

Protocol

Maximizing Engagement, Trust, and Clinical Benefit of AI-Generated Recovery Support Messages for Alcohol Use Disorder: Protocol for an Optimization Study

Kendra Wyant¹, MS; Sarah J Sant'Ana¹, MS; Claire E Punturieri¹, MS; Jiachen Yu¹, MS; Gaylen E Fronk^{1,2}, PhD; C Michael Maggard^{1,3}; Christopher Janssen¹, BS; Susan E Wanta¹, MA; Rachel Kornfield⁴, PhD; Lyn M van Swol⁵, PhD; John J Curtin¹, PhD

¹Department of Psychology, University of Wisconsin-Madison, Madison, WI, United States

²Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charleston, SC, United States

³Department of Electrical and Computer Engineering, University of Wisconsin-Madison, Madison, WI, United States

⁴Department of Preventive Medicine, Northwestern University, Chicago, IL, United States

⁵Department of Communication Arts, University of Wisconsin-Madison, Madison, WI, United States

Corresponding Author:

John J Curtin, PhD

Department of Psychology

University of Wisconsin-Madison

1202 W Johnson Street

Madison, WI, 53706

United States

Phone: 1 608 262 0387

Email: jjcurtin@wisc.edu

Abstract

Background: Successful recovery from alcohol use disorder requires long-term lapse risk monitoring. Self-monitoring is difficult, given the dynamic, complex interplay of the many risk factors over time. An automated recovery monitoring support system embedded with a machine learning lapse prediction model could improve sustained, adaptive, and personalized self-monitoring by delivering daily support messages.

Objective: We propose to optimize the components included in daily support messages to increase engagement with a recovery monitoring support system.

Methods: The participants will include 304 US adults with moderate to severe alcohol use disorder. Participants will complete daily surveys and provide geolocation data for 17 weeks. Participants will receive daily support messages, starting in week 2, that convey a combination of individualized information from a lapse prediction model. Manipulated message components include (1) lapse probability and lapse probability change, (2) an important model feature, (3) a risk-relevant recommendation, and (4) message personalization on tone preference.

Results: The National Institute on Alcohol Abuse and Alcoholism funded this project (R01AA031762) on August 9, 2024, with a funding period from August 20, 2024, to July 31, 2029. The institutional research board of the University of Wisconsin-Madison Health Sciences approved this project (IRB #2024-0869). Enrollment will begin in December 2025.

Conclusions: Message components that either increase engagement or improve clinical outcomes will be recommended for use in future recovery monitoring support systems and digital therapeutics.

International Registered Report Identifier (IRRID): PRR1-10.2196/81697

(*JMIR Res Protoc* 2025;14:e81697) doi: [10.2196/81697](https://doi.org/10.2196/81697)

KEYWORDS

substance use disorders; precision mental health; continuing care; relapse prevention; MOST; Multiphase Optimization Strategy

Introduction

Alcohol Use Disorder

Alcohol use disorder is highly prevalent and costly. Over 30 million adults in the United States had an active alcohol use disorder in 2021, and 23.3% reported engaging in past-month binge drinking [1]. Alcohol ranks as the third leading preventable cause of death in the United States, accounting for approximately 140,000 fatalities [2,3] and economic costs that exceed US \$249 billion annually [4].

Relapse Prevention

Alcohol use disorder is a chronic condition. Relapse rates following initial treatment are high [5,6]. Lapses (ie, single instances of goal-inconsistent alcohol use) are a necessary precursor for relapse (ie, full return to harmful levels of alcohol use). As a result, preventing lapses and guiding behavior change immediately following a lapse are key goals in both acute and continuing care alcohol use disorder treatment.

Marlatt and Gordon's seminal relapse prevention model [7,8] provides the backbone for contemporary, clinician-delivered interventions for alcohol use disorder (eg, cognitive behavioral therapy [9,10] and mindfulness-based relapse prevention [11]) that have the highest level of empirical support [12-14]. The relapse prevention model provides a detailed framework to understand how emotions, events, and situations can lead to lapses and relapse back to alcohol use. This framework includes both distal and proximal precipitants of lapse such as lifestyle imbalances, craving, high-risk situations, stressors, negative affect, self-efficacy or confidence, and abstinence violation effects. The model's influence is clearly seen in numerous, efficacious intervention modules and supports that are included in cognitive behavioral and mindfulness-based interventions (eg, coping skills training, lapse management, urge surfing, stimulus control techniques, and positive lifestyle changes).

The relapse prevention model highlights that relapse is a complex, nonlinear function of many risk and protective factors that combine and interact to affect relapse timing and severity [15-19]. Many of these factors are transient, leading to fluctuating relapse risk. Urges, mood, lifestyle imbalances, self-efficacy, and motivation all vary over time. Social networks evolve to be more protective or risky. High-risk situations can occur at any time.

Much like other chronic conditions, clinical observation and research suggest that successful recovery from alcohol use disorder requires lifelong monitoring [15,17,19-21]. However, individual (eg, clinician bandwidth) and systemic (eg, overburdened health care system) resources are insufficient for long-term clinician-guided care. Additionally, self-monitoring is difficult, given the dynamic, complex interplay of many risk factors over time.

Recovery Monitoring Support System

An automated recovery monitoring support system that includes embedded machine learning lapse prediction models powered by personal sensing could enable sustained, adaptive, and personalized monitoring. For example, it could help an

individual track their risk of lapsing, alert an individual to important drivers of their lapse risk, and potentially recommend personalized recovery activities and other supports.

Moment-by-moment personal sensing of intra- and interpersonal risk factors for alcohol use disorder is already feasible [22-33]. Self-report sensing methods, like ecological momentary assessment (EMA), offer privileged access into subjective factors, such as craving, affect, valence, and self-efficacy, that may be difficult to quantify reliably through other sensing methods. More novel sensing methods, such as tracking geolocation and cellular communications, could provide a window into risk-relevant information difficult to obtain with self-report. For example, individuals may not have the insight to report on subtle changes in routine or changes to one's social circle. Other information, such as time spent in risky locations, could not possibly be collected solely by self-report without drastically increasing the burden of EMA (eg, by increasing the number of questions or frequency of prompts). Irrespective of the method used, sensing lapse risk factors can provide ongoing monitoring that cannot be accomplished in real time by clinicians and is difficult for patients to implement consistently on their own.

Machine learning models can predict future alcohol use [28,33,34] and lapses [32,35] using these sensed features. Our group has developed a model that uses EMA, collected 4 times daily, to predict future lapses back to drinking in the next day [35]. Critically, the model's performance exceeded general benchmarks for excellent performance [36], with an area under the receiver operating characteristic curve of 0.91, when rigorously assessed on new data from new individuals.

Methods from interpretable machine learning [37] can be used to identify patient-specific predictors of lapse risk at any moment in time [35]. These outputs can be anchored within the relapse prevention model to identify specific interventions and supports that are risk-relevant for each patient—much like a clinician would do if they were available in the moment. For example, during sensed periods of high stress, guided mindfulness and body scans could be recommended. Moreover, these systems can harness valuable information that a clinician could not realistically capture, like real-time changes in activity patterns. If increased time with risky people or locations is driving lapse risk, the recovery monitoring support system could provide information on local Alcoholics Anonymous or other support meetings.

Beyond exceptional performance and interpretable inputs, it is imperative that feedback from these machine learning models is intentionally tailored such that patients use the information and follow its recommendations. Communication research has begun to explore how humans communicate with artificial intelligence and other embedded algorithms [38,39] and how qualities of these communications influence trust, acceptance, and the use of “advice” [39-44].

The degree of transparency of the model's outputs (ie, the system's reasoning) is one potential area for tailoring. Research has demonstrated that making predictions or recommendations from “black-box” machine learning models more transparent can improve perceptions, acceptance, and trust of these

embedded machine learning models [45-51]. In the context of a recovery monitoring support system, some transparency, such as revealing key factors driving an individual's lapse risk, may also promote patient learning and insight to yield additional clinical benefits [46,52-54]. However, in some instances, overly complex or otherwise detailed explanations of how a system works may erode confidence in that system because incremental feedback provides more information than necessary, reveals errors, or leads users to question the system even when it was correct [45,48,51,52,55-57].

Research also suggests that people are more likely to trust, prefer, and engage with automated systems that portray human-like traits, emotions, and intentions [58-60]. Tailoring the linguistic style and tone may be one way to deliver feedback in a manner that more closely resembles an authentic supportive interaction.

Automated communications written with an informal linguistic style have been shown to be related to higher perceptions of human-likeness and trust compared to formal styles. However, this relationship between style and trust is likely context-dependent. In studies where interactions with automated agents are more task- and goal-oriented (eg, during a customer service interaction), informal style was associated with perceived human-likeness, but not trust or overall satisfaction [61,62].

Substantial research exists on human-written (rather than automated) supportive messages and advice. Messages written in tones that acknowledge the recipient's feelings, help the recipient feel accepted, validate experiences of distress, and emphasize the feasibility of following the advice or recommendation have all been shown to be important [63-65]. Given that people tend to anthropomorphize automated messaging systems, these tones are likely important when extending to a supportive messaging system. It is less clear, however, which tones are most appropriate and would be best received in the context of a recovery monitoring support system that delivers important information about lapse risk.

Additionally, important individual differences in linguistic style and tone preferences exist. Research shows that people tend to rate automated messages and agents more favorably when they reflect traits similar to their own (ie, the similarity-attraction effect [66]). These findings have been demonstrated to exist broadly at the group level (eg, culture [67], gender [68,69], and age [70]) and narrowly at the individual level (eg, cognitive and personality style [71,72]).

In the context of digital mental health applications, more broadly, promoting trust is important for eventual therapeutic benefit [73], and mistrust is a key factor in abandonment of mobile health apps [74]. Moreover, declines in user engagement are a known issue in the context of digital mental health supports and may not be exclusively related to trust in the system [75]. Therefore, for a recovery monitoring support system to succeed, we must first explicitly evaluate and optimize the feedback from these embedded machine learning models to maximize patient engagement, trust, and clinical benefit.

Objectives

Our broad goal is to develop a recovery monitoring support system with personalized daily support messages for people with alcohol use disorder. In this project, we propose to optimize the components included in these daily support messages to increase engagement with the support system.

Our approach is guided by the Multiphase Optimization Strategy (MOST) framework [76-78]. MOST has become highly influential for the optimization and evaluation of adaptive and multicomponent interventions in digital health. Its core assertions are that (1) interventions should be explicitly optimized to meet specific criteria; (2) intervention optimization and intervention evaluation are different phases of research, pursue different specific aims, and require different methodological approaches; and (3) the optimization of an intervention should precede its evaluation. The research in this project focuses on the MOST optimization phase. Completion of study objectives sets the stage for future, programmatic MOST research to evaluate optimized smart digital therapeutics using appropriate designs (eg, randomized controlled trial) for subsequent phases of development.

We will manipulate 4 candidate components of daily support messages that convey transparent, individualized, risk-relevant information from our machine learning lapse prediction model to participants. These message components include (1) the user's current lapse probability for that day and trends in that daily lapse probability over the past 2 weeks, (2) an important lapse feature contributing to their current lapse probability, (3) a risk-relevant recommendation for a recovery activity to complete that day, and (4) linguistic style and tone personalization for the support message. These components use output that would be available from any machine learning lapse prediction model such that conclusions about the impact of these components on engagement can generalize beyond our system.

Models will be optimized on a measure of engagement (days of engagement with the daily support messages). We will use a MOST factorial experiment to determine which of the 4 message components encourages individuals to continue engaging with the support system. We will also look at secondary optimizing outcomes of trust in the support system, perceived usefulness of the support messages, and digital working alliance with the support system, as these are likely to drive long-term engagement.

