This study demonstrates the utility of short-form daily EMA for predicting suicidal and self-harm ideation in a community sample. Our deep learning model showed high predictive accuracy for suicidal ideation, with performance improving as more days of EMA data were accumulated.

Even brief EMA data—consisting of just seven daily items—enabled the accurate identification of individuals at elevated suicide risk. In addition to prediction performance, descriptive comparisons across EMA variables revealed clinically meaningful group-level differences.

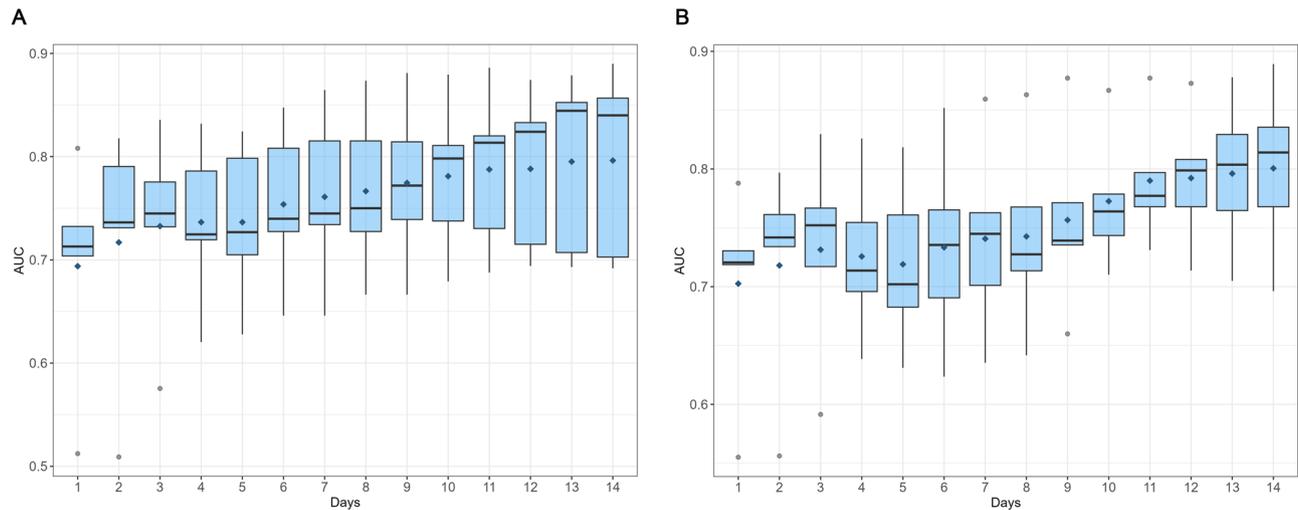


Fig. 1 Change in AUC for predicting suicidal ideation using daily EMA data. Changes in the area under the curve (AUC) as (A) the four basic daily EMA items and (B) the full EMA dataset were cumulatively added to predict ideation over 14 days

