Predicting short-term suicidal thoughts in adolescents using machine learning: Developing decision tools to identify daily-level risk after hospitalization

E.K. Czyz, Ph.D.[a], **H.J.M.S. Koo**[a], **N. Al-Dajani, Ph.D.**[a], **C. A. King, Ph.D.**[a,b], **I. Nahum-Shani, Ph.D.**[c]

[a.] Department of Psychiatry, University of Michigan

[b.] Department of Psychology, University of Michigan

[c.] Institute for Social Research, University of Michigan

Abstract

Background: Mobile technology offers unique opportunities for monitoring short-term suicide risk in daily life. In this study of suicidal adolescent inpatients, theoretically-informed risk factors were assessed daily following discharge to predict near-term suicidal ideation and inform decision algorithms for identifying elevations in daily-level risk, with implications for real-time suicide-focused interventions.

Methods: Adolescents (N=78; 67.9% female) completed brief surveys texted daily for four weeks after discharge (n=1621 observations). Using multi-level Classification and Regression Trees (CARTSs) with repeated 5-fold cross-validation, we tested (a) a simple prediction model incorporating previous-day scores for each of 10 risk factors, and (b) a more complex model incorporating, for each of these factors, a time-varying person-specific mean over prior days together with deviation from that mean. Models also incorporated missingness and contextual (study week, day of the week) indicators. The outcome was presence/absence of next-day suicidal ideation.

Results: The best-performing model (cross-validated AUC=0.86) was a complex model that included ideation duration, hopelessness, burdensomeness, and self-efficacy to refrain from suicidal action. An equivalent model that excluded ideation duration had acceptable overall performance (cross-validated AUC=0.78). Models incorporating only previous-day scores, with and without ideation duration (cross-validated AUC of 0.82 and 0.75, respectively), showed relatively weaker performance.

Conclusions: Results suggest that specific combinations of dynamic risk factors assessed in adolescents' daily life have promising utility in predicting next-day suicidal thoughts. Findings represent an important step in the development of decision tools identifying short-term risk as well as guiding timely interventions sensitive to proximal elevations in suicide risk in daily life.

Keywords

suicidal ideation; adolescents; daily diary; ecological momentary assessment; machine learning; risk algorithm; just-in-time adaptive intervention

The prevention of suicide deaths and related outcomes in youth, including non-lethal attempts and suicidal thoughts, is an urgent public health priority. The prevalence of suicide deaths among adolescents in the United States has been on the rise, showing a nearly 60% increase between 2007 and 2018 (Curtin, 2020). Periods of high-risk transition, such as following psychiatric hospitalization, are associated with particularly elevated suicide risk (Chung et al., 2019). Approximately 20% of discharged youth experience a suicide-related event (e.g., rehospitalization, suicide attempt) within three months and between 30%-40% experience these events within six months after discharge (Czyz et al., 2016; Kennard et al., 2018; Yen et al., 2013; Yen et al., 2019). Discharged adolescents are also vulnerable to suicidal thoughts of varying intensity and chronicity (Prinstein et al., 2008; Wolff et al., 2018), with more persistent suicidal ideation patterns showing an especially strong association with suicide attempts (Czyz & King, 2015; Prinstein et al., 2008), and with prior research highlighting notable fluctuations in day-to-day suicidal thoughts shortly after hospitalization (Czyz et al., 2018). Given that suicidal ideation and related outcomes are heterogeneous and time-varying, there is a critical need for strategies that can accurately detect elevations in suicidal ideation and risk more broadly as well as guide the timely delivery of interventions, particularly during high-risk periods.

Identifying Elevations in Suicide Risk in Daily Life

The pervasiveness of mobile phones has paved the way for use of intensive longitudinal assessment approaches—such as daily diaries (once-a-day assessments) and ecological momentary assessments (EMAs) characterized by repeated measurements within a day— that allow for sampling individuals' day-to-day experiences (Shiffman et al., 2008). By allowing for repeated and frequent assessment of experiences in real-world environments and over time, these approaches reduce recall bias and lead to more ecologically valid information about thoughts, behaviors, or feelings compared to traditional assessments spaced over longer time intervals (Shiffman et al., 2008). Daily diary and EMA studies are increasingly common in suicide prevention research (Davidson et al., 2017; Gee et al., 2020; Kleiman & Nock, 2018), as they enable assessment of dynamic and short-term precursors of suicide risk. In addition to demonstrating the highly dynamic nature of suicidal thoughts (Hallensleben et al., 2017; Kleiman et al., 2017), this growing research has also examined short-term correlates and precipitants of suicidal ideation across different domains (e.g., situational, affective, interpersonal) among adults (Armey et al., 2018; Ben-Zeev et al., 2012; Coppersmith et al., 2019; Hallensleben et al., 2019; Husky et al., 2017; Kleiman et al., 2017) and, to a lesser extent, among adolescents (Czyz et al., 2018; Nock et al., 2009). Of note, existing daily diary and EMA studies have largely examined individual indicators of suicidal ideation or relatively limited interactions guided by theory (Czyz et al., 2018; Hallensleben et al., 2019). A focus on individual predictors alone, or solely pre-specified interactions, may obscure important patterns in data that could be useful in

detecting short-term risk. For example, a proof-of-concept study has shown that models that incorporate multiple risk and protective factors assessed daily after hospitalization predict near-term suicidal crises more accurately relative to models that incorporate single factors (Czyz et al., 2020).

In line with meta-analytic studies indicating that individual risk factors are only modestly predictive of suicidal thoughts and behavior (Franklin et al., 2016; Ribeiro et al., 2016), more research considering multiple risk factors (e.g., King et al., 2020) and their complex interactions may be needed to improve our understanding of real-time, real-world conditions that represent near-term suicide risk. Although not yet widely utilized in this context, data-driven approaches (such as machine learning) could aid in identifying such complex relationships, with implications for improving short-term risk detection and for guiding the delivery of timely support.

Potential Value of Decision Algorithms Identifying Short-Term Risk

Machine learning approaches (e.g., support vector machine, random forests, and artificial neural network) have been increasingly applied in suicide prevention research (see review Burke et al., 2019). These methods may be highly useful in identifying states of vulnerability representing elevations in suicide risk based on complex intensive longitudinal data by estimating or "learning" parameters that can produce reliable prediction of an adverse proximal outcome (e.g., near-term suicidal thoughts) (Nahum-Shani et al., 2021). These methods have the important advantage of handling large volumes of predictors and combining them in interactive and non-linear ways (Athey, 2017), yielding algorithmically optimized prediction models. Machine learning has shown promise in predicting suicidal ideation and behavior among adults (Chen et al., 2020; de la Garza et al., 2021; Ribeiro et al., 2019; Walsh et al., 2017) and adolescents (Hill et al., 2017; Miché et al., 2020; Walsh et al., 2018). Moreover, a study of adult inpatients responding to EMAs during hospitalization used machine learning models to predict suicide attempts two to four weeks later (Wang et al., 2021). To date, however, machine learning approaches have not been applied to predict near-term suicidal thoughts in daily life.