In addition to the engagement outcome, we will test message component effects on clinical outcomes because the risk-relevant information from our machine learning model may provide direct benefits to participants through mechanisms other than engagement (eg, information about relapse processes highlighted by important model features may promote gradual adaptive lifestyle and behavioral adjustments that are difficult to quantify). Message components that either increase engagement or improve clinical outcomes will be recommended for use in future recovery monitoring support systems and digital therapeutics.

Methods

Overview

The primary goal of this study is to optimize daily support messages from our Smart Technology for Addiction Recovery (STAR) recovery monitoring support system to increase participant engagement with the system. To this end, participants will receive daily support messages from the STAR system for 16 weeks. We will manipulate 4 separate components of these support messages independently in a fully crossed, 4-way factorial, between-subjects design. Each component is operationalized to be binary (eg, include or exclude) so that each is either included or excluded in the support message delivered to the participant. This yields 16 between-subjects conditions, with each participant assigned to receive support messages defined by one of these conditions for the entire study period.

The first 3 components of the support messages involve personalized information based on output from the lapse risk prediction model embedded within the STAR system. The three prediction model-based components are (1) the user's current lapse probability for that day and trends in that daily lapse probability over the past 2 weeks, (2) an important lapse feature contributing to their current lapse probability, and (3) a risk-relevant recommendation for a recovery activity to complete that day.

The fourth component of the support message is linguistic style and tone personalization. The support messages are created using a large language model (LLM) that takes the relevant prediction model-based components as inputs along with a prompt that dictates the linguistic tone and style of the message. Participants either receive support messages using a linguistic tone and style that matches their prespecified preferences or is yoked to another participant's preferences.

The primary analyses for this study evaluate the main effects and 2-way interactions among these 4 message components on days of engagement with the STAR system daily support messages. The secondary analyses examine the effects of the components on other perceptions of the STAR system (message usefulness, system trust, and system digital working alliance) and clinical outcomes (eg, days of drinking and heavy drinking and flourishing).

Participants

Participants must be 18 years of age or older and meet criteria for alcohol use disorder with at least moderate severity (≥ 4 *Diagnostic and Statistical Manual of Mental Disorders [Fifth Edition; DSM-5]* symptoms). All participants will report a goal of abstinence from alcohol, with their most recent use of alcohol

between 1 week and 3 months in the past at study intake [79]. Participants with medical and psychiatric comorbidities will not be excluded. However, participants must be able to read in English and will be excluded if they have disabilities that prevent the use of a smartphone (eg, uncorrected vision, hearing problems, or profound cognitive impairment). Participants must have a smartphone and cellular plan.

We will recruit 304 individuals to participate for up to 17 weeks. This includes a 1-week burn-in period to collect input features prior to using the lapse risk prediction model for personalized support messages followed by 16 weeks of daily support messages from the STAR system combined with data collection of study outcomes. Participants will be recruited using social media and through laboratory-affiliated treatment centers. A diverse sample with respect to age, sex, race or ethnicity, and population density (eg, urban, suburban, and rural regions) will be recruited. We will track recruitment success and make ongoing adjustments to meet our target sample characteristics throughout the recruitment period.

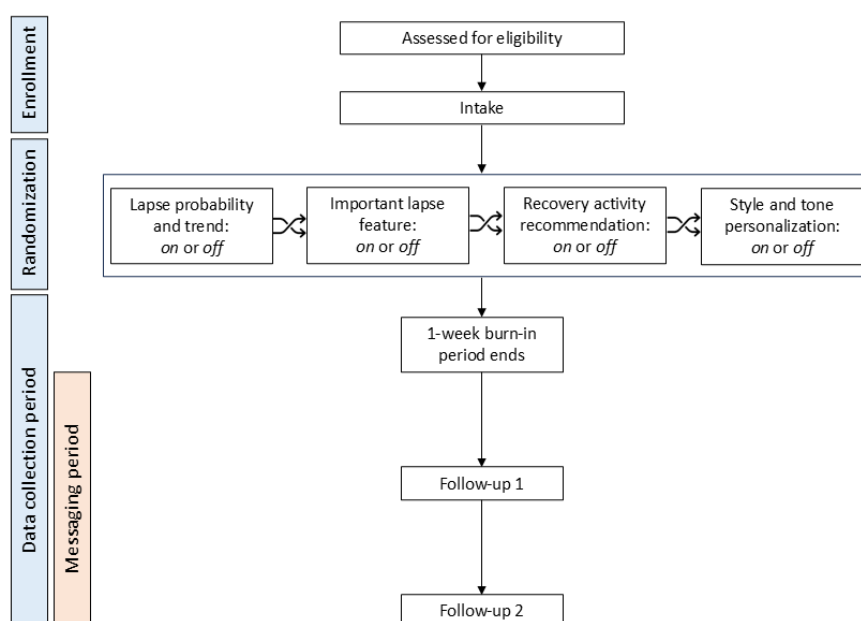
Participants will be paid US \$40 for the intake visit and US \$20 for phone visits at 8 and 16 weeks into the data collection period. Participants will also be paid up to US \$105 each month on study for completing daily EMAs (up to US \$50 per month depending on adherence), for sharing geolocation data (US \$5 per month), and to offset costs of cell phone service (US \$50 per month).

Procedure

Participants will first complete an intake session by phone or videoconference according to their preference. During this session, study staff will describe study procedures, requirements, and participant compensation. Study staff will also confirm the participant's alcohol use disorder diagnosis (via module E of the Structured Clinical Interview for *DSM-5*) and assess other inclusion or exclusion criteria to determine study eligibility (eg, abstinence goal and duration). Eligible participants will be randomly assigned to 1 of 16 possible message conditions (a crossed factorial design of 4 message components included or not included in the messages).

Following the intake session, participants will start to complete daily EMAs and provide geolocation data. After a 1-week burn-in period, participants will begin receiving personalized daily support messages based on output from the lapse risk prediction model embedded in the STAR system (see Daily Support Messages section). Participants will continue to complete EMAs, share geolocation, and receive daily support messages for 16 weeks following the burn-in period. Participants will complete 2 phone follow-up visits at 8 and 16 weeks into this data collection period to measure other study outcomes. Figure 1 presents a flow diagram of study participation.

Figure 1. Participant flow diagram. During enrollment, participants are assessed for eligibility and complete baseline measures. Eligible participants are subsequently randomized to 1 of 16 possible messaging conditions. Each of the 4 message components, lapse probability and trend, important lapse feature, recovery activity recommendation, and style and tone personalization, can be turned on or off (ie, $2 \times 2 \times 2 \times 2$ factorial design). Participants then immediately begin providing ecological momentary assessment and geolocation data. After a 1-week burn-in period, participants begin receiving daily support messages. Participants complete 2 follow-up visits at 8 and 16 weeks after messaging begins. Each follow-up visit consists of an interview and a battery of self-report measures.



Lapse Risk Prediction Machine Learning Model

Following the methods and analysis workflow used in Wyant et al [35], we have developed a machine learning model that uses both geolocation and daily EMAs to predict the probability of a lapse back to alcohol use in the next day. This elastic net model was trained using approximately 300,000 labeled observations from 146 participants with moderate to severe alcohol use disorder [31,35]. Critically, the model has high combined sensitivity and specificity as indicated by its area under the receiver operating characteristic curve that exceeds 0.90 for held-out observations (using grouped k-fold cross-validation). This model can be used to provide both the probability of a lapse back to alcohol use in the next 24 hours and the change in that lapse probability over the past 2 weeks (eg, increasing, decreasing, or stable). A complete description of the model, including model coefficients and feature engineering details, is available on the Open Science Framework (OSF) repository for this project [80].

Local feature importance [37] for risk categories (eg, craving, stress, and risky situations) is quantified in log-odds using model coefficients and scores for features within each risk category that are associated with each model prediction (ie, a predicted lapse probability for a specific participant on a specific day). Locally important features are features that substantially increase or decrease the predicted probability of a lapse for that participant at that time from the aggregate lapse probability (ie, average probability across all participants and times in the original training data). We consider features that increase or decrease the log-odds of a lapse by $>|0.1|$, clinically important for binary classification models [37]. By extracting local (ie, for a specific, single prediction) feature importance, we can use the model to determine not only the probability of a lapse and

its change over time but also information about why that lapse is likely or unlikely to happen.

Our model features tap into key constructs from the relapse prevention model such as craving, affect, stressors, lifestyle imbalances, high-risk situations, self-efficacy or confidence, and abstinence violation effects. This allows us to use the relapse prevention model to identify interventions and other supports that are personally relevant to that participant at that moment in time to address their risk. Following methods implemented by Fisher et al [81] and Fernandez et al [82] (also see [83-86]), combinations of lapse probability (categorized into low, moderate, or high risk) and important feature categories are mapped to specific recommendations using a mapping matrix based on clinical expertise guided by the relapse prevention model. Combinations of lapse probability and important features (ie, cells of the mapping matrix) often map to more than 1 appropriate module. For example, if craving is important and lapse probability is low, multiple useful urge management recommendations exist (eg, guided urge surfing and distracting games and activities). Thus, this procedure allows us to use the lapse prediction model to identify the set of interventions and supports that are most personally relevant to the participant at that moment in time, given their recent experiences, current lapse probability, and important risk features.

Daily Support Messages

Overview

All participants will receive daily support messages beginning after the 1-week burn-in period. Each day, participants are sent an SMS text message that will contain a Qualtrics (Qualtrics XM) link. The daily support message is accessed and viewed or read in Qualtrics by clicking on that link. The use of Qualtrics

to display the support message allows us to have a time-stamped confirmation that the support message has been accessed that day. Text messages with this Qualtrics link will be sent out at 5 AM each morning in the participant's own time zone so that the support message is available at the start of the new day upon waking. A reminder is also sent by text message at 11 AM and 4 PM if the participant has not yet accessed the support message in Qualtrics by those times.

The support messages are generated for each participant each night using an LLM (GPT-4o) that is accessed by the STAR system through the Microsoft 365 Copilot application programming interface. LLM prompts and example messages are provided on the OSF repository for this project [80]. The core of each support message is a simple statement that encourages the participant to engage with their recovery that day. In addition to the core statement, the message may include up to 4 additional components of personalized information derived from their linguistic tone or style preferences and the lapse risk prediction model to guide their recovery efforts that day. These components are (1) their current lapse probability for that day and trends in that daily lapse probability over the past 2 weeks, (2) an important lapse feature contributing to their current lapse probability, (3) a risk-relevant recommendation for a recovery activity to complete that day, and (4) linguistic style and tone personalization. These 4 message components are operationalized as follows.

Current Lapse Probability and Trends in Daily Lapse Probability Over the Past 2 Weeks

If included, the support message will contain categorical information about the participant's lapse probability in the next 24 hours based on output from the lapse risk prediction model. Lapse probability will be provided using a 3-level categorization of "low risk" ($P \leq .05$), "moderate risk" ($.05 < P \leq .20$), or "high risk" ($P > .20$). These probability cut points were determined both from clinician input and distributional information in the sample of 151 participants with alcohol use disorder where we developed our model [35]. The cut points represent approximately the 10th and 60th percentiles in the overall distribution of the ~300,000 observed probabilities. The decision to convey lapse probability categorically (rather than numerically) follows current best practice recommendations to avoid complex numerical probability statements, especially for those with low numeracy, and provide the minimal information needed for individuals to assess the magnitude of risk, weigh options, and act [87-91]. It also follows the format from an existing app that characterizes the risk associated with various blood alcohol concentration levels [92].