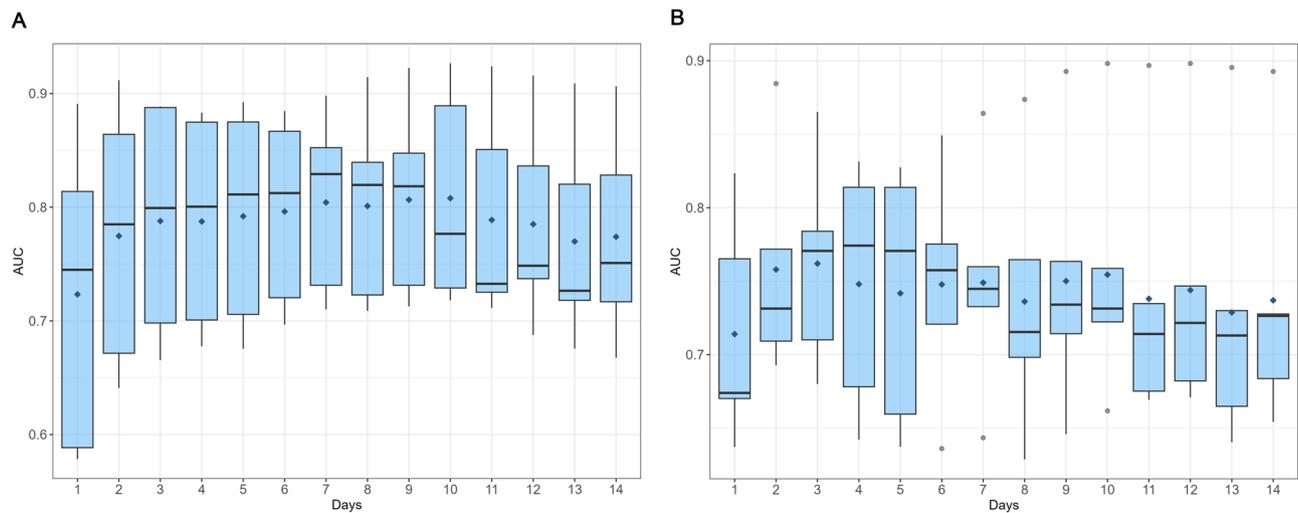


Fig. 2 Change in AUC for predicting self-harm ideation using daily EMA data. Changes in the area under the curve (AUC) as (A) the four basic daily EMA items and (B) the full EMA dataset were cumulatively added to predict ideation over 14 days

Participants with suicidal or self-harm ideation reported significantly lower scores across all EMA indicators. Notably, the self-harm ideation group demonstrated the lowest scores across all domains, suggesting more severe impairments in daily functioning and affective stability. These findings reinforce the clinical value of even brief, indirect EMA assessments in capturing real-time psychological vulnerability.

The combined model, which included EMA and baseline characteristics, significantly outperformed the baseline-only model (AUC=0.873). Model performance improved steadily as EMA data accumulated, suggesting that emotional and behavioral fluctuations over time are valuable for risk prediction. Clinically meaningful AUC levels were reached within 6–9 days of data collection.

Prediction of self-harm ideation was more variable, with lower AUCs overall. This may reflect the impulsive and context-dependent nature of self-harm, which is often more difficult to anticipate through affective patterns alone [22]. In addition, the relatively low number of self-harm ideation cases in our sample may have constrained model performance. Of the 998 total observations, only 75 were classified as indicating self-harm ideation, potentially limiting the ability of the model to learn stable and generalizable patterns. These findings highlight the need to explore additional data streams such as passive sensing or physiological monitoring to enhance prediction of self-harm risk. Prior EMA research has shown that affective instability and high intra-day variability—core features often associated with

self-harm—impair prediction accuracy, as they contribute to rapid and irregular shifts in risk states [9, 19, 23, 24].

Furthermore, prior work has emphasized the heterogeneity of suicidal ideation, including variations in intensity, duration, and motivational underpinnings [25]. By relying on indirect affective and behavioral signals, our model may have captured diverse trajectories of ideation that are not readily articulated in conventional assessments.

Unlike previous EMA studies conducted in high-risk clinical populations and often incorporating direct questions about suicidal ideation [23, 26, 27], this study employed indirect affective and behavioral indicators in a community sample. This reflects a non-invasive and scalable approach to suicide risk detection, particularly suited for individuals who do not disclose ideation in clinical settings or avoid professional help. Prior research has shown that repetitive questioning about suicidality may evoke reactivity in some individuals, and that suicidal ideation disclosed through EMA is often denied in retrospective reports [26, 28]. By avoiding direct questions and using EMA to monitor daily fluctuations in mood and behavior, this study may have captured early-stage or passive suicidal ideation—signals that are less likely to emerge in conventional assessments. Such an approach is especially relevant in contexts with high stigma around mental illness, such as Korea, where help-seeking is low despite high suicide prevalence [29].

