

Sense2Stop: A Micro-randomized Trial Using Wearable Sensors to Optimize a Just-In-Time-
Adaptive Stress Management Intervention for Smoking Relapse Prevention

Samuel L. Battalio¹, David E. Conroy^{1,2}, Walter Dempsey³, Peng Liao⁴, Marianne Menictas⁴,
Susan Murphy⁴, Inbal Nahum-Shani³, Tianchen Qian⁶, Santosh Kumar⁵, & Bonnie Spring¹

1. Department of Preventive Medicine, Northwestern University Feinberg School of
Medicine, 680 N. Lakeshore Drive, Suite 1400, Chicago, IL 60611
2. Department of Kinesiology, Penn State University, 266 Recreation Building, University
Park, PA 16802
3. Survey Research Center, University of Michigan, 426 Thompson Street, Room 2464,
Ann Arbor, MI 48106
4. Department of Statistics, Harvard University, Science Center 400 Suite, One Oxford
Street, Cambridge, MA 02138
5. Department of Computer Science, University of Memphis, 319 Dunn Hall, Memphis, TN
38152
6. Department of Statistics, University of California, Irvine, Irvine CA 92697

Clinical Trials Registration #NCT03184389

CORRESPONDING AUTHOR: Bonnie Spring, Ph.D., Department of Preventive Medicine,
680 N. Lakeshore Drive, Suite 1400, Chicago, IL 60611, bspring@northwestern.edu

Abstract

Background. Relapse to smoking is commonly triggered by stress, but behavioral interventions have shown only modest efficacy in preventing stress-related relapse. Continuous digital sensing to detect states of smoking risk and intervention receptivity may make it feasible to increase treatment efficacy by adapting intervention timing.

Objective. Aims are to investigate whether the delivery of a prompt to perform stress management behavior, as compared to no prompt, reduces the likelihood of (a) being stressed and (b) smoking in the subsequent two hours, and (c) whether current stress moderates these effects.

Study Design. A micro-randomized trial will be implemented with 75 adult smokers who wear Autosense chest and wrist sensors and use the mCerebrum suite of smartphone apps to report and respond to ecological momentary assessment (EMA) questions about smoking and mood for 4 days before and 10 days after a quit attempt and to access a set of stress-management apps. Sensor data will be processed on the smartphone in real time using the cStress algorithm to classify minutes as probably stressed or probably not stressed. Stressed and non-stressed minutes will be micro-randomized to deliver either a prompt to perform a stress management exercise via one of the apps or no prompt (2.5 - 3 stress management prompts will be delivered daily). Sensor and self-report assessments of stress and smoking will be analyzed to optimize decision rules for a just-in-time adaptive intervention (JITAI) to prevent smoking relapse.

Significance. Sense2Stop will be the first digital trial using wearable sensors and micro-randomization to optimize a just-in-time adaptive stress management intervention for smoking relapse prevention.

KEYWORDS: micro-randomized trial, optimization, mHealth, smoking, stress, digital intervention, just-in-time adaptive intervention

1. Introduction

Cigarette smoking remains responsible for one in five deaths¹ and U.S. healthcare costs estimated at \$300 billion per year.¹ Despite research and public policy efforts, smoking prevalence has remained stable since 2015.² Although more than half of smokers attempt to quit each year, only 7.5% succeed.³ Most relapses to smoking occur in the first week after quitting, demonstrating that timely, effective, scalable intervention is needed to prevent relapse.⁴

Many smokers attribute relapses to stress and report that⁵⁻⁷ self-medicating with nicotine dispels the distress caused by nicotine withdrawal⁸ and other stressors.⁹ For smokers who do quit, stress jeopardizes abstinence by diminishing self-confidence about the ability to manage negative emotions and resist temptation to smoke.^{10,11} Behavioral interventions encourage mastering self-regulation strategies besides smoking to manage stress,^{9,12} but to date have shown only modest efficacy. One reason for their modest efficacy may be that gold standard smoking and stress management interventions are static (i.e., not delivered in moments and contexts where they might be maximally effective). Therefore, Sense2Stop was designed to examine whether treatment efficacy might be enhanced by adapting the timing of intervention¹⁵ delivery to align with changing states of vulnerability to smoking and receptivity to intervention.