If included, the support message will also contain categorical information about the trend in the predicted probabilities of a lapse over the past 2 weeks. Lapse probability change will be reported using a 3-level categorization of "decreasing risk" (lapse probability decreases by $\geq .05$ over 2 weeks), "increasing risk" (probability increases by $\geq .05$ over 2 weeks), or "stable risk" (probability changes by $< .05$ over 2 weeks). Lapse probability change is quantified as the difference in mean weekly predicted lapse probability over the 2 previous weeks (ie, week 2 mean-week 1 mean). We use a simple weekly mean difference

to capture both linear and more complex (eg, increasing monotonic) trends across these 2 weeks. The 2-week lapse probability trend is provided because trajectory information may make subtle changes in risk level more apparent, can relieve cognitive load associated with monitoring over time, and can provide incremental value beyond absolute risk levels alone [93]. The decision to convey this information categorically follows the same best practice recommendations as for lapse probability [87-92].

Important Lapse Feature

If included, the support message will contain a description of 1 important model feature category that contributes to the participant's predicted lapse probability that day. The feature category will be selected from among the set of important feature categories that day as defined based on the categories' log-odds values (ie, feature categories that increase or decrease the log-odds of a lapse by $> |0.1|$). Among this set of important features, features will be ranked or sorted by how often they have been included in recent support messages and then the magnitude of their importance score. This will prioritize providing the participant with information about new but also important risk features. We include this message component because providing risk feature information may help participants both identify which issues to target for change and select appropriate support tools matched to these issues [94]. Emerging best practices from interpretable machine learning suggest that this information may make the machine learning model less of a "black box," which may build trust, or allow individuals to engage more thoughtfully (eg, questioning false positives and scrutinizing model behavior) and gain insight [95-97].

Risk-Relevant Recommendation for a Recovery Activity

If included, the support message will contain a suggestion to consider completing a specific recovery activity that has been identified by the lapse risk prediction model as personally risk-relevant, given their lapse probability and important model features. All recovery activities are drawn from existing empirically based treatment protocols for relapse prevention, including matrix [98], cognitive-behavioral coping skills therapy [99], and motivational enhancement therapy [100]. As with the feature information mentioned earlier, when multiple risk-relevant recovery activities are available and appropriate, given the lapse risk prediction model output, activities that have not been recommended recently will be selected. We include this message component following best practice suggestions to pair risk information with specific action recommendations to address the risk [101]. This component follows directly from the guiding thesis of our research that improved clinical outcomes will result from increases in adaptive, personally relevant engagement rather than simply more engagement. This component also follows recent examples of digital therapeutics that used machine learning prediction models to generate intervention recommendations for medical or health issues other than alcohol use disorder [102,103].

Linguistic Style and Tone Personalization

In addition to the relevant lapse risk prediction model-based message components for that participant, the LLM prompt for

support message creation also includes details about the linguistic style (formal or informal) and tone (legitimizing, caring and supportive, self-efficacy support, acknowledging, value affirming, and normalizing) for the support message. These styles and tones were selected based on research on linguistic factors that can affect the acceptance and use of advice during algorithmic or computer-mediated communications [44,104], medical decision-making [105-107], and communications more generally [108-111].

At intake, participants rate how much they would like to receive messages written in each of the available tones and styles independently on a 7-point Likert scale with 1=strongly disagree and 7=strongly agree as anchors. Participants are assigned to either receive support messages matched to their preferences

(ie, across days, support messages will be written using any of the tones or styles that were rated higher than the neutral midpoint of the scale) or yoked to receive messages that match the preferences of another participant.

Measures

Overview

Detailed descriptions of the measure items, sources, and administration are available on the project's OSF repository [80]. Our description of study measures is organized into 5 categories: demographic and other stable characteristics, lapse prediction model inputs, primary or optimization outcome, secondary system outcomes, and secondary clinical outcomes. A summary of these measures is also presented in Table 1.

Table 1. Measure constructs, sources, frequency, and use.

Constructs	Source	Measurement frequency	Measure use
Geolocation	FollowMee app	Continuous	Model input
Daily experiences and events (eg, craving, affect, stressors, and risky situations)	EMA ^a	Daily	Model input
System engagement	Confirmation support message accessed	Daily	Primary or optimization outcome
Demographics	Laboratory-created	Intake visit	Baseline characterization or covariate
Alcohol use disorder symptoms	<i>Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition)</i>	Intake visit	Baseline characterization or covariate
Alcohol use history	Laboratory-created	Intake visit	Baseline characterization or covariate
Lifetime substance use	World Health Organization's Alcohol, Smoking, and Substance Involvement Screening Test	Intake visit	Baseline characterization or covariate
Trust in automated systems	Adapted from Propensity to Trust Questionnaire and Trust in Automation Scale	Intake visit	Baseline characterization or covariate
Message usefulness	Laboratory-created	8- and 16-week follow-up visits	Secondary system outcome
System trust	Adapted from Trust of Automated Systems Test and Trust in Automation Scale	8- and 16-week follow-up visits	Secondary system outcome
System digital working alliance	Digital Working Alliance Inventory	8- and 16-week follow-up visits	Secondary system outcome
Number of drinking days (past 30 days)	Alcohol Timeline Follow-Back and EMA	8- and 16-week follow-up visits	Secondary clinical outcome
Number of heavy drinking days (past 30 days)	Alcohol Timeline Follow-Back and EMA	8- and 16-week follow-up visits	Secondary clinical outcome
Anxiety	Generalized Anxiety Disorder-7	Intake and 8- and 16-week follow-up visits	Secondary clinical outcome
Depression	Patient Health Questionnaire-9	Intake and 8- and 16-week follow-up visits	Secondary clinical outcome
Human flourishing	Flourishing Measure	Intake and 8- and 16-week follow-up visits	Secondary clinical outcome
Recovery capital	Multidimensional Inventory of Recovery Capital	Intake and 8- and 16-week follow-up visits	Secondary clinical outcome

^aEMA: ecological momentary assessment.

Demographic and Other Baseline Characteristics

At the intake visit, we will collect self-report information about demographics (age, sex at birth, gender identity, sexual orientation, race, ethnicity, household income, education level, marital status, and number of individuals living in the household), *DSM-5* alcohol use disorder symptoms, general alcohol use history characteristics (eg, age of first use, years of regular use, number of quit attempts, and previous treatment received), and lifetime substance use (World Health Organization's Alcohol, Smoking and Substance Involvement Screening Test [112]). We also measure individual differences in generic trust in automated systems using a subset of items from the Propensity to Trust Questionnaire (6 items [113]) and the Trust in Automation scale (7 items [114]).

Lapse Prediction Model Inputs

Input features for the lapse risk prediction model are engineered from two sources: (1) 1 time daily EMA and (2) passively sensed geolocation. Participants will complete 1 brief (less than 1 minute to complete) EMA each day. EMAs will be sent to the participant by SMS text message at 5 PM each night in the participant's own time zone. These SMS text messages will include a link to a Qualtrics survey optimized for completion on their smartphone. A reminder to complete the survey is sent by text message at 7 PM if a participant has not yet completed it.

On each EMA, participants will report their current mood (ie, affective valence and arousal), their peak alcohol craving since their last EMA, and the occurrence and intensity of any positive events and any stressors or hassles since their last EMA. They also report how likely they are to encounter risky situations (people, places, or things) and pleasant and stressful events in the next week and how confident they are about abstaining from alcohol use in the next week. They conclude each EMA by reporting any alcohol use that they have not yet reported, the time of the start of that use, the duration of the drinking episode, and the number of drinks consumed.

Participants' location will be continuously sensed passively using the FollowMee (FollowMee LLC) geolocation tracking app. We will increase the predictive signal from geolocation by gathering contextual information about the locations that participants visit frequently (≥ 2 times per month). For frequent locations, participants will report the location type (eg, home of friend, bar, restaurant, workplace, and Alcoholics Anonymous meeting location), if alcohol is typically present at that location, if they drank there previously, their typical emotional experience at that location (pleasant, unpleasant, mixed, or neutral), and the perceived risk to their recovery associated with that location. This contextual information can be gathered quickly by appending these additional questions about a newly detected frequent location to their next daily EMA.

Primary or Optimization Outcome

Our primary outcome for this study is STAR system engagement. System engagement is measured by counting the days that participants access the daily support message by following the link to the support message provided to them by text message. We will calculate message engagement scores

that count the number of days the support message is accessed across the full 16 weeks of data collection. Support message engagement serves as the optimization outcome for this study.

Secondary System Outcomes

We measure 3 secondary STAR system outcomes at 8 and 16 weeks into the data collection period: daily support message usefulness, system trust, and system digital working alliance. Daily support message usefulness is measured with 7 items scored on a 7-point Likert scale. These items ask participants to rate the perceived helpfulness, general liking, and personal relevance of the messages [115-118]. Trust in the overall STAR recovery monitoring support system is measured with 10 items scored on a 7-point Likert scale. These items come from the Trust of Automated Systems Test (9 items [119]) and the Trust in Automation scale (1 item [114]). The system digital working alliance between the participant and the STAR system is measured with the Digital Working Alliance Inventory [120]. This measure consists of 6 items measured on a 7-point Likert scale.

Secondary Clinical Outcomes

We will measure 2 primary clinical outcomes recommended by the Food and Drug Administration [121] for the evaluation of interventions for alcohol use disorder: number of drinking days and number of heavy drinking days in the past 30 days. These 2 outcomes are quantified at 8 and 16 weeks into the data collection period. Information about participant alcohol use is obtained from 2 sources. First, participants report episodes of alcohol use in the daily EMAs as described earlier. Second, study staff conduct a 30-day alcohol Timeline Follow-Back (TLFB) procedure [122] at the study phone visits at 8 and 16 weeks in the data collection period. The TLFB procedure is a calendar-assisted retrospective reconstruction of how many standard alcoholic drinks were consumed by the participant each day in the assessment period. This procedure is further enhanced by adding previously reported drinking episodes (dates, times, and number of drinks) from the daily EMAs to the calendar used with the participant. This provides an opportunity to validate those previous reports and collect any additional drinking episodes that may have been missed by the EMAs. The number of drinking days is defined as the number of days that any alcohol is consumed during the relevant 30-day period. The number of heavy drinking days is defined as the number of days that more than 3 or 4 standard drinks are consumed for women and men, respectively, during the relevant 30-day period. The TLFB will also be administered by study staff during the intake visit to determine inclusion or exclusion criteria (ie, time since last drink).

We will also collect secondary measures of anxiety (General Anxiety Disorder-7 [123]), depression (Patient Health Questionnaire-9 [124]), human flourishing (total score and subscales from the Flourish Measure [125]), and recovery capital (Multidimensional Inventory of Recovery Capital [126]). Each of these measures will be collected at intake and at 8 and 16 weeks into the data collection period.

Data Analytic Plan

Overview

All data preprocessing, visualization, and exploratory analyses will be done using the tidyverse ecosystem [127] in R (R Foundation for Statistical Computing). General or generalized mixed effects models will be fit using the *lme4* package [128] in R. Multiple imputation will be performed using the *mice* package [129].

Primary Analyses

MOST [77] highlights the importance of optimizing any intervention prior to evaluating it with a formal randomized controlled trial. The goal of this project is to optimize support message components in the STAR system by identifying which of the several message components increase engagement with the daily support messages. Such information will be crucial to develop future digital therapeutics for alcohol and other substance use disorders that use personalized supportive messages based on machine learning lapse risk prediction models.

Collins [130] advocates for the use of the factorial experiment as a highly efficient and powerful design to optimize intervention components generally. We use this design to evaluate the effects of our 4 message components (current lapse probability and trends in daily lapse probability over the past 2 weeks, important feature contributing to current lapse probability, risk-relevant recommendation for a recovery activity, and linguistic style and tone personalization) on the primary or optimization outcome (count of days of support message engagement) measured at the end of the data collection period (week 16). The goal of these analyses is to determine the effect of each support message component and whether the effect of one component varies depending on the level or setting of another component (eg, Is the effect of a recovery activity recommendation greater if given in combination with lapse probability or an important feature?).