Applying such data-driven approaches, based on intensive longitudinal data, to detect proximal elevations in suicide risk may be especially valuable if it can aid in identifying opportune times to provide timely interventions that are sensitive to individuals' changing suicide risk levels in real-world conditions. For example, just-in-time adaptive interventions (JITAIs) use dynamically changing information about the individual's internal state and context to recommend whether, when, and how to deliver interventions in daily life using mobile technology (Nahum-Shani et al., 2015; Nahum-Shani et al., 2017). JITAIs employ decision rules that link dynamic information about the individual to specific intervention options, specifying the conditions in which an intervention should be delivered. The goal is to provide the intervention option that is best for an individual at a given time point, while avoiding unnecessary intervention and burden (Nahum-Shani et al., 2015). Identifying states of vulnerability to an adverse proximal outcome (here, proximal suicidal thoughts) play a critical role in the formulation of effective JITAIs. Many JITAIs and similar ecological interventions focusing on mental health (see reviews: Balaskas et al., 2021; Bidargaddi et al.,

2020; Wang & Miller, 2020) are motivated to break the link between states of vulnerability (i.e. conditions that represent heightened risk) and a specific adverse proximal outcome via the delivery of a timely intervention. However, to inform the development of real-time interventions for suicide prevention, more research is needed to empirically identify states of vulnerability or the conditions in which adolescents at risk for suicide may benefit from such interventions (i.e. *when* an intervention should be delivered). Hence, the current study seeks to leverage intensive longitudinal data and machine learning to arrive at a decision algorithm for detecting near-term suicidal thoughts, which, in turn, could be used to guide the formulation of real-time interventions (Nahum-Shani et al., 2015).

The Current Study

In this daily diary study of psychiatrically hospitalized adolescents, multi-level classification and regression tree (CART) was applied to predict proximally emerging suicidal ideation, operationalized as next-day suicidal ideation, during a high-risk post-discharge period. The multi-level CART was selected, instead of other machine learning strategies (e.g., random forest, elastic net, neural network), due its ability to generate readily interpretable results (e.g., see Boudreaux et al., 2021 for an overview) while simultaneously considering intensively sampled predictors (i.e. multi-level data structure) and their complex combinations (Fokkema et al., 2020). Risk and protective factors were selected based on theoretical considerations and clinical relevance (Bandura, 1977; Klonsky & May, 2015; Rudd et al., 2006; Shneidman, 1993; Van Orden et al., 2010; Wenzel & Beck, 2008) and were assessed via electronic surveys each day over the course of one month after psychiatric hospitalization. The multi-level CART was used to arrive at a decision tool (i.e., algorithm) predicting the occurrence of next-day suicidal ideation. Two approaches were considered: (1) a simple prediction model incorporating previous-day scores for each individual factor, and (2) a more complex model incorporating, for each of these factors, a person-specific mean over the days prior together with the deviation (change score) from that person-specific mean. To the best of our knowledge, this is the first study that leveraged machine learning to generate decision algorithms for predicting short-term suicidal ideation in daily life. Such decision algorithms represent an important step toward the development of suicide-focused interventions addressing real-time, real-world changes in suicide risk.

Methods

Participants and Procedures

Participants were psychiatrically hospitalized adolescents (ages 13–17) who were eligible to participate based on last-month suicide attempt and/or last-week suicidal ideation (with thoughts of either method, intent, and/or plan; based on the Columbia-Suicide-Severity Rating Scale [Posner et al., 2011]). Exclusion criteria were severe cognitive impairment or altered mental status (psychosis, mania), transfer to medical unit or residential placement, no availability of a legal guardian (ward of sate), or adolescents not having a cell phone. Described in detail elsewhere (Czyz et al., 2021), participants, were recruited between March 2019 and January 2020 as part of a psychosocial intervention pilot study focusing on feasibility and acceptability[1]. Of those who provided study consent/assent (n=82 or 87.2%),

80 completed baseline assessment and continued in the study. Adolescents completed daily surveys for four weeks, which were texted to their phone each evening beginning on the first day after discharge. Adolescents could respond to surveys between 5–8pm and received $4 for completing each survey. This study's analytic sample includes 78 adolescents (97.5%) who completed at least two consecutive daily surveys; survey completion rate was 74.2% (1621 out of 2,184 possible daily surveys) over the four-week period. On average, adolescents completed 20.78 daily surveys (SD = 6.92). The study's procedures complied with ethical standards for human subjects research and were approved by the participating university's Institutional Review Board.

Measures

Daily Outcome—Each day, adolescents responded to questions assessing frequency of their suicidal thoughts in reference to the last 24 hours ("How many times did you have thoughts of killing yourself?"), with responses ranging from 0 (not at all) to 4 (all the time). Modeled after the Columbia-Suicide Severity Rating Scale (C-SSRS) (Posner et al., 2011), the item was previously adapted to assess daily suicidal ideation in adolescents (Czyz et al., 2018). A binary outcome was created to indicate absence/presence (0 versus 1) of next-day suicidal ideation.

Daily Predictors—Predictors were assessed with the following brief measures, adapted based on existing validated scales, to reduce response burden given that data were collected daily. Adolescents responded to all items in reference to the last 24 hours.

Hopelessness.: Participants rated the extent to which they felt hopeless ("I see only bad things ahead of me, not good things") using a 4-point scale from "strongly disagree" to "strongly agree." This item was modeled after the 6-item Brief Hopelessness Scale (Bolland et al., 2001).

Connectedness to Family and Friends.: Using a 7-point scale (from "not at all true for me" to "very true for me"), adolescents rated the extent to which they felt close to their friends and, separately, to their family. These items were modeled after the Interpersonal Needs Questionnaire (INQ) (Van Orden et al., 2012), which measures thwarted belongingness and perceived burdensomeness. The INQ closeness item was adapted to reference sense of closeness with peers (item 1) and with family (item 2).

Burdensomeness.: Participants rated their perceived sense of burdensomeness ("The people in my life would be happier without me") on a 7-point scale ranging from "not at all true for me" to "very true for me." This item was similarly based on the INQ (Van Orden et al., 2012)

Agitation.: Participants rated their agitation ("I felt so stirred up inside I wanted to scream") on a 7-point scale. The item was modeled after an item from the Brief Agitation Measure

---

[1]The results did not differ when the pilot intervention conditions were controlled for in the models. Thus, intervention conditions were removed from the final models.

(Ribeiro et al., 2011). The item phrasing was adjusted to reference experiences in the 24-hour period and response options ranged from "not at all" to "very much."

Worry and Rumination.: Using a scale ranging from 1 ("not at all") to 7 ("very much"), participants reported levels of worry and rumination using two items. Rumination was assessed with the item "I was dwelling on my feelings and problems," and worry was assessed with the item "I was worried about things that could happen." These items were based on previous EMA studies of rumination and worry (Kircanski et al., 2015).

Self-efficacy to Refrain from Suicidal Action.: Confidence to refrain from suicidal action ("How confident are you that you will be able to keep yourself from attempting suicide?") was assessed with an item from the 3-item Self-Assessed Expectations of Suicide Risk Scale (Czyz et al., 2016), which showed predictive validity for suicide attempts in youth. Responses were rated from 0 ("not at all confident") to 10 ("completely confident").