Stressors that trigger smoking relapse occur precipitously in the hours immediately preceding a smoking event.¹³ Efforts to predict or detect when stress management is needed have been impeded by reliance on retrospective, bias-prone self-reports of stress and smoking.^{15,16} To address this challenge, the Sense2Stop trial measures stress¹⁷ objectively via wearable sensors, and smoking¹⁸ via both wearable sensors and subjective self-report. Psychophysiologic and movement signals from sensors transmit wirelessly to a smartphone, where algorithms process them in near real-time to detect the probability that stress has occurred. Stress is hypothesized to

increase a person's state of risk of smoking, while simultaneously creating cognitive interference that decreases their receptivity to intervention.^{19,20} Hence, stress is conceptualized as a potential tailoring variable that determines whether and when intervention should occur. The Just-in-Time-Adaptive Intervention [JITAI] to be optimized in Sense2Stop will be a digital prompt: the unobtrusive opening of one of several smartphone apps that guides the user through stress management exercises.

Sense2Stop is a micro-randomized trial (MRT) whose purpose is to inform the development of an effective JITAI decision rule to deliver intervention when needed to prevent a smoking lapse. The protocol is to randomly, in real time, assign stressed and non-stressed minutes to receive a digital intervention prompt or no prompt. The aim is to learn whether the intervention prompt should occur under a state of stress (increased risk of smoking/need for stress regulation) or a state of no stress (increased intervention receptivity). The primary hypothesis is that delivery of a prompt to perform a stress management exercise via one of the mobile apps, as compared to no prompt, will reduce the likelihood of being stressed in the subsequent two hours, and this effect will be stronger if the prompt is delivered when the individual is stressed. The secondary hypothesis is that delivery of a prompt to perform a stress management exercise will reduce the odds of smoking for the next two hours, and this effect will be stronger if the prompt is delivered when the individual is stressed.

2. Materials and Methods

2.1 Study Design

The Sense2Stop study will enroll 75 adult smokers who express willingness to quit. For the first 4 days of the 14-day study period, participants will receive cognitive-behavioral smoking cessation treatment consisting of in-person and telephone coaching in conjunction with

a digital intervention that includes a set of smartphone stress-management apps. Participants will wear ambulatory sensors that generate signals for which detection algorithms have been validated previously to detect smoking puffs and states of physiologic stress. These sensors will be wirelessly connected to the participant's smartphone, which will apply algorithms to detect the participant's physiologic stress in real time. Participants will be coached to prepare for quitting on days 1-3, to quit on day 4, and, thereafter, to attempt to abstain from smoking.

For the micro-randomized trial analyses, sensor data for the 10-day post-quit period will be processed to classify whether the participant is probably-stressed or probably-not-stressed during each minute of each day. Minutes, stratified by stress classification and time of day, will be randomly allocated (micro-randomized) to treatment, which involves delivery of either a stress management prompt or no prompt. Each day, 0-6 stress management prompts will be delivered, averaging to 2.5 - 3 prompts per day based on our micro-randomization algorithm (see Section 2.12). Hence, the independent variable (i.e., treatment condition) is whether the intervention (stress management prompt) is delivered versus not delivered. The moderating (potential tailoring) variable is whether the individual is stressed or not stressed at the time of micro-randomization. The primary (proximal) outcome is whether stress occurs in the 2-hour window following micro-randomization. The secondary outcome is whether smoking occurs in the 2-hour window following micro-randomization. Sense2Stop Study's overarching objective is to inform the development of a JITAI that prompts smokers attempting to quit to perform stress management exercises in a manner that prevents smoking relapse.