We will analyze support message engagement in a generalized linear model (with Poisson or negative binomial distribution dependent on the distribution diagnostics for the count data). We will use unit-weighted orthogonal contrasts to code for main effects and 2-way interactions among the 4 message components. Support message engagement will be indexed as the count of days on which the support message was accessed by the participant across the full 16 weeks in the data collection period. Baseline covariates will be included in the model, as described in the Covariates section. This analysis will include all participants who were randomized into the study, following intention-to-treat principles. If participants discontinue use of the STAR system during the data collection period, they will get a 0 for each subsequent day they do not engage with the message.

Secondary Analyses

Analyses for secondary outcomes will follow the same analytic plan as for the primary analyses but will use the secondary outcomes (eg, daily support message usefulness, system trust, system digital working alliance, and clinical outcomes) as the

outcome variable. These analyses will allow us to determine whether the effects of message components on engagement also extend to other perceptions of the STAR system and clinical outcomes. The error distribution (Gaussian, Poisson, and negative binomial) for these generalized linear models will be selected based on the distribution of the outcome variable. These secondary analyses will use mean scores for each outcome across the 16 weeks (ie, average of scores at 8 and 16 weeks) as the dependent measure because the effects of time are not a primary focus, given the relatively short duration of the study. These analyses will only include participants who provided at least 1 measurement for the secondary outcomes during the data collection period (ie, at 8 or 16 weeks). For participants missing 1 measurement, scores will be imputed using multiple imputation (see Missing Data section).

Covariates

The use of covariates has been demonstrated to improve estimation efficiency and increase power to test parameter estimates in linear models [131-135]. Given this, the inclusion of baseline characteristics measured prior to assignment to levels for message components will be considered as covariates in the primary and secondary analyses. When available, the baseline scores associated with the study outcome for each model (eg, baseline flourishing for the model examining flourishing) will be included in the model to control for baseline differences on that measure. In addition, following well-established practices in our laboratory [136-138], other baseline characteristics (eg, demographics and alcohol use or history measures and baseline scores on other outcomes) will be included in the model if they demonstrate a significant relationship with the outcome variable for that model, independent of (ie, controlling for) message component effects. This allows for the selection of covariates that will increase statistical power but not bias the parameter estimates for the message component effects.

Sample Size Planning

The primary optimization outcome (count of days of support message engagement) will be analyzed in a generalized linear model (with Poisson or negative binomial distribution) and contrast coding for the main effects and 2-way interactions of the 4 message components. We planned the sample size for this study by simulating the power to detect main effects or interactions of the message components across varying effect sizes and baseline (all message components off) support message engagement rates. We simulated count data using the Poisson distribution for support message engagement across 112 days (16 weeks), with λ (mean count for distribution) calculated as the number of days of observation multiplied by the baseline engagement rate. We added a message component effect to these data by increasing the count to reflect an increased rate of support message engagement across days (eg, on 5% more days) when the component was included in the support message. These simulations indicated that 304 participants would provide 85% power to detect a message component main effect or interaction that increased support message engagement rates by 3% of days when the baseline engagement rate was 85% of days. Power was higher still when message effects were larger (ie, >3%) or the baseline engagement rate was lower (eg, <85%

of days). We believe that message components that increase engagement rates by less than 3% would not be sufficiently large to warrant including that component in future versions of the STAR system. Furthermore, if baseline engagement exceeds 85%, there would not be much need for support messages to increase engagement further. Therefore, we believed these bounds to be appropriate to represent the minimal clinically meaningful contribution from message components. It should also be noted that this power estimate is likely conservative because it does not include potential covariates that will be included in the analysis model.

Missing Data

We pursue a variety of methods to minimize biases that can occur when participants are dropped from analyses because of discontinued participation or missing data. First, we follow guidelines and procedures that have been recommended for clinical trials [139,140].

Second, our primary optimization outcome (support message engagement) will use all participants who were randomized to the study because days of support message engagement can be measured for all participants who were randomized, even if they do not complete the full 16 weeks of data collection.

Third, our Food and Drug Administration–recommended clinical outcomes (days of drinking and heavy drinking) are somewhat robust to missing data because they can be scored independently from 2 separate methods (daily EMA and TLFB at 8- and 16-week visits). Ideally, both measurement methods will contribute to the most reliable measurement of these outcomes. However, either is sufficient to score these clinical outcomes and can be used in isolation if the other method is missing (due to missing periods for daily EMA or missing study visits).

Finally, our secondary analyses use outcomes that are averaged across 2 measurement periods. This will provide a more reliable measurement for the outcomes. In instances where participants are missing a measurement, we will use multiple imputation to estimate the missing score. Multiple imputation does not attempt to estimate each missing value with a single imputed value but instead uses a random sample of imputed values. This allows for valid statistical inferences that properly reflect the uncertainty that results from missing values, such as valid SEs and CIs for parameters [141]. In brief, we will use *mice* to impute missing values in the dataset using predictive mean matching. Five distinct imputed datasets will be generated. Relevant statistical models for each secondary analysis will be fit to these imputed datasets, and parameter estimates will be averaged.

Ethical Considerations

The institutional research board of the University of Wisconsin-Madison Health Sciences approved this project (IRB #2024-0869). All participants will provide written informed consent and will be told they can discontinue at any time without penalty. Participants will share location data through the use of a third-party app (FollowMee). Their location data will only be identified by a randomly generated Device ID created when the app is installed. Location data are encrypted and sent to the FollowMee servers, where it will be automatically deleted after

14 days. The University of Wisconsin Cybersecurity Office has reviewed the FollowMee software, its security practices, data loss history, and encryption system for transmitting data and judged the risk of a data breach to be low. All raw data (including geolocation) will be stored on Health Insurance Portability and Accountability Act (HIPAA)–compliant servers protected behind a firewall. Raw data will be labeled with an anonymous study identification number. We also have obtained a certificate of confidentiality, which prevents the disclosure of identifiable, sensitive research information to anyone not connected to the study except with explicit consent by the participant. While we use an LLM to generate personalized support messages, no identifiable participant data will be passed into the LLM. It will only be given up to 4 pieces of information about the participant: lapse probability (high, moderate, or low), trend in lapse probability (increasing, decreasing, or stable), one important lapse feature category selected from several categories, and preferences for style and tone.

Results

The National Institute on Alcohol Abuse and Alcoholism funded this project (R01AA031762) on August 9, 2024, with a funding period from August 20, 2024, to July 31, 2029. Enrollment will begin in December 2025. We plan to recruit participants for approximately 3 years.

Discussion

Clinical observation and research suggest that successful recovery from alcohol use disorder requires long-term and perhaps lifelong self-monitoring of lapse risk. Self-monitoring, however, is difficult given the dynamic and complex interplay of numerous risk factors over time.

Machine learning lapse prediction models are now emerging. Recovery monitoring support systems that synthesize these developments can potentially guide patients toward personally risk-relevant engagement that is adaptive, efficient, and more effective. Research is needed to determine the optimal feedback from embedded machine learning prediction models such that patients make the best use of these automated capabilities.

The optimization study in this project is guided by the MOST framework. The research in this project focuses on the MOST optimization phase. We propose to optimize feedback from a lapse prediction model (via daily engagement messages) both to increase engagement with a recovery monitoring support system and to improve clinical outcomes. We will evaluate the effects of 4 message components (lapse probability or lapse probability change, an important feature, a risk-relevant module recommendation, and tone personalization) that comprise the daily engagement messages and can make the output from our machine learning model more transparent to participants. Completion of our specific aims sets the stage for future, programmatic MOST research to evaluate optimized recovery monitoring support systems using appropriate designs (eg, randomized controlled trial) for that subsequent phase of development.

Additionally, the optimization components use output that would be available from any machine learning lapse prediction model such that conclusions about the impact of these components on engagement can generalize beyond our specific machine learning model. Similarly, engagement messages including these message components could be used in any recovery monitoring system

or digital therapeutic for alcohol use disorder, allowing conclusions to generalize to current and future system variants. At the conclusion of the grant period, we will also deliver this optimized recovery monitoring support system as a tangible product and model for how to embed sensing and machine learning into other existing systems.

Acknowledgments

This protocol is funded by the National Institute on Alcohol Abuse and Alcoholism (R01AA031762). The project has been funded for a 5-year period starting in August 2024.

Data Availability

The datasets generated or analyzed during this study will be available in the Open Science Framework repository upon completion of the study [80].

Authors' Contributions

JJC contributed to funding acquisition, methodology, writing the original draft, supervision, and software. KW contributed to the methodology and writing of the original draft. SJS contributed to the methodology. CMM and CJ contributed to the software. SEW contributed to project administration. All authors contributed to the conceptualization and reviewing and editing the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Peer review report by the National Institute on Alcohol Abuse and Alcoholism (R01AA031762).

[PDF File (Adobe PDF File), 154 KB-Multimedia Appendix 1]

References

1. 2021 National Survey on Drug Use and Health (NSDUH-2021-DS0001). SAMHSA Center for Behavioral Health Statistics and Quality. Substance Abuse and Mental Health Services Administration URL: <https://www.samhsa.gov/data/data-we-collect/nsduh-national-survey-drug-use-and-health/datafiles/2021-2023> [accessed 2025-10-21]
2. Centers for Disease ControlPrevention (CDC). Annual average for United States 2011-2015 alcohol-attributable deaths due to excessive alcohol use, all ages. 2022 Alcohol Related Disease Impact (ARDI) Application Website. URL: <https://tinyurl.com/3x6z9pc9> [accessed 2025-10-21]
3. Esser MB, Leung G, Sherk A, Bohm MK, Liu Y, Lu H, et al. Estimated deaths attributable to excessive alcohol use among US adults aged 20 to 64 years, 2015 to 2019. JAMA Netw Open. 2022;5(11):e2239485. [FREE Full text] [doi: [10.1001/jamanetworkopen.2022.39485](https://doi.org/10.1001/jamanetworkopen.2022.39485)] [Medline: [36318209](https://pubmed.ncbi.nlm.nih.gov/36318209/)]
4. Substance Abuse and Mental Health Services Administration (US), Office of the Surgeon General (US). Facing Addiction in America. Washington, DC. US Department of Health and Human Services; 2016.
5. Dennis M, Scott CK. Managing addiction as a chronic condition. Addict Sci Clin Pract. 2007;4(1):45-55. [FREE Full text] [doi: [10.1151/ascp074145](https://doi.org/10.1151/ascp074145)] [Medline: [18292710](https://pubmed.ncbi.nlm.nih.gov/18292710/)]
6. McLellan AT, Lewis DC, O'Brien CP, Kleber HD. Drug dependence, a chronic medical illness: implications for treatment, insurance, and outcomes evaluation. JAMA. 2000;284(13):1689-1695. [doi: [10.1001/jama.284.13.1689](https://doi.org/10.1001/jama.284.13.1689)] [Medline: [11015800](https://pubmed.ncbi.nlm.nih.gov/11015800/)]
7. Marlatt GA, Gordon JR. Relapse Prevention: Maintenance Strategies in the Treatment of Addictive Behaviors. New York, NY. The Guilford Press; 1985.
8. Marlatt GA, Donovan DM. Relapse Prevention, Second Edition: Maintenance Strategies in the Treatment of Addictive Behaviors. New York, NY. The Guilford Press; 2007.
9. McHugh RK, Hearon BA, Otto MW. Cognitive behavioral therapy for substance use disorders. Psychiatr Clin North Am. 2010;33(3):511-525. [FREE Full text] [doi: [10.1016/j.psc.2010.04.012](https://doi.org/10.1016/j.psc.2010.04.012)] [Medline: [20599130](https://pubmed.ncbi.nlm.nih.gov/20599130/)]
10. Liese BS, Beck AT. Cognitive-Behavioral Therapy of Addictive Disorders. 1st Edition. New York, NY. The Guilford Press; 2022.
11. Bowen S, Chawla N, Grow J, Marlatt GA. Mindfulness-Based Relapse Prevention for Addictive Behaviors: A Clinician's Guide. New York, NY. The Guilford Press; 2021.