Psychological Pain.: Participants rated the extent to which they felt miserable as a proxy for psychological pain. Responses were rated on a 5-point scale (from "very slightly or not at all" to "extremely"). This question was adapted from the 10-item Positive and Negative Affect Schedule for Children (PANAS-C) (Ebesutani et al., 2012).

Suicidal Ideation Duration.: Each day, adolescents reporting any suicidal ideation subsequently rated the duration of their suicidal thoughts ("*How long did these thoughts last?*") on a 5-point scale (from "a few seconds or minutes" to "more than 8 hours/ continuous"). The duration of ideation item was based on the C-SSRS (Posner et al., 2011). We created a continuous scale for ideation duration ranging from 0 (no ideation) to 5 (continuous ideation).

Predictor Preparation

For each of the 10 risk and protective factors of interest, we calculated the cumulative person-specific mean (each person's own mean) for each day and the deviation from that mean. Specifically, for each of the 10 factors, a person-specific mean was calculated for each day t based on daily responses up to and including day t (i.e. within-person sum of responses to the specific factor up to and including day t, divided by the number of completed surveys up to and including t). Additionally, for each factor, the deviation of day t response from the cumulative person-specific mean was calculated. The relatively simple prediction model included previous-day t responses for each of the 10 factors, whereas the more complex model included each factor's cumulative person-specific mean for each day t as well as day t deviation from that mean.

Data Analytic Strategy

We fit a series of multi-level CART models to predict next-day suicidal ideation. Designed to accommodate multi-level and longitudinal data structures, these CART models employ generalized linear mixed model (GLMM) and decision trees; specifically, a GLMM tree algorithm uses an unbiased recursive partitioning method that selects the splitting variable based on the lowest p value until there are no more splitting predictors that are below our

pre-specified α of 0.05 (Fokkema et al., 2020). A tree like-structure is produced using a process of partitioning data into subgroups such that observations in each group become more similar in terms of the outcome within groups, a process that is repeated until no additional improvements in classification of the outcome are identified based on available predictors. The CART has the advantage of not requiring any assumptions of the data and allowing for a large number of predictors to be included in the model. Moreover, CARTs provide easily interpretable results.

We initially fit two separate CART models: (a) a simple model including previous-day response for each of the 10 factors; and (2) a more complex model including, for each factor, the previous-day cumulative person-specific mean and deviation from that mean. The models also included indicators for week in the study (ranging from week 1–4), day of the week (from Mon-Sun), and a variable indicating whether previous-day survey was missing (previous-day missingness). To investigate the added value of measuring ideation duration as a predictive factor, we fit two additional models that were identical to the first two but excluded predictors based on ideation duration. The outcome across all models was presence/absence of next-day suicidal ideation (binary outcome).

To evaluate model performance, we used blocked 5-fold cross-validation. This method includes random division of the data into five sets, such that four models are created and then tested on a single selected set. This process is repeated five times, with a different test set each time. Traditional random k-fold cross-validation methods do not consider dependent data structures (observations nested within individuals) and may thus provide too optimistic results, leading to incorrect confidence in predictions (Roberts et al., 2017). Also, unbalanced values of the outcome (i.e., higher proportion of zeros to ones) can be problematic for cross-validation (Thabtah et al., 2020). Thus, cross-validation was based on stratified blocked folds (i.e. individual participants serving as blocks) with similar prevalence of the outcome (Roberts et al., 2017). Folds were created such that all observations from a particular participant were within the same fold, and the proportion of high and low risk individuals (based on the median split of the proportion of days ideation was endorsed) were kept consistent across the folds. The blocked 5-fold cross validation process was repeated 10 items to improve stability and consistency of results. All the analyses were conducted in R version 4.0.2 (2020–06–22) utilizing glmertree, caret, pROC, and groupdata2 packages.

Model performance was evaluated using different metrics, including sensitivity, specificity, area under the receiver operating characteristic curve (AUC), and positive predictive value (PPV). Sensitivity refers to the probability of identifying correctly those who would experience the outcome (i.e. next-day suicidal ideation) while specificity refers to the probability of identifying correctly those who would not experience the outcome. AUC captures the average sensitivity over all values of false positive rates (i.e., 1-Specificity) of different cutoff points of predicted probability; AUC thus provides a measure of discriminatory ability (accuracy of differentiating cases with the outcome from those without), with values ranging from 0 to 1, where 0.5 indicates no discriminative ability (similar to chance) and 1 indicates perfect discrimination. Of note, to obtain maximal sensitivity and specificity for selected models, we sought to determine cutoffs on predicted

probabilities using the "closest top left" criterion implemented in the R package pROC (Robin et al., 2011). Finally, the PPV refers to the probability that those identified as having the outcome actually have the outcome (i.e. proportion of true positives over the number of cases predicted to be positive based on the model). Unlike sensitivity and specificity, PPV is directly related to prevalence rate of the outcome in the population; hence, PPV will increase with increasing prevalence of the outcome in the population. Similar to AUC, sensitivity, specificity, and PPV values can range from 0 to 1, with values closer to 1 indicating better performance.

Results

Participant Characteristics

Over 60% of participants were biological females (67.9%; n=53), with a mean age of 15.19 (SD=1.35) years. The distribution of race/ethnicity was as follows (more than one category could be selected): 65 (83.3%) White, 5 (6.4%) African American/Black, 4 (5.1%) Asian, 4 (5.1%) American Indian or Alaska Native, and 1 (1.3%) Native Hawaiian or Other Pacific Islander. Nine participants (11.5%) self-identified as Hispanic. At the time of hospitalization, all adolescents had last-week suicidal ideation, with a mean ideation rating of 3.90 (SD=0.91) (range 0–5, where 0 is no ideation and 5 corresponds to ideation with intent and plan). In addition, half of participants (n=39) had at least one lifetime suicide attempt. Over the course of four weeks after discharge, 64 (82.1%) participants reported a total of 631 instances of suicidal ideation (38.9%), out of 1621 completed daily surveys, with an average of 9.86 (SD=7.91) instances. Descriptive information for the predictors and the outcome are provided in supplemental Table 1.

Model Performance

As shown in Table 1, the best-performing model includes the cumulative mean and deviations from that mean as predictors. This model has a cross-validated mean AUC value of 0.86 (standard error [SE]=0.002), mean PPV of 0.74 (SE=0.006), mean sensitivity of 0.81 (SE=0.005), and mean specificity of 0.82 (SE=0.006). Its corresponding tree is presented in Figure 1. The simple model (previous-day scores) yielded a cross-validated mean AUC value of 0.82 (SE=0.002), mean PPV of 0.69 (SE=0.005), mean sensitivity of 0.79 (SE=0.008), and mean specificity of 0.78 (SE=0.006). Thus, the complex model resulted in 4% improvement in accurately identifying suicidal ideation events, which is important to maximize in the context of suicide prevention[2]. The best-performing model (see Figure 1) included a combination of cumulative mean and/or deviation from that mean for the following factors: ideation duration, hopelessness, burdensomeness, and self-efficacy. For instance, as shown in Figure 1, one pathway leading to next-day suicidal ideation involves the following: 0.95 or lower cumulative mean of ideation duration, combined with 2.19 or lower cumulative mean of hopelessness, together with greater than 1.95 cumulative mean of burdensomeness and with −1.59 or lower change score of self-efficacy. Table

---

[2]In supplemental analyses, we considered whether the best-performing model could be further improved by adding previous-day scores (i.e. including all 3 features: cumulative means, deviations, and previous-day scores). Notably, this model (cross validated AUC=0.85; PPV=0.74; Sensitivity=0.76; specificity=0.81) did not yield improved prediction, suggesting previous-day scores did not contribute additional information to the best-performing model.