While suicide risk is often assessed by directly asking individuals about their thoughts, this approach has limitations. Traditional self-report methods may fail to capture concealed or passive ideation due to stigma or fear of hospitalization [30]. In one study, 60% of participants who reported suicidal ideation during EMA later denied it in retrospective self-report, highlighting how EMA can uncover ideation that may otherwise go unreported [28]. By capturing daily mood and behavioral fluctuations in real-world settings, EMA enables the detection of subtle precursors to suicidal ideation that might not emerge in clinical interviews [13]. This indirect approach may be particularly effective for identifying subtle risk states not readily disclosed in clinical interviews. It also reduces the potential for reactivity sometimes observed with repeated direct questioning [26]. Together, these strengths position EMA-based monitoring as a promising strategy for proactive suicide prevention, particularly in contexts such as Korea, where suicide rates remain among the highest in the Organization for Economic Cooperation and Development (OECD) [31]. To enhance clinical applicability, future systems may integrate EMA data with clinician dashboards or automated alerts, enabling timely interventions guided by real-time affective trends. For example, coupling EMA platforms with

clinician-facing dashboards or automated alert systems could enable timely interventions based on real-time mood dynamics. This type of digital integration could help bridge the gap between passive risk detection and active clinical decision-making [14]. These systems could serve a dual role: first, enabling population-level screening in community settings—such as universities, public health centers, or digital wellness platforms—to identify individuals at elevated risk based on their EMA profiles; and second, triggering clinical referral pathways when sustained distress or high-risk trajectories are detected. In such cases, alerts could be generated to guide timely psychiatric evaluation or facilitate linkage to appropriate mental health services, ensuring that individuals receive clinical support even in the absence of spontaneous help-seeking. Although the current app does not yet incorporate active intervention features, future iterations could integrate alert-triggered responses that facilitate clinical linkage upon detection of high-risk patterns, thereby bridging passive monitoring with timely mental health care.

Importantly, these results should be viewed in light of the strong adherence observed in this study. A compliance rate of 94.0% over the EMA period reflects a high level of participant engagement, especially when compared to previous EMA studies reporting more moderate rates of adherence [26, 32]. This suggests that even low-burden EMA protocols, when well designed and clearly communicated, can be feasible and acceptable in non-clinical populations.

This study has several limitations, which also suggest directions for future improvement. First, both the predictors and outcome were based on self-report measures. While EMA provides ecologically valid real-time data, the K-CESD-R used for outcome measurement relies on a two-week retrospective window, which may be subject to recall bias. This limits the interpretability of predictive accuracy, as the target may not reflect momentary experiences with precision. Nonetheless, using self-report tools allowed us to efficiently capture a wide range of participants in a non-clinical setting, offering scalability that clinician-based assessment may not afford. Future studies should compare EMA-based predictions against clinician-rated outcomes to evaluate clinical relevance.

Second, our model predicted ideation occurring two weeks later, which does not directly translate to short-term or crisis-level intervention. The ability to detect next-day or same-day risk would be highly valuable. Prior research suggests that affective markers related to self-harm can change dramatically within hours, indicating the potential of near-term prediction to enhance safety planning or just-in-time interventions [19]. Future studies should explore ultra-short-term prediction frameworks using high-frequency EMA, or hybrid models that

combine EMA with passive data such as GPS, phone use, or wearable devices.

Third, although the model predicted future ideation status, it did not explicitly distinguish between different types of clinical state transitions, such as onset, persistence, or remission of suicidal or self-harm ideation. Clinically, the emergence of new ideation represents a particularly critical event. Future research should therefore adopt transition-based or multi-state modeling frameworks that explicitly differentiate these trajectories and evaluate transition-specific predictive performance.

Fourth, the use of only indirect EMA items limited our ability to detect explicit suicidal planning or intent. While this design reduces potential response fatigue and reactivity, it may miss certain risk signals. Periodic inclusion of direct items, or adaptive questioning algorithms, may help balance intrusiveness with specificity.

Finally, model interpretability remains an issue for clinical translation. While LSTM networks are powerful, their black-box nature may hinder clinician trust. Integration of explainability methods such as SHapley Additive exPlanations (SHAP) or Local Interpretable Model-agnostic Explanations (LIME), as demonstrated in recent EMA studies of suicidality, will be critical to ensure practical uptake in mental health settings.

Beyond these methodological limitations, adherence observed under research conditions may have been partially influenced by participant compensation and may be lower in real-world implementations where such incentives are not provided. This highlights the importance of user-centered design and engagement strategies when translating EMA-based monitoring systems into routine clinical or community settings.