12. Yaghubi M, Zargar F. Effectiveness of mindfulness-based relapse prevention on quality of life and craving in methadone-treated patients: a randomized clinical trial. *Addict Health*. 2018;10(4):250-259. [[FREE Full text](#)] [doi: [10.22122/ahj.v10i4.573](https://doi.org/10.22122/ahj.v10i4.573)] [Medline: [31263524](#)]
13. Ramadas E, Lima MPD, Caetano T, Lopes J, Dixe MDA. Effectiveness of mindfulness-based relapse prevention in individuals with substance use disorders: a systematic review. *Behav Sci (Basel)*. 2021;11(10):133. [[FREE Full text](#)] [doi: [10.3390/bs11100133](https://doi.org/10.3390/bs11100133)] [Medline: [34677226](#)]
14. Goldberg SB, Tucker RP, Greene PA, Davidson RJ, Wampold BE, Kearney DJ, et al. Mindfulness-based interventions for psychiatric disorders: a systematic review and meta-analysis. *Clin Psychol Rev*. 2018;59:52-60. [[FREE Full text](#)] [doi: [10.1016/j.cpr.2017.10.011](https://doi.org/10.1016/j.cpr.2017.10.011)] [Medline: [29126747](#)]
15. Hendershot CS, Witkiewitz K, George WH, Marlatt GA. Relapse prevention for addictive behaviors. *Subst Abuse Treat Prev Policy*. 2011;6:17. [[FREE Full text](#)] [doi: [10.1186/1747-597X-6-17](https://doi.org/10.1186/1747-597X-6-17)] [Medline: [21771314](#)]
16. Witkiewitz K, Marlatt GA. Relapse prevention for alcohol and drug problems: that was Zen, this is Tao. *Am Psychol*. 2004;59(4):224-235. [doi: [10.1037/0003-066X.59.4.224](https://doi.org/10.1037/0003-066X.59.4.224)] [Medline: [15149263](#)]
17. Hufford MR, Witkiewitz K, Shields AL, Kodya S, Caruso JC. Relapse as a nonlinear dynamic system: application to patients with alcohol use disorders. *J Abnorm Psychol*. 2003;112(2):219-227. [doi: [10.1037/0021-843x.112.2.219](https://doi.org/10.1037/0021-843x.112.2.219)] [Medline: [12784831](#)]
18. Witkiewitz K, van der Maas HLJ, Hufford MR, Marlatt GA. Nonnormality and divergence in posttreatment alcohol use: reexamining the project MATCH data "another way.". *J Abnorm Psychol*. 2007;116(2):378-394. [[FREE Full text](#)] [doi: [10.1037/0021-843X.116.2.378](https://doi.org/10.1037/0021-843X.116.2.378)] [Medline: [17516769](#)]
19. Witkiewitz K, Marlatt GA. Modeling the complexity of post-treatment drinking: it's a rocky road to relapse. *Clin Psychol Rev*. 2007;27(6):724-738. [[FREE Full text](#)] [doi: [10.1016/j.cpr.2007.01.002](https://doi.org/10.1016/j.cpr.2007.01.002)] [Medline: [17355897](#)]
20. Brandon TH, Vidrine JI, Litvin EB. Relapse and relapse prevention. *Annu Rev Clin Psychol*. 2007;3:257-284. [doi: [10.1146/annurev.clinpsy.3.022806.091455](https://doi.org/10.1146/annurev.clinpsy.3.022806.091455)] [Medline: [17716056](#)]
21. Witkiewitz KA, Marlatt GA. *Therapist's Guide to Evidence-Based Relapse Prevention*. 1st Edition. Amsterdam. Academic Press; 2007.
22. Epstein DH, Tyburski M, Kowalczyk WJ, Burgess-Hull AJ, Phillips KA, Curtis BL, et al. Prediction of stress and drug craving ninety minutes in the future with passively collected GPS data. *NPJ Digit Med*. 2020;3:26. [[FREE Full text](#)] [doi: [10.1038/s41746-020-0234-6](https://doi.org/10.1038/s41746-020-0234-6)] [Medline: [32195362](#)]
23. Suchting R, Hébert ET, Ma P, Kendzor DE, Businelle MS. Using elastic net penalized cox proportional hazards regression to identify predictors of imminent smoking lapse. *Nicotine Tob Res*. 2019;21(2):173-179. [[FREE Full text](#)] [doi: [10.1093/ntr/ntx201](https://doi.org/10.1093/ntr/ntx201)] [Medline: [29059349](#)]
24. Hébert ET, Suchting R, Ra CK, Alexander AC, Kendzor DE, Vidrine DJ, et al. Predicting the first smoking lapse during a quit attempt: a machine learning approach. *Drug Alcohol Depend*. 2021;218:108340. [[FREE Full text](#)] [doi: [10.1016/j.drugalcdep.2020.108340](https://doi.org/10.1016/j.drugalcdep.2020.108340)] [Medline: [33092911](#)]
25. Engelhard M, Xu H, Carin L, Oliver JA, Hallyburton M, McClernon FJ. Predicting smoking events with a time-varying semi-parametric Hawkes process model. *Proc Mach Learn Res*. 2018;85:312-331. [[FREE Full text](#)] [Medline: [30899917](#)]
26. Mohr DC, Zhang M, Schueller SM. Personal sensing: understanding mental health using ubiquitous sensors and machine learning. *Annu Rev Clin Psychol*. 2017;13:23-47. [[FREE Full text](#)] [doi: [10.1146/annurev-clinpsy-032816-044949](https://doi.org/10.1146/annurev-clinpsy-032816-044949)] [Medline: [28375728](#)]
27. Businelle MS, Ma P, Kendzor DE, Frank SG, Wetter DW, Vidrine DJ. Using intensive longitudinal data collected via mobile phone to detect imminent lapse in smokers undergoing a scheduled quit attempt. *J Med Internet Res*. 2016;18(10):e275. [[FREE Full text](#)] [doi: [10.2196/jmir.6307](https://doi.org/10.2196/jmir.6307)] [Medline: [27751985](#)]
28. Soyster PD, Ashlock L, Fisher AJ. Pooled and person-specific machine learning models for predicting future alcohol consumption, craving, and wanting to drink: a demonstration of parallel utility. *Psychol Addict Behav*. 2022;36(3):296-306. [doi: [10.1037/adb0000666](https://doi.org/10.1037/adb0000666)] [Medline: [35041441](#)]
29. Hébert ET, Stevens EM, Frank SG, Kendzor DE, Wetter DW, Zvolensky MJ, et al. An ecological momentary intervention for smoking cessation: the associations of just-in-time, tailored messages with lapse risk factors. *Addict Behav*. 2018;78:30-35. [[FREE Full text](#)] [doi: [10.1016/j.addbeh.2017.10.026](https://doi.org/10.1016/j.addbeh.2017.10.026)] [Medline: [29121530](#)]
30. Moshontz H, Colmenares AJ, Fronk GE, Sant'Ana SJ, Wyant K, Wanta SE, et al. Prospective prediction of lapses in opioid use disorder: protocol for a personal sensing study. *JMIR Res Protoc*. 2021;10(12):e29563. [[FREE Full text](#)] [doi: [10.2196/29563](https://doi.org/10.2196/29563)] [Medline: [34559061](#)]
31. Wyant K, Moshontz H, Ward SB, Fronk GE, Curtin JJ. Acceptability of personal sensing among people with alcohol use disorder: observational study. *JMIR Mhealth Uhealth*. 2023;11:e41833. [[FREE Full text](#)] [doi: [10.2196/41833](https://doi.org/10.2196/41833)] [Medline: [37639300](#)]
32. Chih MY, Patton T, McTavish FM, Isham AJ, Judkins-Fisher CL, Atwood AK, et al. Predictive modeling of addiction lapses in a mobile health application. *J Subst Abuse Treat*. 2014;46(1):29-35. [[FREE Full text](#)] [doi: [10.1016/j.jsat.2013.08.004](https://doi.org/10.1016/j.jsat.2013.08.004)] [Medline: [24035143](#)]