3 provides detailed information about interpreting this model's results, including specific thresholds in the decision tool shown in Figure 1.

As shown in Table 2, models without duration of suicidal thoughts had acceptable—but relatively poorer—performance. The simple model without suicidal ideation duration was least predictive, with cross-validated mean AUC of 0.75 (SE=0.005), mean PPV of 0.61 (SE=0.006), mean sensitivity of 0.73 (SE=0.009), and mean specificity of 0.71 (SE=0.009). As shown in Figure 2, the best-performing model that excluded suicidal ideation as a predictor (cross-validated mean AUC=0.78) identified the following predictors for next-day ideation: burdensomeness, hopelessness, self-efficacy, and study week. Supplemental Table 2 provides information about interpreting this model's results, including specific thresholds in the decision tool, shown in Figure 2.

Discussion

In this study of suicidal adolescent inpatients, we applied machine learning methods (multi-level CARTs) to daily data collected over four weeks during the post-discharge period. The goal was to develop a decision tool that leverages these intensive longitudinal data to indicate each day whether the adolescent is likely to experience suicidal ideation on the next day. To our knowledge, this is the first study that used machine learning to identify different combinations of intensely sampled factors (and their specific thresholds) that predicted, with relatively high accuracy, the emergence of next-day suicidal ideation during the critical post-discharge period. The results further highlighted that such algorithmically optimized combinations hold promise for identifying specific states of vulnerability to near-term suicidal ideation that could, in turn, be used to guide *when* (i.e., the conditions in which) an intervention might be needed. The study's key findings and implications are discussed below.

First, we found that the best-performing CART model (AUC of 0.86)—which included time-varying cumulative means of risk factors together with deviations from the means—predicted next-day suicidal ideation with greater accuracy relative to a simpler model that included previous-day ratings of risk factors. The fact that prediction was improved by incorporating different features of risk factors (i.e. mean and deviation) is largely consistent with previous daily diary and EMA studies, where more complex models exhibited better prediction of suicide-related outcomes two to four weeks later (Czyz et al., 2020; Wang et al., 2021). Others have similarly shown that, in contrast to simpler (univariate) models, machine learning models accounting for complex combinations among predictors showed stronger prediction of suicidal thoughts and attempts 3, 14, and 28 days later (Ribeiro et al., 2019). However, in contrast to previous machine learning studies of suicide-related outcomes assessed over longer intervals (Chen et al., 2020; de la Garza et al., 2021; Ribeiro et al., 2019; Walsh et al., 2017; Walsh et al., 2018), CARTs in this study achieved relatively good predictive accuracy without using a large number of predictors (e.g., 50–3000 predictors in prior studies). Attending to the issues of parsimony is an important consideration for ultimately translating algorithms into clinical practice, particularly if such algorithms incorporate frequent assessments. In the current study, it was also notable that person-specific cumulative means of factors, in combination with more proximally

occurring experiences (represented by deviations from the cumulative mean), emerged as being important in identifying next-day ideation. While somewhat intuitive (e.g., enduring hopelessness over prior days together with more recent increases in hopelessness make for a powerful combination), we are not aware of previous studies that examined how dynamically changing factors accumulate to influence near-term suicidal ideation or risk more broadly. It may be that examining short-term changes in risk and protective factors, together with their accumulation, provides a fuller picture of the processes impacting suicidal ideation in daily life. For example, previous studies have demonstrated that within-person changes in individual risk factors alone have limited prospective prediction of short-term suicidal thoughts (Ben-Zeev et al., 2012; Coppersmith et al., 2019; Kleiman et al., 2017).

Second, there were multiple pathways of risk factor combinations associated with next-day suicidal ideation, although duration of suicidal thinking emerged as the top risk factor. This is in line with previous daily and EMA studies demonstrating that suicidal ideation tends to have the most consistent prospective link with suicidal thoughts (Coppersmith et al., 2019; Hallensleben et al., 2019; Kleiman et al., 2017). However, we also found that when excluding predictors based on ideation duration, next-day suicidal thoughts can still be adequately predicted based on other factors (e.g., one such pathway incorporated, at specific thresholds of each: cumulative hopelessness mean, cumulative burdensomeness mean, and change in self-efficacy score). In other words, the CART analysis identifies other pathways that could "flag" next-day suicidal ideation. Beyond identifying *when* an intervention could potentially be provided, a promising application of identifying multiple pathways of near-term ideation risk could include the personalization of intervention content based on the unique constellation of risk factors (specific warning signs of near-term ideation). This is consistent with expert consensus highlighting a range of warning signs related to suicide risk (Fowler, 2012; Rudd et al., 2006) as well as empirical evidence indicating that individuals exibit different warning signs prior to attempting suicide (Bagge et al., 2013; Bagge et al., 2017). Here, we show that near-term suicidal ideation is similarly multidetermined and thus could benefit from interventions addressing different risk pathways in daily life.

Finally, findings highlight that intensive assessment of prior ideation duration may not be needed to identify subsequent suicidal ideation risk. The more complex model without ideation duration yielded reasonably good predictive accuracy (AUC of 0.78). Previous studies have similarly found that models without suicidal ideation performed well in identifying the occurrence of suicidal thoughts and related crises within a few days to weeks (Czyz et al., 2020; Ribeiro et al., 2019). Excluding ideation duration in prediction models may offer practical advantages if frequent assessment of suicidal thoughts is not feasible or if there are concerns that at-risk individuals may not disclose ideation if queried directly (Bernecker et al., 2019; Drum et al., 2009). However, as found in the current study, excluding ideation may lead to some tradeoffs in model performance (e.g., lower specificity). Additional practical considerations include sustaining response adherence (especially over longer periods of time) and minimizing response burden. Prediction models that are parsimonious and do not solely rely on self-report responses may be especially valuable. For example, real-time passive data not requiring direct input from individuals (e.g., sensors, geolocation, communication logs) may offer unique advantages for predicting

near-term risk (Allen et al., 2019; Kleiman et al., 2019; Torous, et al., 2018), although the utility of such data in predicting suicide-related outcomes, particularly in everyday life, still needs to be established to guide clinical decision-making.