In conclusion, this study demonstrates that brief, indirect, and self-reported EMA items can be used to predict suicidal and self-harm ideation in a non-clinical population. The integration of EMA with baseline characteristics significantly enhanced prediction accuracy, especially for suicidal ideation. Notably, this was achieved with just seven daily items, and this was further supported by the high compliance rate observed (94.0%), indicating strong feasibility in real-world settings. These findings support the utility of digital monitoring tools as scalable and context-sensitive strategies for early suicide risk detection. Future work should extend these models toward transition-focused and clinically actionable prediction frameworks, including behavioral outcomes such as suicide attempts.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12888-026-07815-6>.

Supplementary Material 1

Acknowledgements

The authors would like to thank all participants who generously contributed their time to this study.

Author contributions

Heeyeon Kim: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Writing – original draft. Seok-jae Heo: Data curation, Formal analysis, Visualization. Sehwon Park: Investigation, Software. Jooho Lee: Data curation, Formal analysis. Gangho Do: Software, Resources, Project administration. Jin Young Park: Conceptualization, Methodology, Supervision, Funding acquisition, Writing – review & editing. All authors read and approved the final manuscript.

Funding

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: RS-2023-KH135442).

Data availability

The datasets generated and/or analyzed during the current study are not publicly available due to privacy restrictions but are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

This study was approved by the Public Institutional Bioethics Committee designated by the Ministry of Health and Welfare of the Korean Government (MOHW; approval number P01-202411-01-038). All procedures contributing to this work complied with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. All participants provided informed consent prior to participation.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 29 September 2025 / Accepted: 12 January 2026

Published online: 10 February 2026

References

- Liu RT, Bettis AH, Burke TA. Characterizing the phenomenology of passive suicidal ideation: a systematic review and meta-analysis of its prevalence, psychiatric comorbidity, correlates, and comparisons with active suicidal ideation. *Psychol Med*. 2020;50(3):367–83.
- Wiklander M, Samuelsson M, Åsberg M. Shame reactions after suicide attempt. *Scand J Caring Sci*. 2003;17(3):293–300.
- Dunkley C, Borthwick A, Bartlett R, Dunkley L, Palmer S, Gleeson S, et al. Hearing the suicidal patient's emotional pain. *Crisis*. 2017.
- Wilson TD. Know thyself. *Perspect Psychol Sci*. 2009;4(4):384–9.
- Franklin JC, Ribeiro JD, Fox KR, Bentley KH, Kleiman EM, Huang X, et al. Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychol Bull*. 2017;143(2):187.
- Chapman AL, Gratz KL, Brown MZ. Solving the puzzle of deliberate self-harm: the experiential avoidance model. *Behav Res Ther*. 2006;44(3):371–94.
- Klonsky ED. The functions of deliberate self-injury: A review of the evidence. *Clin Psychol Rev*. 2007;27(2):226–39.
- Woodford R, Spittal MJ, Milner A, McGill K, Kapur N, Pirks J, et al. Accuracy of clinician predictions of future self-harm: a systematic review and meta-analysis of predictive studies. *Suicide Life-Threatening Behav*. 2019;49(1):23–40.
- Sedano-Capdevila A, Porras-Segovia A, Bello HJ, Baca-Garcia E, Barrigon ML. Use of ecological momentary assessment to study suicidal thoughts and behavior: a systematic review. *Curr Psychiatry Rep*. 2021;23(7):41.
- Quellec G, Berrouguet S, Morgiève M, Dubois J, Leboyer M, Vaiva G, et al. Predicting suicidal ideation from irregular and incomplete time series of