33. Bae S, Chung T, Ferreira D, Dey AK, Suffoletto B. Mobile phone sensors and supervised machine learning to identify alcohol use events in young adults: implications for just-in-time adaptive interventions. *Addict Behav.* 2018;83:42-47. [FREE Full text] [doi: [10.1016/j.addbeh.2017.11.039](https://doi.org/10.1016/j.addbeh.2017.11.039)] [Medline: [29217132](https://pubmed.ncbi.nlm.nih.gov/29217132/)]
34. Walters ST, Businelle MS, Suchting R, Li X, Hébert ET, Mun EY. Using machine learning to identify predictors of imminent drinking and create tailored messages for at-risk drinkers experiencing homelessness. *J Subst Abuse Treat.* 2021;127:108417. [FREE Full text] [doi: [10.1016/j.jsat.2021.108417](https://doi.org/10.1016/j.jsat.2021.108417)] [Medline: [34134874](https://pubmed.ncbi.nlm.nih.gov/34134874/)]
35. Wyant K, Sant'Ana SJK, Fronk GE, Curtin JJ. Machine learning models for temporally precise lapse prediction in alcohol use disorder. *J Psychopathol Clin Sci.* 2024;133(7):527-540. [doi: [10.1037/abn0000901](https://doi.org/10.1037/abn0000901)] [Medline: [39172368](https://pubmed.ncbi.nlm.nih.gov/39172368/)]
36. Hosmer Jr DW, Lemeshow S. *Applied Logistic Regression*. 1st Edition. New York, NY: Wiley-Interscience; 1989.
37. Molnar C. *Interpretable Machine Learning: A Guide For Making Black Box Models Explainable*. Munich. Independently Published; 2022.
38. Fortunati L, Edwards A. Opening space for theoretical, methodological, and empirical issues in human-machine communication. *Hum-Mach Commun.* 2020;1:7-18. [doi: [10.30658/hmc.1.1](https://doi.org/10.30658/hmc.1.1)]
39. Banks J, de Graaf M. Toward an agent-agnostic transmission model: synthesizing anthropocentric and technocentric paradigms in communication. *Hum-Mach Commun.* 2020;1:19-36. [doi: [10.30658/hmc.1.2](https://doi.org/10.30658/hmc.1.2)]
40. Hoff KA, Bashir M. Trust in automation: integrating empirical evidence on factors that influence trust. *Hum Factors.* 2015;57(3):407-434. [doi: [10.1177/0018720814547570](https://doi.org/10.1177/0018720814547570)] [Medline: [25875432](https://pubmed.ncbi.nlm.nih.gov/25875432/)]
41. Lutz C, Tamó-Larrieux A. The robot privacy paradox: understanding how privacy concerns shape intentions to use social robots. *Hum-Mach Commun.* 2020;1:87-111. [doi: [10.30658/hmc.1.6](https://doi.org/10.30658/hmc.1.6)]
42. Ibrahim MA, Assaad Z, Williams E. Trust and communication in human-machine teaming. *Front Phys.* 2022;10. [doi: [10.3389/fphy.2022.942896](https://doi.org/10.3389/fphy.2022.942896)]
43. Guzman A. Ontological boundaries between humans and computers and the implications for human-machine communication. *Hum-Mach Commun.* 2020;1:37-54. [doi: [10.30658/hmc.1.3](https://doi.org/10.30658/hmc.1.3)]
44. Prael A, Van Swol L. Out with the humans, in with the machines?: Investigating the behavioral and psychological effects of replacing human advisors with a machine. *Hum-Mach Commun.* 2021;2:209-234. [doi: [10.30658/hmc.2.11](https://doi.org/10.30658/hmc.2.11)]
45. Kizilcec RF. How much information?: Effects of transparency on trust in an algorithmic interface. 2016. Presented at: CHI-16: CHI Conference on Human Factors in Computing Systems; May 7-12, 2016:2390-2395; San Jose, CA, United States. [doi: [10.1145/2858036.2858402](https://doi.org/10.1145/2858036.2858402)]
46. Kulesza T, Burnett M, Wong WK, Stumpf S. Principles of explanatory debugging to personalize interactive machine learning. 2015. Presented at: IUI-15 20th International Conference on Intelligent User Interfaces; March 29-April 1, 2015:A-137; Atlanta, GA, United States. [doi: [10.1145/2678025.2701399](https://doi.org/10.1145/2678025.2701399)]
47. Lim BY, Dey AK, Avrahami D. Why and why not explanations improve the intelligibility of context-aware intelligent systems. 2009. Presented at: CHI-09: CHI Conference on Human Factors in Computing Systems; April 4-9, 2009:2119-2128; Boston, MA, United States. [doi: [10.1145/1518701.1519023](https://doi.org/10.1145/1518701.1519023)]
48. Burke M, Amento B, Isenhour P. Error correction of voicemail transcripts in SCANMail. 2006. Presented at: CHI06: CHI 2006 Conference on Human Factors in Computing Systems; April 22-27, 2006:339-348; Montréal, Québec, Canada. [doi: [10.1145/1124772.1124823](https://doi.org/10.1145/1124772.1124823)]
49. Yang R, Newman MW. Learning from a learning thermostat: lessons for intelligent systems for the home. 2013. Presented at: UbiComp-13: The 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 8-12, 2013:93-102; Zurich, Switzerland. [doi: [10.1145/2493432.2493489](https://doi.org/10.1145/2493432.2493489)]
50. Lim B, Dey A. Assessing demand for intelligibility in context-aware applications. 2009. Presented at: Ubicomp '09: The 11th International Conference on Ubiquitous Computing; September 30-October 3, 2009:195-204; Orlando, FL, United States. [doi: [10.1145/1620545.1620576](https://doi.org/10.1145/1620545.1620576)]
51. Lim BY, Dey AK. Investigating intelligibility for uncertain context-aware applications. 2011. Presented at: Ubicomp '11: The 2011 ACM Conference on Ubiquitous Computing; September 17-21, 2011:415-424; Beijing, China. [doi: [10.1145/2030112.2030168](https://doi.org/10.1145/2030112.2030168)]
52. Springer A, Whittaker S. Progressive disclosure: when, why, and how do users want algorithmic transparency information? *ACM Trans Interact Intell Syst.* 2020;10(4):1-32. [doi: [10.1145/3374218](https://doi.org/10.1145/3374218)]
53. Kulesza T, Wong WK, Stumpf S, Perona S, White R, Burnett M, et al. Fixing the program my computer learned: barriers for end users, challenges for the machine. 2009. Presented at: IUI09: 14th International Conference on Intelligent User Interfaces Sanibel; February 8-11, 2009:187-196; Island, FL, United States. [doi: [10.1145/1502650.1502678](https://doi.org/10.1145/1502650.1502678)]
54. Ribeiro MT, Singh S, Guestrin C. "Why should i trust you": explaining the predictions of any classifier. 2016. Presented at: KDD-16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 13-17, 2016:1135-1144; San Francisco, CA, United States. [doi: [10.1145/2939672.2939778](https://doi.org/10.1145/2939672.2939778)]
55. Abdul A, Vermeulen J, Wang D, Lim B, Kankanhalli M. Trends and trajectories for explainable, accountable and intelligible systems: an HCI research agenda. 2018. Presented at: CHI-18: CHI Conference on Human Factors in Computing Systems; April 21-26, 2018:1-18; Montreal, Québec, Canada. [doi: [10.1145/3173574.3174156](https://doi.org/10.1145/3173574.3174156)]

56. Bunt A, Lount M, Lauzon C. Are explanations always important? A study of deployed, low-cost intelligent interactive systems. 2012. Presented at: Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces; February 14-17, 2012:169-178; New York, NY, United States. [doi: [10.1145/2166966.2166996](https://doi.org/10.1145/2166966.2166996)]
57. Wickens CD, Helton WS, Hollands JG, Banbury S. Engineering Psychology and Human Performance. 5th Edition. New York, NY. Routledge; 2022.
58. Reeves B, Nass CI. The Media Equation: How People treat Computers, Television, and New Media Like Real People and Places. New York, NY. Cambridge University Press; 1996:305.
59. Nass C, Moon Y. Machines and mindlessness: social responses to computers. *J Soc Issues*. 2002;56(1):81-103. [doi: [10.1111/0022-4537.00153](https://doi.org/10.1111/0022-4537.00153)]
60. de Visser EJ, Monfort SS, McKendrick R, Smith MAB, McKnight PE, Krueger F, et al. Almost human: anthropomorphism increases trust resilience in cognitive agents. *J Exp Psychol Appl*. 2016;22(3):331-349. [doi: [10.1037/xap0000092](https://doi.org/10.1037/xap0000092)] [Medline: [27505048](https://pubmed.ncbi.nlm.nih.gov/27505048/)]
61. Liebrecht C, Sander L, van Hooijdonk C. Too informal? How a chatbot's communication style affects brand attitude and quality of interaction. 2020. Presented at: Chatbot Research and Design: 4th International Workshop, CONVERSATIONS 2020; November 23-24, 2020:16-31; Virtual Event. [doi: [10.1007/978-3-030-68288-0_2](https://doi.org/10.1007/978-3-030-68288-0_2)]
62. Araujo T. Living up to the chatbot hype: the influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Comput Hum Behav*. 2018;85:183-189. [doi: [10.1016/j.chb.2018.03.051](https://doi.org/10.1016/j.chb.2018.03.051)]
63. Burleson BR, Samter W. Consistencies in theoretical and naive evaluations of comforting messages. *Commun Monogr*. 2009;52(2):103-123. [doi: [10.1080/03637758509376099](https://doi.org/10.1080/03637758509376099)]
64. MacGeorge EL, Guntzville LM, Hanasono LK, Feng B. Testing advice response theory in interactions with friends. *Commun Res*. 2013;43(2):211-231. [doi: [10.1177/0093650213510938](https://doi.org/10.1177/0093650213510938)]
65. MacGeorge EL, Foley KA, Figgins EPC, Vanderbilt RR, Worthington AK, Hackman NM. "Watchful Waiting" advice for pediatric ear infections. *J Lang Soc Psychol*. 2020;39(3):414-424. [doi: [10.1177/0261927x20912292](https://doi.org/10.1177/0261927x20912292)]
66. Montoya RM, Horton RS. A meta-analytic investigation of the processes underlying the similarity-attraction effect. *J Soc Pers Relat*. 2012;30(1):64-94. [doi: [10.1177/0265407512452989](https://doi.org/10.1177/0265407512452989)]
67. Kim Y, Baek TH, Yoon S, Oh S, Choi YK. Assertive environmental advertising and reactance: differences between South Koreans and Americans. *J Advert*. 2017;46(4):550-564. [doi: [10.1080/00913367.2017.1361878](https://doi.org/10.1080/00913367.2017.1361878)]
68. Eyssel F, Kuchenbrandt D, Bobinger S, de Ruiter L, Hegel F. If you sound like me, you must be more human?: On the interplay of robot and user features on human-robot acceptance and anthropomorphism. 2012. Presented at: Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction; March 5-8, 2012:125-126; New York, NY, United States. [doi: [10.1145/2157689.2157717](https://doi.org/10.1145/2157689.2157717)]
69. Lee EJ, Nass C, Brave S. Can computer-generated speech have gender? An experimental test of gender stereotype. 2000. Presented at: CHI '00 Extended Abstracts on Human Factors in Computing Systems; April 1-6, 2000:289-290; New York, NY, United States. [doi: [10.1145/633292.633461](https://doi.org/10.1145/633292.633461)]
70. Edwards C, Edwards A, Stoll B, Lin X, Massey N. Evaluations of an artificial intelligence instructor's voice: social identity theory in human-robot interactions. *Comput Hum Behav*. 2019;90:357-362. [doi: [10.1016/j.chb.2018.08.027](https://doi.org/10.1016/j.chb.2018.08.027)]
71. Moon Y, Nass C. How "real" are computer personalities? Psychological responses to personality types in human-computer interaction. *Commun Res*. 1996;23(6):651-674. [doi: [10.1177/009365096023006002](https://doi.org/10.1177/009365096023006002)]
72. Nass C, Lee KM. Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *J Exp Psychol*. 2001;7(3):171-181. [doi: [10.1037/1076-898x.7.3.171](https://doi.org/10.1037/1076-898x.7.3.171)]
73. Torous J, Roberts LW. Needed innovation in digital health and smartphone applications for mental health: transparency and trust. *JAMA Psychiatry*. 2017;74(5):437-438. [doi: [10.1001/jamapsychiatry.2017.0262](https://doi.org/10.1001/jamapsychiatry.2017.0262)] [Medline: [28384700](https://pubmed.ncbi.nlm.nih.gov/28384700/)]
74. Murnane EL, Huffaker D, Kossinets G. Mobile health apps: adoption, adherence, and abandonment. 2015. Presented at: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers—UbiComp '15; September 7-11, 2015:261-264; Osaka, Japan. [doi: [10.1145/2800835.2800943](https://doi.org/10.1145/2800835.2800943)]
75. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. *Evid Based Ment Health*. 2018;21(3):116-119. [FREE Full text] [doi: [10.1136/eb-2018-102891](https://doi.org/10.1136/eb-2018-102891)] [Medline: [29871870](https://pubmed.ncbi.nlm.nih.gov/29871870/)]
76. Collins LM, Murphy SA, Strecher V. The Multiphase Optimization Strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. *Am J Prev Med*. 2007;32(5 Suppl):S112-S118. [FREE Full text] [doi: [10.1016/j.amepre.2007.01.022](https://doi.org/10.1016/j.amepre.2007.01.022)] [Medline: [17466815](https://pubmed.ncbi.nlm.nih.gov/17466815/)]
77. Collins LM. Optimization of Behavioral, Biobehavioral, and Biomedical Interventions: The Multiphase Optimization Strategy (MOST). Cham. Springer International Publishing; 2018.
78. Collins LM, Kugler KC. Optimization of Behavioral, Biobehavioral, and Biomedical Interventions: Advanced Topics. Cham. Springer International Publishing; 2018.
79. Hagman BT, Falk D, Litten R, Koob GF. Defining recovery from alcohol use disorder: development of an NIAAA research definition. *Am J Psychiatry*. 2022;179(11):807-813. [doi: [10.1176/appi.ajp.21090963](https://doi.org/10.1176/appi.ajp.21090963)] [Medline: [35410494](https://pubmed.ncbi.nlm.nih.gov/35410494/)]