All study procedures will be approved by the Northwestern University Institutional Review Board (IRB) prior to study recruitment (IRB Number STU00201566). A diagram showing the sequence of study procedures appears in Figure 1.

Figure 1. Sense2Stop Study Protocol Sequence.

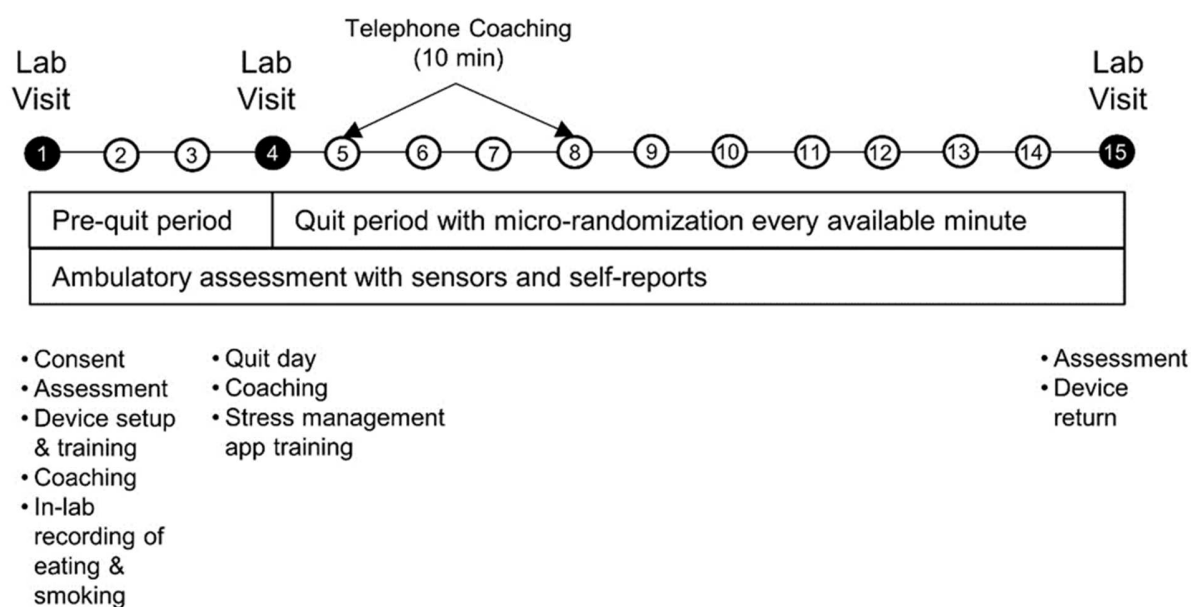


Figure 1 caption: Shows periods of sensor wear and EMA assessments during 4-day pre-quit period and 10-day post-quit period. Laboratory visits and coaching calls also displayed.

2.2. Eligibility

At the time of screening, all participants must be between the ages of 18 and 65 years and be an active smoker of one or more tobacco cigarettes per day for the past year. Participants are not permitted to use any other form of tobacco product or nicotine replacement during the study period, and must express willingness to quit smoking for at least 48 hours during the quit phase of the 15-day study. Use or intention to use pharmaceutical smoking cessation aids (e.g. varenicline, bupropion SR, nicotine replacement) during the study period is also exclusionary. Additionally, participants will not be enrolled if they are unable to wear study devices (e.g., due to skin irritation, size limitations), are planning to move outside of the Chicagoland area during

the study period, or cannot read, write, or speak English. Prisoners and women who are pregnant or planning to become pregnant will not be enrolled.