Study Limitations and Future Directions

The study's findings should be considered in light of key limitations. Adolescents in this study were primarily White and were recruited from a single inpatient unit, which might limit generalizability. The findings should be replicated in other samples. While the risk factors included in analyses were selected based on clinical and theoretical considerations, they do not offer an exhaustive list of constructs that can be explored. Future research could consider additional risk and protective factors along with additional variable features. Given limitations inherent in self-report data (e.g., social desirability, recall bias, response burden), additional research should also consider examining predictors based on other types of assessment, such as passively collected data. Future extensions of this work, particularly involving large samples, could examine the extent to which near-term risk algorithms incorporating intensive longitudinal data may vary as a function of participants' initial characteristics (e.g., suicide attempt history) as well as investigate their predictive performance beyond the near-term window (ideation outcome defined as 2 days later, 3 days later, etc.). In addition to detecting its near-term presence, future extensions of this work could also consider predicting proximal change in suicidal ideation (e.g., continued absence, continued presence, decline, or emergence). Moreover, while we tested the stability of results using repeated 5-fold cross-validation, replicating these findings in an independent sample would offer the most rigorous validation approach. While beyond the scope of this study, future research is also needed to determine the practical utility of such prediction algorithms in informing intervention development, including how to intervene when individuals are vulnerable to next-day suicidal thoughts and whether individuals at risk for suicide are receptive to such real-time interventions.

Conclusions

With the goal of developing a decision tool that leverages intensive longitudinal data to identify near-term suicidal ideation in a high-risk clinical sample, this study advances prior research focusing on prediction of short-term suicide risk. Among recently discharged adolescents, we applied a series of multi-level CART models incorporating dynamically changing risk and protective factors to predict next-day suicidal thoughts. While simpler models that included previous-day ratings of risk and protective factors yielded adequate predictive accuracy, CART models achieved improved performance when including as predictors both the cumulative means of factors over prior days and deviations from these means. These results illustrate that intensive longitudinal data could guide the delivery of real-time interventions for suicide prevention by identifying specific states of vulnerability to near-term suicidal thoughts. Additional research is needed to validate results in independent samples as well as build on this work by improving generalizability to other populations and settings, incorporating different risk factors and data collection methods, as well as by attending to practical issues that could impact translation of prediction algorithms into clinical settings.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

Allen NB, Nelson BW, Brent D, & Auerbach RP (2019). Short-term prediction of suicidal thoughts and behaviors in adolescents: Can recent developments in technology and computational science provide a breakthrough? Journal of Affective Disorders, 250, 163–169. [PubMed: 30856493]

Armey MF, Brick L, Schatten HT, Nugent NR, & Miller IW (2018). Ecologically assessed affect and suicidal ideation following psychiatric inpatient hospitalization. General Hospital Psychiatry, 63, 89–96. [PubMed: 30297091]

Athey S (2017). Beyond prediction: Using big data for policy problems. Science, 355(6324), 483–485. [PubMed: 28154050]

Bagge CL, Glenn CR, & Lee H (2013). Quantifying the impact of recent negative life events on suicide attempts. Journal of Abnormal Psychology, 122(2), 359–368. [PubMed: 23088374]

Bagge CL, Littlefield AK, & Glenn CR (2017). Trajectories of affective response as warning signs for suicide attempts: An examination of the 48 hours prior to a recent suicide attempt. Clinical Psychological Science, 5(2), 259–271.

Balaskas A, Schueller SM, Cox AL, & Doherty G (2021). Ecological momentary interventions for mental health: A scoping review. PloS One, 16(3), e0248152. [PubMed: 33705457]

Bandura A (1977). Self-efficacy: Toward a unifying theory of behavioral change. Psychological Review, 84(2), 191. [PubMed: 847061]

Ben-Zeev D, Young MA, & Depp CA (2012). Real-time predictors of suicidal ideation: Mobile assessment of hospitalized depressed patients. Psychiatry Research, 197(1), 55–59. [PubMed: 22397912]

Bernecker SL, Zuromski KL, Gutierrez PM, Joiner TE, King AJ, Liu H, . . . Stein MB. (2019). Predicting suicide attempts among soldiers who deny suicidal ideation in the army study to assess risk and resilience in servicemembers (Army STARRS). Behaviour Research and Therapy, 120, 103350. [PubMed: 30598236]

Bidargaddi N, Schrader G, Klasnja P, Licinio J, & Murphy S (2020). Designing m-health interventions for precision mental health support. Translational Psychiatry, 10(1), 1–8. [PubMed: 32066695]

Bolland J, McCallum D, Lian B, Bailey C, & Rowan P (2001). Hopelessness and violence among inner-city youths. Maternal and Child Health Journal, 5(4), 237–244. [PubMed: 11822525]

Boudreaux ED, Rundensteiner E, Liu F, Wang B, Larkin C, Agu E, Ghosh S, Semeter J, Simon G, & Davis-Martin RE (2021). Applying Machine Learning Approaches to Suicide Prediction Using Healthcare Data: Overview and Future Directions. Frontiers in Psychiatry, 12, 707916. [PubMed: 34413800]

Burke TA, Ammerman BA, & Jacobucci R (2019). The use of machine learning in the study of suicidal and non-suicidal self-injurious thoughts and behaviors: A systematic review. Journal of Affective Disorders, 245, 869–884. [PubMed: 30699872]

Chen Q, Zhang-James Y, Barnett EJ, Lichtenstein P, Jokinen J, D'Onofrio BM, . . . Fazel S. (2020). Predicting suicide attempt or suicide death following a visit to psychiatric specialty care: A machine learning study using Swedish national registry data. PLoS Medicine, 17(11), e1003416. [PubMed: 33156863]

Chung D, Hadzi-Pavlovic D, Wang M, Swaraj S, Olfson M, & Large M (2019). Meta-analysis of suicide rates in the first week and the first month after psychiatric hospitalisation. BMJ Open, 9(3), e023883.

Coppersmith DD, Kleiman EM, Glenn CR, Millner AJ, & Nock MK (2019). The dynamics of social support among suicide attempters: A smartphone-based daily diary study. Behaviour Research and Therapy, 120, 103348. [PubMed: 30594300]

Curtin SC (2020). State suicide rates among adolescents and young adults aged 10–24: United States, 2000–2018. National Vital Statistics Report, 69(11). Hyattsville, MD: National Center for Health Statistics.

Czyz EK, Berona J, & King CA (2016). Rehospitalization of suicidal adolescents in relation to course of suicidal ideation and future suicide attempts. Psychiatric Services, 67(3), 332–338. [PubMed: 26725287]

Czyz EK, Horwitz AG, Arango A, & King CA (2018). Short-term change and prediction of suicidal ideation among adolescents: A daily diary study following psychiatric hospitalization. Journal of Child Psychology and Psychiatry, 60(7), 731–741.

Czyz EK, Horwitz AG, & King CA (2016). Self-rated expectations of suicidal behavior predict future suicide attempts among adolescent and young adult psychiatric emergency patients. Depression and Anxiety, 33, 512–519. [PubMed: 27110663]

Czyz EK, King CA, Prouty D, Micol VJ, Walton M, & Nahum-Shani I (2021). Adaptive intervention for prevention of adolescent suicidal behavior after hospitalization: A pilot sequential multiple assignment randomized trial. Journal of Child Psychology and Psychiatry. In press.

Czyz EK, Yap JR, King CA, & Nahum-Shani I (2020). Using intensive longitudinal data to identify early predictors of suicide-related outcomes in high-risk adolescents: Practical and conceptual considerations. Assessment, 1–11.