80. Wyant K, Sant'Ana S, Punturieri C, Yu J, Fronk G, Maggard C, et al. Maximizing engagement, trust, and clinical benefit of AI-generated recovery support messages for alcohol use disorder: protocol for an optimization study. Open Science Framework Repository. 2025. URL: <https://osf.io/8ascm/overview> [accessed 2025-10-27]
81. Fisher AJ, Bosley HG, Fernandez KC, Reeves JW, Soyster PD, Diamond AE, et al. Open trial of a personalized modular treatment for mood and anxiety. *Behav Res Ther*. 2019;116:69-79. [doi: [10.1016/j.brat.2019.01.010](https://doi.org/10.1016/j.brat.2019.01.010)] [Medline: [30831478](https://pubmed.ncbi.nlm.nih.gov/30831478/)]
82. Fernandez KC, Fisher AJ, Chi C. Development and initial implementation of the Dynamic Assessment Treatment Algorithm (DATA). *PLoS One*. 2017;12(6):e0178806. [FREE Full text] [doi: [10.1371/journal.pone.0178806](https://doi.org/10.1371/journal.pone.0178806)] [Medline: [28654667](https://pubmed.ncbi.nlm.nih.gov/28654667/)]
83. Shah RV, Grennan G, Zafar-Khan M, Alim F, Dey S, Ramanathan D, et al. Personalized machine learning of depressed mood using wearables. *Transl Psychiatry*. 2021;11(1):338. [FREE Full text] [doi: [10.1038/s41398-021-01445-0](https://doi.org/10.1038/s41398-021-01445-0)] [Medline: [34103481](https://pubmed.ncbi.nlm.nih.gov/34103481/)]
84. Cohen ZD, DeRubeis RJ. Treatment selection in depression. *Annu Rev Clin Psychol*. 2018;14:209-236. [doi: [10.1146/annurev-clinpsy-050817-084746](https://doi.org/10.1146/annurev-clinpsy-050817-084746)] [Medline: [29494258](https://pubmed.ncbi.nlm.nih.gov/29494258/)]
85. DeRubeis RJ, Cohen ZD, Forand NR, Fournier JC, Gelfand LA, Lorenzo-Luaces L. The Personalized Advantage Index: translating research on prediction into individualized treatment recommendations. A demonstration. *PLoS One*. 2014;9(1):e83875. [FREE Full text] [doi: [10.1371/journal.pone.0083875](https://doi.org/10.1371/journal.pone.0083875)] [Medline: [24416178](https://pubmed.ncbi.nlm.nih.gov/24416178/)]
86. Hall-Flavin DK, Winner JG, Allen JD, Jordan JJ, Nesheim RS, Snyder KA, et al. Using a pharmacogenomic algorithm to guide the treatment of depression. *Transl Psychiatry*. 2012;2(10):e172-e172. [FREE Full text] [doi: [10.1038/tp.2012.99](https://doi.org/10.1038/tp.2012.99)] [Medline: [23047243](https://pubmed.ncbi.nlm.nih.gov/23047243/)]
87. Zikmund-Fisher BJ. The right tool is what they need, not what we have: a taxonomy of appropriate levels of precision in patient risk communication. *Med Care Res Rev*. 2013;70(1 Suppl):37S. [doi: [10.1177/1077558712458541](https://doi.org/10.1177/1077558712458541)] [Medline: [22955699](https://pubmed.ncbi.nlm.nih.gov/22955699/)]
88. Sanneman L, Shah JA. A situation awareness-based framework for design and evaluation of explainable AI. 2020. Presented at: Explainable, Transparent Autonomous Agents and Multi-Agent Systems. Second International Workshop, EXTRAAMAS 2020; May 9-13, 2020:94-110; Auckland, New Zealand. [doi: [10.1007/978-3-030-51924-7_6](https://doi.org/10.1007/978-3-030-51924-7_6)]
89. Fagerlin A, Ubel PA, Smith DM, Zikmund-Fisher BJ. Making numbers matter: present and future research in risk communication. *Am J Health Behav*. 2007;31 Suppl 1:S47-S56. [doi: [10.5555/ajhb.2007.31.suppl.S47](https://doi.org/10.5555/ajhb.2007.31.suppl.S47)] [Medline: [17931136](https://pubmed.ncbi.nlm.nih.gov/17931136/)]
90. Zipkin DA, Umscheid CA, Keating NL, Allen E, Aung K, Beyth R, et al. Evidence-based risk communication: a systematic review. *Ann Intern Med*. 2014;161(4):270-280. [doi: [10.7326/M14-0295](https://doi.org/10.7326/M14-0295)] [Medline: [25133362](https://pubmed.ncbi.nlm.nih.gov/25133362/)]
91. Garcia-Retamero R, Cokely ET. Communicating health risks with visual aids. *Curr Dir Psychol Sci*. 2013;22(5):392-399. [doi: [10.1177/0963721413491570](https://doi.org/10.1177/0963721413491570)]
92. Gajecki M, Berman AH, Sinadinovic K, Rosendahl I, Andersson C. Mobile phone brief intervention applications for risky alcohol use among university students: a randomized controlled study. *Addict Sci Clin Pract*. 2014;9(1):11. [FREE Full text] [doi: [10.1186/1940-0640-9-11](https://doi.org/10.1186/1940-0640-9-11)] [Medline: [24985342](https://pubmed.ncbi.nlm.nih.gov/24985342/)]
93. Eini-Porat B, Amir O, Eytan D, Shalit U. Tell me something interesting: clinical utility of machine learning prediction models in the ICU. *J Biomed Inform*. 2022;132:104107. [FREE Full text] [doi: [10.1016/j.jbi.2022.104107](https://doi.org/10.1016/j.jbi.2022.104107)] [Medline: [35688332](https://pubmed.ncbi.nlm.nih.gov/35688332/)]
94. Huckvale K, Venkatesh S, Christensen H. Toward clinical digital phenotyping: a timely opportunity to consider purpose, quality, and safety. *NPJ Digit Med*. 2019;2(1):88. [FREE Full text] [doi: [10.1038/s41746-019-0166-1](https://doi.org/10.1038/s41746-019-0166-1)] [Medline: [31508498](https://pubmed.ncbi.nlm.nih.gov/31508498/)]
95. Lauritsen SM, Kristensen M, Olsen MV, Larsen MS, Lauritsen KM, Jørgensen MJ, et al. Explainable artificial intelligence model to predict acute critical illness from electronic health records. *Nat Commun*. 2020;11(1):3852. [FREE Full text] [doi: [10.1038/s41467-020-17431-x](https://doi.org/10.1038/s41467-020-17431-x)] [Medline: [32737308](https://pubmed.ncbi.nlm.nih.gov/32737308/)]
96. da Cruz HF, Pfahringer B, Martensen T, Schneider F, Meyer A, Böttinger E, et al. Using interpretability approaches to update "black-box" clinical prediction models: an external validation study in nephrology. *Artif Intell Med*. 2021;111:101982. [doi: [10.1016/j.artmed.2020.101982](https://doi.org/10.1016/j.artmed.2020.101982)] [Medline: [33461682](https://pubmed.ncbi.nlm.nih.gov/33461682/)]
97. Shin D. The effects of explainability and causability on perception, trust, and acceptance: implications for explainable AI. *Int J Hum-Comput Studies*. 2021;146:102551. [doi: [10.1016/j.ijhcs.2020.102551](https://doi.org/10.1016/j.ijhcs.2020.102551)]
98. Center for Substance Abuse Treatment. Counselor's Treatment Manual: Matrix Intensive Outpatient Treatment for People With Stimulant Use Disorders. HHS Publication No. (SMA) 13-4152. Rockville, MD. Substance Abuse and Mental Health Services Administration; 2006.
99. DeMarce JM, Gnys M, Raffa SD, Karlin BE. Cognitive Behavioral Therapy for Substance Use Disorders Among Veterans: Therapist Manual (Rev ed.). Washington, DC. US Department of Veterans Affairs; 2023.
100. Miller WR, Zweben A, DiClemente CC, Rychtarik R, Mattson M, editor. Motivational Enhancement Therapy Manual: A Clinical Research Guide for Therapists Treating Individuals With Alcohol Abuse and Dependence. Rockville, MD. National Institute on Alcohol Abuse and Alcoholism; 1999.
101. Maloney EK, Lapinski MK, Witte K. Fear appeals and persuasion: a review and update of the extended parallel process model. *Soc Person Psychol*. 2011;5(4):206-219. [doi: [10.1111/j.1751-9004.2011.00341.x](https://doi.org/10.1111/j.1751-9004.2011.00341.x)]
102. Mitchell EG, Heitkemper EM, Burgermaster M, Levine ME, Miao Y, Hwang ML, et al. From reflection to action: combining machine learning with expert knowledge for nutrition goal recommendations. 2021. Presented at: Proceedings of the 2021