- questionnaires in a smartphone-based suicide prevention platform: a pilot study. *Sci Rep.* 2024;14(1):20870.
11. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annu Rev Clin Psychol.* 2008;4(1):1–32.
 12. Torous J, Staples P, Shanahan M, Lin C, Peck P, Keshavan M, et al. Utilizing a personal smartphone custom app to assess the patient health questionnaire-9 (PHQ-9) depressive symptoms in patients with major depressive disorder. *JMIR Mental Health.* 2015;2(1):e3889.
 13. Arney MF, Schatten HT, Haradhvala N, Miller IW. Ecological momentary assessment (EMA) of depression-related phenomena. *Curr Opin Psychol.* 2015;4:21–5.
 14. Kamath J, Barriera RL, Jain N, Keisari E, Wang B. Digital phenotyping in depression diagnostics: integrating psychiatric and engineering perspectives. *World J Psychiatry.* 2022;12(3):393.
 15. Ben-Zeev D, Young MA, Depp CA. Real-time predictors of suicidal ideation: mobile assessment of hospitalized depressed patients. *Psychiatry Res.* 2012;197(1–2):55–9.
 16. Nock MK, Prinstein MJ, Sterba SK. Revealing the form and function of self-injurious thoughts and behaviors: A real-time ecological assessment study among adolescents and young adults. 2010.
 17. Links PS, Eynan R, Heisel MJ, Barr A, Korzekwa M, McMain S, et al. Affective instability and suicidal ideation and behavior in patients with borderline personality disorder. *J Personal Disord.* 2007;21(1):72–86.
 18. Husky M, Olié E, Guillaume S, Genty C, Swendsen J, Courtet P. Feasibility and validity of ecological momentary assessment in the investigation of suicide risk. *Psychiatry Res.* 2014;220(1–2):564–70.
 19. Victor SE, Scott LN, Stepp SD, Goldstein TR. I want you to want me: interpersonal stress and affective experiences as within-person predictors of nonsuicidal self-injury and suicide urges in daily life. *Suicide Life-Threatening Behav.* 2019;49(4):1157–77.
 20. Elsworth S, Güttel S. Time series forecasting using LSTM networks: A symbolic approach. *ArXiv Preprint arXiv:200305672.* 2020.
 21. Vlachas PR, Byeon W, Wan ZY, Sapsis TP, Koumoutsakos P. Data-driven forecasting of high-dimensional chaotic systems with long short-term memory networks. *Proc R Soc A.* 2018;474(2213):20170844.
 22. Rodham K, Hawton K, Evans E. Reasons for deliberate self-harm: comparison of self-poisoners and self-cutters in a community sample of adolescents. *J Am Acad Child Adolesc Psychiatry.* 2004;43(1):80–7.
 23. Hallensleben N, Spangenberg L, Forkmann T, Rath D, Hegerl U, Kersting A, et al. Investigating the dynamics of suicidal ideation. *Crisis.* 2017.
 24. Rizk MM, Choo T-H, Galfalvy H, Biggs E, Brodsky BS, Oquendo MA, et al. Variability in suicidal ideation is associated with affective instability in suicide attempters with borderline personality disorder. *Psychiatry.* 2019;82(2):173–8.
 25. Klonsky ED, Saffer BY, Bryan CJ. Ideation-to-action theories of suicide: a conceptual and empirical update. *Curr Opin Psychol.* 2018;22:38–43.
 26. Kivelä L, Fiß F, Van der Does W, Antypa N. Examination of acceptability, feasibility, and iatrogenic effects of ecological momentary assessment (EMA) of suicidal ideation. *Assessment.* 2024;31(6):1292–308.
 27. Andrewes HE, Hulbert C, Cotton SM, Betts J, Chanen AM. An ecological momentary assessment investigation of complex and conflicting emotions in youth with borderline personality disorder. *Psychiatry Res.* 2017;252:102–10.
 28. Gratch I, Choo TH, Galfalvy H, Keilp JG, Itzhaky L, Mann JJ, et al. Detecting suicidal thoughts: the power of ecological momentary assessment. *Depress Anxiety.* 2021;38(1):8–16.
 29. An S, Lee H, Lee J, Kang S. Social stigma of suicide in South Korea: A cultural perspective. *Death Stud.* 2023;47(3):259–67.
 30. Klonsky ED, May AM, Saffer BY. Suicide, suicide attempts, and suicidal ideation. *Ann Rev Clin Psychol.* 2016;12(1):307–30.
 31. OECD. Society at a glance 2024: OECD social indicators. Paris: OECD Publishing; 2024.
 32. Porras-Segovia A, Díaz-Oliván I, Barrigón ML, Moreno M, Artes-Rodríguez A, Perez-Rodríguez MM, et al. Real-world feasibility and acceptability of real-time suicide risk monitoring via smartphones: A 6-month follow-up cohort. *J Psychiatr Res.* 2022;149:145–54.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.