Davidson CL, Anestis MD, & Gutierrez PM (2017). Ecological momentary assessment is a neglected methodology in suicidology. Archives of Suicide Research, 21(1), 1–11. [PubMed: 26821811]

de la Garza Ángel García, Blanco C, Olfson M, & Wall MM. (2021). Identification of suicide attempt risk factors in a national US survey using machine learning. JAMA Psychiatry, 78(4), 398–406. [PubMed: 33404590]

Drum DJ, Brownson C, Burton Denmark A, & Smith SE (2009). New data on the nature of suicidal crises in college students: Shifting the paradigm. Professional Psychology: Research and Practice, 40(3), 213–222.

Ebesutani C, Regan J, Smith A, Reise S, Higa-McMillan C, & Chorpita BF (2012). The 10-item positive and negative affect schedule for children, child and parent shortened versions: Application of item response theory for more efficient assessment. Journal of Psychopathology and Behavioral Assessment, 34(2), 191–203.

Fokkema M, Edbrooke-Childs J, & Wolpert M (2020). Generalized linear mixed-model (GLMM) trees: A flexible decision-tree method for multilevel and longitudinal data. Psychotherapy Research, 31(3), 329–341.

Fowler JC (2012). Suicide risk assessment in clinical practice: Pragmatic guidelines for imperfect assessments. Psychotherapy, 49(1), 81–90. [PubMed: 22369082]

Franklin JC, Ribeiro JD, Fox KR, Bentley KH, Kleiman EM, Huang X, . . . Nock MK. (2016). Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. Psychological Bulletin, 143(2), 187–232. [PubMed: 27841450]

Gee BL, Han J, Benassi H, & Batterham PJ (2020). Suicidal thoughts, suicidal behaviours and self-harm in daily life: A systematic review of ecological momentary assessment studies. Digital Health, 6, 1–38.

Hallensleben N, Glaesmer H, Forkmann T, Rath D, Strauss M, Kersting A, & Spangenberg L (2019). Predicting suicidal ideation by interpersonal variables, hopelessness and depression in real-time. an ecological momentary assessment study in psychiatric inpatients with depression. European Psychiatry, 56(1), 43–50. [PubMed: 30530103]

Hallensleben N, Spangenberg L, Forkmann T, Rath D, Hegerl U, Kersting A, . . . Glaesmer H. (2017). Investigating the dynamics of suicidal ideation. Crisis, 39(1), 65–69. [PubMed: 28468557]

Hill RM, Oosterhoff B, & Kaplow JB (2017). Prospective identification of adolescent suicide ideation using classification tree analysis: Models for community-based screening. Journal of Consulting and Clinical Psychology, 85(7), 702–711. [PubMed: 28414489]

Husky M, Swendsen J, Ionita A, Jaussent I, Genty C, & Courtet P (2017). Predictors of daily life suicidal ideation in adults recently discharged after a serious suicide attempt: A pilot study. Psychiatry Research, 256, 79–84. [PubMed: 28624676]

Kennard BD, Goldstein T, Foxwell AA, McMakin DL, Wolfe K, Biernesser C, . . . Wentroble E. (2018). As safe as possible (ASAP): A brief app-supported inpatient intervention to prevent postdischarge suicidal behavior in hospitalized, suicidal adolescents. American Journal of Psychiatry, 175(9), 864–872. [PubMed: 30021457]

King CA, Brent D, Grupp-Phelan J, Shenoi R, Page K, Mahabee-Gittens EM, . . . McGuire TC. (2020). Five profiles of adolescents at elevated risk for suicide attempts: Differences in mental health service use. Journal of the American Academy of Child & Adolescent Psychiatry, 59(9), 1058–1068. [PubMed: 31830523]

Kircanski K, Thompson RJ, Sorenson JE, Sherdell L, & Gotlib IH (2015). Rumination and worry in daily life: Examining the naturalistic validity of theoretical constructs. Clinical Psychological Science, 3(6), 926–939. [PubMed: 26783506]

Kleiman EM, Glenn CR, & Liu RT (2019). Real-time monitoring of suicide risk among adolescents: Potential barriers, possible solutions, and future directions. Journal of Clinical Child & Adolescent Psychology, 48(6), 934–946. [PubMed: 31560584]

Kleiman EM, & Nock MK (2018). Real-time assessment of suicidal thoughts and behaviors. Current Opinion in Psychology, 22, 33–37. [PubMed: 30122275]

Kleiman EM, Turner BJ, Fedor S, Beale EE, Huffman JC, & Nock MK (2017). Examination of real-time fluctuations in suicidal ideation and its risk factors: Results from two ecological momentary assessment studies. Journal of Abnormal Psychology, 126(6), 726–738. [PubMed: 28481571]

Klonsky ED, & May AM (2015). The three-step theory (3ST): A new theory of suicide rooted in the "ideation-to-action" framework. International Journal of Cognitive Therapy, 8(2), 114–129.

Miché M, Studerus E, Meyer AH, Gloster AT, Beesdo-Baum K, Wittchen H, & Lieb R (2020). Prospective prediction of suicide attempts in community adolescents and young adults, using regression methods and machine learning. Journal of Affective Disorders, 265, 570–578. [PubMed: 31786028]

Nahum-Shani I, Hekler EB, & Spruijt-Metz D (2015). Building health behavior models to guide the development of just-in-time adaptive interventions: A pragmatic framework. Health Psychology, 34(S), 1209–1219.

Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, & Murphy SA (2017). Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. Annals of Behavioral Medicine, 52(6), 446–462.

Nock MK, Prinstein MJ, & Sterba SK (2009). Revealing the form and function of self-injurious thoughts and behaviors: A real-time ecological assessment study among adolescents and young adults. Journal of Abnormal Psychology, 118(4), 816–827. [PubMed: 19899851]

Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, . . . Shen S. (2011). The Columbia–Suicide severity rating scale: Initial validity and internal consistency findings from three multisite studies with adolescents and adults. American Journal of Psychiatry, 168(12), 1266–1277. [PubMed: 22193671]

Prinstein MJ, Nock MK, Simon V, Aikins JW, Cheah CS, & Spirito A (2008). Longitudinal trajectories and predictors of adolescent suicidal ideation and attempts following inpatient hospitalization. Journal of Consulting and Clinical Psychology, 76(1), 92–103. [PubMed: 18229987]

Ribeiro JD, Franklin JC, Fox KR, Bentley KH, Kleiman EM, Chang BP, & Nock MK (2016). Self-injurious thoughts and behaviors as risk factors for future suicide ideation, attempts, and death: A meta-analysis of longitudinal studies. Psychological Medicine, 46(2), 225–236. [PubMed: 26370729]

Ribeiro JD, Bender TW, Selby EA, Hames JL, & Joiner TE (2011). Development and validation of a brief self-report measure of agitation: The brief agitation measure. Journal of Personality Assessment, 93(6), 597–604. [PubMed: 21999383]

Ribeiro JD, Huang X, Fox KR, Walsh CG, & Linthicum KP (2019). Predicting imminent suicidal thoughts and nonfatal attempts: The role of complexity. Clinical Psychological Science, 7(5), 941–957.