- CHI Conference on Human Factors in Computing Systems; May 8-13, 2021:1-17; New York, NY, United States. [doi: [10.1145/3411764.3445555](https://doi.org/10.1145/3411764.3445555)]
103. Goldstein SP, Evans BC, Flack D, Juarascio A, Manasse S, Zhang F, et al. Return of the JITAI: applying a just-in-time adaptive intervention framework to the development of m-Health solutions for addictive behaviors. *Int J Behav Med*. 2017;24(5):673-682. [FREE Full text] [doi: [10.1007/s12529-016-9627-y](https://doi.org/10.1007/s12529-016-9627-y)] [Medline: [28083725](https://pubmed.ncbi.nlm.nih.gov/28083725/)]
 104. Prah A, Van Swol L. Understanding algorithm aversion: when is advice from automation discounted? *J Forecast*. 2017;36(6):691-702. [doi: [10.1002/for.2464](https://doi.org/10.1002/for.2464)]
 105. Dexter F, Van Swol LM. Influence of data and formulas on trust in information from journal articles in an operating room management course. *A & A Case Rep*. 2016;6(11):329-334. [doi: [10.1213/xa.0000000000000298](https://doi.org/10.1213/xa.0000000000000298)]
 106. Dexter F, Epstein RH, Fahy BG, Van Swol LM. With directed study before a 4-day operating room management course, trust in the content did not change progressively during the classroom time. *J Clin Anesth*. 2017;42:57-62. [doi: [10.1016/j.jclinane.2017.08.003](https://doi.org/10.1016/j.jclinane.2017.08.003)] [Medline: [28822884](https://pubmed.ncbi.nlm.nih.gov/28822884/)]
 107. Ahn PH, Dexter F, Fahy BG, Van Swol LM. Demonstrability of analytics solutions and shared knowledge of statistics and operating room management improves expected performance of small teams in correctly solving problems and making good decisions. *Perioper Care Oper Room Manag*. 2020;19:100090. [doi: [10.1016/j.pcorn.2020.100090](https://doi.org/10.1016/j.pcorn.2020.100090)]
 108. Sniezek JA, Van Swol LM. Trust, confidence, and expertise in a judge-advisor system. *Organ Behav Hum Decis Process*. 2001;84(2):288-307. [doi: [10.1006/obhd.2000.2926](https://doi.org/10.1006/obhd.2000.2926)] [Medline: [11277673](https://pubmed.ncbi.nlm.nih.gov/11277673/)]
 109. van Swol LM, Sniezek JA. Factors affecting the acceptance of expert advice. *Br J Soc Psychol*. 2011;44(3):443-461. [doi: [10.1348/014466604x17092](https://doi.org/10.1348/014466604x17092)]
 110. Van SL, Paik JE, Prah A. *Advice Recipients: The Psychology of Advice Utilization*. New York, NY. Oxford University Press; 2018:21-41.
 111. MacGeorge EL, Swol LMV. *The Oxford Handbook of Advice*. New York, NY. Oxford University Press; 2018.
 112. WHO ASSIST Working Group. The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): development, reliability and feasibility. *Addiction*. 2002;97(9):1183-1194. [FREE Full text] [doi: [10.1046/j.1360-0443.2002.00185.x](https://doi.org/10.1046/j.1360-0443.2002.00185.x)] [Medline: [12199834](https://pubmed.ncbi.nlm.nih.gov/12199834/)]
 113. Merritt SM, Heimbaugh H, LaChapell J, Lee D. I trust it, but I don't know why: effects of implicit attitudes toward automation on trust in an automated system. *Hum Factors*. 2013;55(3):520-534. [doi: [10.1177/0018720812465081](https://doi.org/10.1177/0018720812465081)] [Medline: [23829027](https://pubmed.ncbi.nlm.nih.gov/23829027/)]
 114. Körber M. Theoretical considerations and development of a questionnaire to measure trust in automation. 2019. Presented at: Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018); August 26-30, 2018:13-30; Florence, Italy. [doi: [10.1007/978-3-319-96074-6_2](https://doi.org/10.1007/978-3-319-96074-6_2)]
 115. Henson P, Wisniewski H, Hollis C, Keshavan M, Torous J. Digital mental health apps and the therapeutic alliance: initial review. *BJPsych Open*. 2019;5(1):e15. [FREE Full text] [doi: [10.1192/bjo.2018.86](https://doi.org/10.1192/bjo.2018.86)] [Medline: [30762511](https://pubmed.ncbi.nlm.nih.gov/30762511/)]
 116. Jensen JD, King AJ, Carcioppolo N, Davis L. Why are tailored messages more effective? A multiple mediation analysis of a breast cancer screening intervention. *J Commun*. 2012;62(5):851-868. [FREE Full text] [doi: [10.1111/j.1460-2466.2012.01668.x](https://doi.org/10.1111/j.1460-2466.2012.01668.x)] [Medline: [26405350](https://pubmed.ncbi.nlm.nih.gov/26405350/)]
 117. Tielman ML, Neerincx MA, Brinkman WP. Design and evaluation of personalized motivational messages by a virtual agent that assists in post-traumatic stress disorder therapy. *J Med Internet Res*. 2019;21(3):e9240. [FREE Full text] [doi: [10.2196/jmir.9240](https://doi.org/10.2196/jmir.9240)] [Medline: [30916660](https://pubmed.ncbi.nlm.nih.gov/30916660/)]
 118. Kocielnik R, Hsieh G. Send me a different message: utilizing cognitive space to create engaging message triggers. 2017. Presented at: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing; February 25-March 1, 2017:2193-2207; New York, NY, United States. [doi: [10.1145/2998181.2998324](https://doi.org/10.1145/2998181.2998324)]
 119. Wojton HM, Porter D, T Lane S, Bieber C, Madhavan P. Initial validation of the trust of automated systems test (TOAST). *J Soc Psychol*. 2020;160(6):735-750. [doi: [10.1080/00224545.2020.1749020](https://doi.org/10.1080/00224545.2020.1749020)] [Medline: [32297844](https://pubmed.ncbi.nlm.nih.gov/32297844/)]
 120. Hatcher RL, Gillaspay JA. Development and validation of a revised short version of the Working Alliance Inventory. *Psychother Res*. 2006;16(1):12-25. [doi: [10.1080/10503300500352500](https://doi.org/10.1080/10503300500352500)]
 121. Alcoholism: developing drugs for treatment (No FDA D-0152-001). Food and Drug Administration. 2015. URL: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/alcoholism-developing-drugs-treatment> [accessed 2025-10-31]
 122. Sobell LC, Sobell MB. Timeline followback: a technique for assessing self-reported alcohol consumption. In: *Measuring alcohol Consumption: Psychosocial and Biological Methods*. Totowa, NJ. Humana Press; 1992:41-72.
 123. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med*. 2006;166(10):1092-1097. [doi: [10.1001/archinte.166.10.1092](https://doi.org/10.1001/archinte.166.10.1092)] [Medline: [16717171](https://pubmed.ncbi.nlm.nih.gov/16717171/)]
 124. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606-613. [FREE Full text] [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
 125. VanderWeele TJ. On the promotion of human flourishing. *Proc Natl Acad Sci USA*. 2017;114(31):8148-8156. [FREE Full text] [doi: [10.1073/pnas.1702996114](https://doi.org/10.1073/pnas.1702996114)] [Medline: [28705870](https://pubmed.ncbi.nlm.nih.gov/28705870/)]
 126. Bowen E, Irish A, Wilding G, LaBarre C, Capozziello N, Nochajski T, et al. Development and psychometric properties of the Multidimensional Inventory of Recovery Capital (MIRC). *Drug Alcohol Depend*. 2023;247:109875. [doi: [10.1016/j.drugalcdep.2023.109875](https://doi.org/10.1016/j.drugalcdep.2023.109875)] [Medline: [37119593](https://pubmed.ncbi.nlm.nih.gov/37119593/)]

127. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the tidyverse. *J Open Source Softw.* 2019;4(43):1686. [doi: [10.21105/joss.01686](https://doi.org/10.21105/joss.01686)]
128. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Soft.* 2015;67(1):1-48. [doi: [10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01)]
129. van Buuren S, Groothuis-Oudshoorn K. Mice: Multivariate imputation by chained equations in R. *J Stat Soft.* 2011;45(3):1-67. [doi: [10.18637/jss.v045.i03](https://doi.org/10.18637/jss.v045.i03)]
130. Collins LM. Introduction to the factorial optimization trial. In: *Optimization of Behavioral, Biobehavioral, and Biomedical Interventions: The Multiphase Optimization Strategy (MOST)*. Cham. Springer International Publishing; 2018:67-113.
131. Koch GG, Tangen CM, Jung JW, Amara IA. Issues for covariance analysis of dichotomous and ordered categorical data from randomized clinical trials and non-parametric strategies for addressing them. *Stat Med.* 1998;17(15-16):1863-1892. [doi: [10.1002/\(sici\)1097-0258\(19980815/30\)17:15/16<1863::aid-sim989>3.0.co;2-m](https://doi.org/10.1002/(sici)1097-0258(19980815/30)17:15/16<1863::aid-sim989>3.0.co;2-m)] [Medline: [9749453](https://pubmed.ncbi.nlm.nih.gov/9749453/)]
132. Moore KL, van der Laan MJ. Covariate adjustment in randomized trials with binary outcomes: targeted maximum likelihood estimation. *Stat Med.* 2009;28(1):39-64. [FREE Full text] [doi: [10.1002/sim.3445](https://doi.org/10.1002/sim.3445)] [Medline: [18985634](https://pubmed.ncbi.nlm.nih.gov/18985634/)]
133. Pocock SJ, Assmann SE, Enos LE, Kasten LE. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. *Stat Med.* 2002;21(19):2917-2930. [doi: [10.1002/sim.1296](https://doi.org/10.1002/sim.1296)] [Medline: [12325108](https://pubmed.ncbi.nlm.nih.gov/12325108/)]
134. Tsiatis AA, Davidian M, Zhang M, Lu X. Covariate adjustment for two-sample treatment comparisons in randomized clinical trials: a principled yet flexible approach. *Stat Med.* 2008;27(23):4658-4677. [FREE Full text] [doi: [10.1002/sim.3113](https://doi.org/10.1002/sim.3113)] [Medline: [17960577](https://pubmed.ncbi.nlm.nih.gov/17960577/)]
135. Zhang M, Tsiatis AA, Davidian M. Improving efficiency of inferences in randomized clinical trials using auxiliary covariates. *Biometrics.* 2008;64(3):707-715. [FREE Full text] [doi: [10.1111/j.1541-0420.2007.00976.x](https://doi.org/10.1111/j.1541-0420.2007.00976.x)] [Medline: [18190618](https://pubmed.ncbi.nlm.nih.gov/18190618/)]
136. Kaye JT, Fronk GE, Zgierska AE, Cruz MR, Rabago D, Curtin JJ. Acute prazosin administration does not reduce stressor reactivity in healthy adults. *Psychopharmacology (Berl).* 2019;236(11):3371-3382. [FREE Full text] [doi: [10.1007/s00213-019-05297-x](https://doi.org/10.1007/s00213-019-05297-x)] [Medline: [31197436](https://pubmed.ncbi.nlm.nih.gov/31197436/)]
137. Moberg CA, Bradford DE, Kaye JT, Curtin JJ. Increased startle potentiation to unpredictable stressors in alcohol dependence: possible stress neuroadaptation in humans. *J Abnorm Psychol.* 2017;126(4):441-453. [FREE Full text] [doi: [10.1037/abn0000265](https://doi.org/10.1037/abn0000265)] [Medline: [28394145](https://pubmed.ncbi.nlm.nih.gov/28394145/)]
138. Hefner KR, Moberg CA, Hachiya LY, Curtin JJ. Alcohol stress response dampening during imminent versus distal, uncertain threat. *J Abnorm Psychol.* 2013;122(3):756-769. [FREE Full text] [doi: [10.1037/a0033407](https://doi.org/10.1037/a0033407)] [Medline: [24016014](https://pubmed.ncbi.nlm.nih.gov/24016014/)]
139. Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, et al. The prevention and treatment of missing data in clinical trials. *N Engl J Med.* 2012;367(14):1355-1360. [FREE Full text] [doi: [10.1056/NEJMsrl203730](https://doi.org/10.1056/NEJMsrl203730)] [Medline: [23034025](https://pubmed.ncbi.nlm.nih.gov/23034025/)]
140. Little RJ, Cohen ML, Dickersin K, Emerson SS, Farrar JT, Neaton JD, et al. The design and conduct of clinical trials to limit missing data. *Stat Med.* 2012;31(28):3433-3443. [FREE Full text] [doi: [10.1002/sim.5519](https://doi.org/10.1002/sim.5519)] [Medline: [22829439](https://pubmed.ncbi.nlm.nih.gov/22829439/)]
141. McCleary L. Using multiple imputation for analysis of incomplete data in clinical research. *Nurs Res.* 2002;51(5):339-343. [doi: [10.1097/00006199-200209000-00012](https://doi.org/10.1097/00006199-200209000-00012)] [Medline: [12352784](https://pubmed.ncbi.nlm.nih.gov/12352784/)]

Abbreviations

DSM-5: Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition)
EMA: ecological momentary assessment
HIPAA: Health Insurance Portability and Accountability Act
LLM: large language model
MOST: Multiphase Optimization Strategy
OSF: Open Science Framework
STAR: Smart Technology for Addiction Recovery
TLFB: Timeline Follow-Back

Edited by J Sarvestan; The proposal for this study was externally peer-reviewed by the National Institute on Alcohol Abuse and Alcoholism (R01AA031762). See the Multimedia Appendix for the peer-review report; Submitted 07.Aug.2025; accepted 17.Oct.2025; published 07.Nov.2025.

Please cite as:

Wyant K, Sant'Ana SJ, Punturieri CE, Yu J, Fronk GE, Maggard CM, Janssen C, Wanta SE, Kornfield R, van Swol LM, Curtin JJ
Maximizing Engagement, Trust, and Clinical Benefit of AI-Generated Recovery Support Messages for Alcohol Use Disorder: Protocol for an Optimization Study

JMIR Res Protoc 2025;14:e81697

URL: <https://www.researchprotocols.org/2025/1/e81697>

doi: [10.2196/81697](https://doi.org/10.2196/81697)

PMID:

©Kendra Wyant, Sarah J Sant'Ana, Claire E Punturieri, Jiachen Yu, Gaylen E Fronk, C Michael Maggard, Christopher Janssen, Susan E Wanta, Rachel Kornfield, Lyn M van Swol, John J Curtin. Originally published in JMIR Research Protocols (<https://www.researchprotocols.org>), 07.Nov.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.researchprotocols.org>, as well as this copyright and license information must be included.