Roberts DR, Bahn V, Ciuti S, Boyce MS, Elith J, Guillera-Arroita G, . . . Thuiller W. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography, 40(8), 913–929.

Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, & Müller M (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics, 12(1), 1–8. [PubMed: 21199577]

Rudd MD, Berman AL, Joiner TE, Nock MK, Silverman MM, Mandrusiak M, . . . Witte T (2006). Warning signs for suicide: Theory, research, and clinical applications. Suicide and Life-Threatening Behavior, 36(3), 255–262. [PubMed: 16805653]

Shiffman S, Stone AA, & Hufford MR (2008). Ecological momentary assessment. Annual Review of Clinical Psychology, 4, 1–32.

Shneidman ES (1993). Suicide as psychache: A clinical approach to self-destructive behavior. Northvale, N.J.: Jason Aronson, Inc.

Thabtah F, Hammoud S, Kamalov F, & Gonsalves A (2020). Data imbalance in classification: Experimental evaluation. Information Sciences, 513, 429–441.

Torous J, Larsen ME, Depp C, Cosco TD, Barnett I, Nock MK, & Firth J (2018). Smartphones, sensors, and machine learning to advance real-time prediction and interventions for suicide prevention: A review of current progress and next steps. Current Psychiatry Reports, 20(7), 4–9. [PubMed: 29387967]

Van Orden KA, Cukrowicz KC, Witte TK, & Joiner TE Jr (2012). Thwarted belongingness and perceived burdensomeness: Construct validity and psychometric properties of the interpersonal needs questionnaire. Psychological Assessment, 24(1), 197–215. [PubMed: 21928908]

Van Orden KA, Witte TK, Cukrowicz KC, Braithwaite SR, Selby EA, & Joiner TE Jr (2010). The interpersonal theory of suicide. Psychological Review, 117(2), 575. [PubMed: 20438238]

Walsh CG, Ribeiro JD, & Franklin JC (2017). Predicting risk of suicide attempts over time through machine learning. Clinical Psychological Science, 5(3), 457–469.

Walsh CG, Ribeiro JD, & Franklin JC (2018). Predicting suicide attempts in adolescents with longitudinal clinical data and machine learning. Journal of Child Psychology and Psychiatry, 59(12), 1261–1270. [PubMed: 29709069]

Wang L, & Miller LC (2020). Just-in-the-moment adaptive interventions (JITAI): A meta-analytical review. Health Communication, 35(12), 1531–1544. [PubMed: 31488002]

Wang SB, Coppersmith DD, Kleiman EM, Bentley KH, Millner AJ, Fortgang R, . . . Nock MK. (2021). A pilot study using frequent inpatient assessments of suicidal thinking to predict short-term postdischarge suicidal behavior. JAMA Network Open, 4(3), e210591. [PubMed: 33687442]

Wenzel A, & Beck AT (2008). A cognitive model of suicidal behavior: Theory and treatment. Applied and Preventive Psychology, 12(4), 189–201.

Wolff JC, Davis S, Liu RT, Cha CB, Cheek SM, Nestor BA, ... & Spirito A. (2018). Trajectories of suicidal ideation among adolescents following psychiatric hospitalization. Journal of Abnormal Child Psychology, 46(2), 355–363. [PubMed: 28349306]

Yen S, Weinstock LM, Andover MS, Sheets ES, Selby EA, & Spirito A (2013). Prospective predictors of adolescent suicidality: 6-month post-hospitalization follow-up. Psychological Medicine, 43(5), 983–993. [PubMed: 22932393]

Yen S, Spirito A, Weinstock LM, Tezanos K, Kolobaric A, & Miller I (2019). Coping long term with active suicide in adolescents: Results from a pilot randomized controlled trial. Clinical Child Psychology and Psychiatry, 24(4), 847–859. stylefix [PubMed: 31064203]
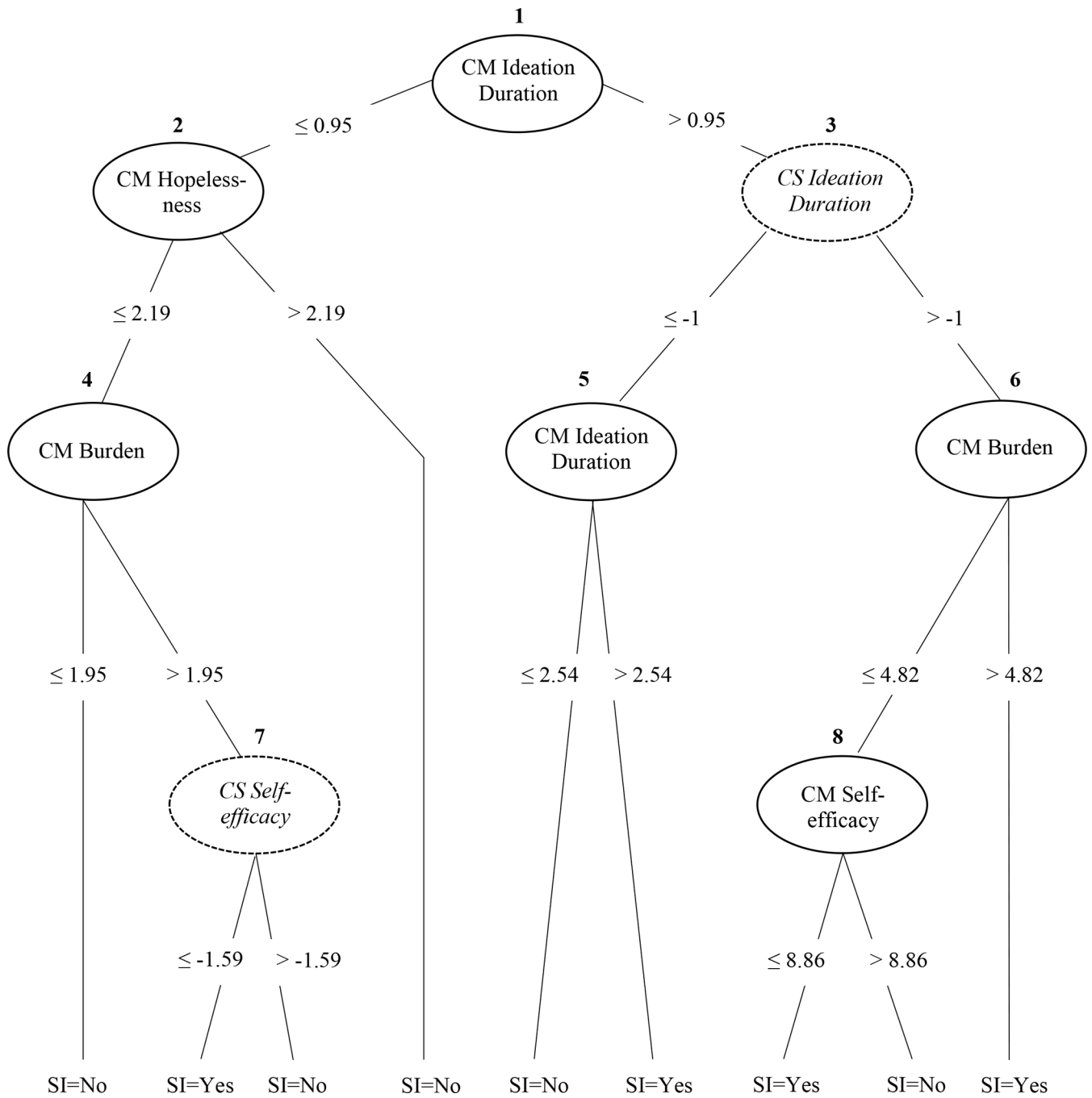
Figure 1.
Prediction rule for next-day suicidal ideation
Notes: SI = Next-day suicidal ideation; CM = Cumulative person-specific mean; *CS* = Change score (deviation from the person-specific mean); Bolded numbers are added to denote nodes for ease of interpretation (see Table 3 for interpretation).
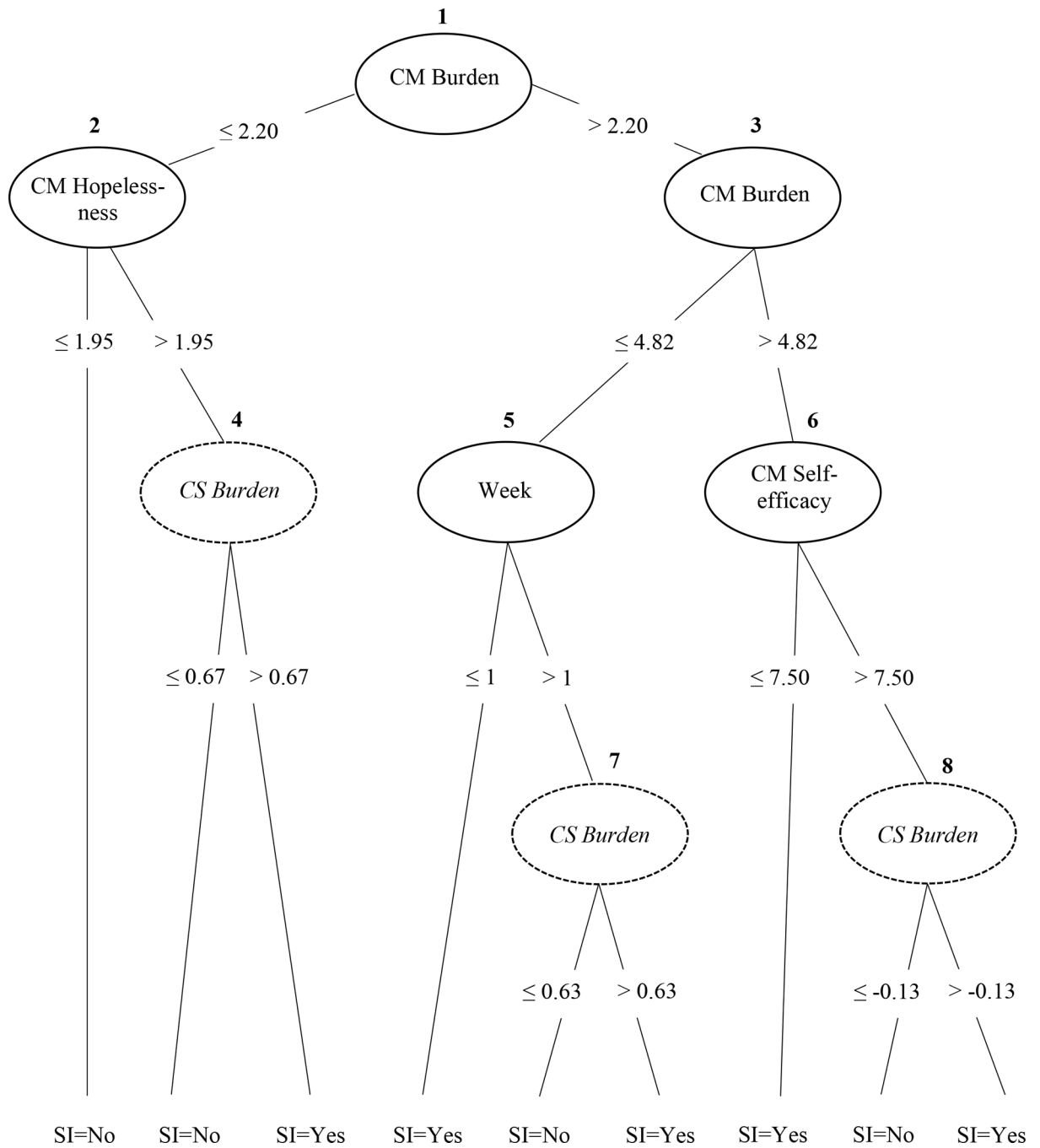
Figure 2.
Prediction rule for next-day suicidal ideation excluding previous-day ideation duration
Notes: SI = Next-day suicidal ideation; CM = Cumulative person-specific mean; *CS* =
Change score (deviation from the person-specific mean); Bolded numbers are added to
denote nodes for ease of interpretation (see Suppl. Table 2 for interpretation).

Table 1.

Performance metrics for models predicting next-day suicidal ideation

	Model based on raw scores	Model based on cumulative mean and change scores
	Mean Metric (SE)	**Mean Metric (SE)**
AUC	0.82 (0.002)	0.86 (0.002)
PPV	0.69 (0.005)	0.74 (0.006)
Sensitivity	0.79 (0.008)	0.81 (0.005)
Specificity	0.78 (0.006)	0.82 (0.006)

Table 2.

Performance metrics for models predicting next-day suicidal ideation excluding suicidal ideation duration as a predictor

	Model based on raw scores	Model based on cumulative mean and change scores
	Mean Metric (SE)	**Mean Metric (SE)**
AUC	0.75 (0.005)	0.78 (0.004)
PPV	0.61 (0.006)	0.60 (0.007)
Sensitivity	0.73 (0.009)	0.84 (0.012)
Specificity	0.71 (0.009)	0.64 (0.014)

Table 3.

Interpretation of the best-performing model shown in Figure 1

Node number (see Figure 1)	Directions
1	a) Within-person cumulative mean of ideation duration greater than 0.95 = Move to node 3 b) Otherwise, move to node 2
2	a) Within-person cumulative mean of hopelessness is greater than 2.19 = No next day ideation b) Otherwise, move to node 4
3	a) Within-person change score of ideation duration is greater than −1 = Move to node 6 b) Otherwise, move to node 5
4	a) Within-person cumulative mean of burdensomeness is greater than 1.95 = Move to node 7 b) Otherwise, no next day ideation
5	a) Within-person cumulative mean of ideation duration is greater than 2.54 = Next day ideation b) Otherwise, no next day ideation
6	a) Within-person cumulative mean of burdensomeness is greater than 4.82 = Next day ideation b) Otherwise, move to node 8
7	a) Within-person change score of self-efficacy is greater than −1.59 = No next day ideation b) Otherwise, next day ideation
8	a) Within-person cumulative mean score of self-efficacy is greater than 8.86 = No next day ideation b) Otherwise, next day ideation