2.3. Recruitment and Screening

Participants will be recruited through flyers posted on the Northwestern University Feinberg School of Medicine campus in Chicago and through social media (i.e., geo-localized advertisements on Facebook.com). Advertisements will target individuals who meet the age inclusion criteria and live within a 25-mile radius of Chicago. Interested volunteers will complete a brief online screening questionnaire. Those who qualify will be contacted by a research assistant and asked to complete a 10-minute telephone screening interview, during which study eligibility is confirmed. An initial visit will be scheduled and two documents will be mailed for participants to read before coming to the Day 1 study visit: 1) the study consent form, and 2) a handout about strategies to achieve smoking cessation. Participants will be asked to provide contact information for two locator people who can be reached in the event of an emergency, lost contact with the participant, or need for device return.

2.4. Day 1 Visit.

The first study visit, held while participants are still smoking, will last approximately 3.5-4 hours. Procedures to be conducted during Visit 1 are: consent process, measurement of height and weight, survey completion, device setup and training, and coaching about forthcoming smoking cessation plans. The visit will take place in an indoor laboratory space with a dedicated ventilation system and a negative pressure room designed to allow smoking.

2.4.1. Questionnaires.

During the Day 1 visit, participants will use REDCap to complete a number of questionnaires and surveys. The surveys assess demographic characteristics, cigarette use

behaviors, and psychosocial characteristics. Questionnaires to be administered will be the Fagerstrom Test for Nicotine Dependence,²¹ Intolerance for Smoking Abstinence Discomfort,^{22,23} the Barratt Impulsiveness Scale (BIS-11),²⁴ job strain²⁵ and parental strain²⁶ assessments, the Center of Epidemiological Studies – Depression (CES-D)^{27,28} scale, the Beck Anxiety Inventory,²⁹ and the Interpersonal Support Evaluation List -12.^{30,31}

2.4.2. Device Setup and Training

All participants will be loaned the Autosense suite of devices to use during the 14-day study. The package of devices includes: 1) two wrist-worn accelerometers, 2) a chest-worn ECG and breathing monitor, and 3) a smartphone. The smartphone contains a suite of mCerebrum apps²⁷ including a Sense2Stop app that records when participants indicate that they began and ended their day and shows participants whether the sensors are properly attached. The app enables participants to complete event-triggered Ecological Momentary Assessments (EMAs) to report when they smoke. The app also prompts them randomly throughout the day to complete EMAs about their smoking behavior, smoking urge, and mood. Finally, the app delivers an end-of-day survey that asks participants to report when they smoked each cigarette throughout the day. In addition to app features that are visible to participants, the smartphone houses algorithms that process incoming sensor data to detect stress and smoking in real time and that trigger transmission of data to study servers at regular intervals.

Study staff will demonstrate how to charge the study devices overnight and how to wear them during the day. They will help participants put on the devices for the first time and, as a memory aid, will provide take-home handouts with photos showing how to wear the devices. Participants will be asked to wear the wrist and chest sensors as often as possible throughout

each day, putting them on when they first wake up, and only taking them off to charge when they go to bed.

2.4.3 Coaching.

Smoking cessation treatment will involve gold standard cognitive behavioral quit smoking counseling in combination with motivational interviewing. Four treatment sessions (the number that Medicare reimburses per quit attempt) will be provided in a hybrid format involving in-person and telehealth delivery, both of which are evidence-based for smoking cessation.^{33,34}

The first of four smoking cessation coaching sessions will occur in person at the Day 1 visit.

During this 60-minute session, participants will meet with a Bachelor's level health coach, trained and supervised by a licensed clinical health psychologist (BS), to provide standard-of-care motivational interviewing and cognitive-behavioral smoking cessation treatment.^{33,34}

Participants will reflect on how their core values relate to the pros and cons of quitting versus not quitting smoking, receive psychoeducation about nicotine withdrawal, and learn effective behavior change techniques to prepare for the post-quit period (i.e., self-monitoring, goal setting, use of social support). They will learn to self-monitor by completing "pack wraps"³⁰ to track their cigarette use and identify smoking triggers. They will also be taught coping strategies to avoid smoking triggers, withstand cravings, and substitute other responses and rewards besides smoking. Additionally, they will learn to practice goal-setting and self-regulation by tapering their smoking intake to increase self-efficacy about being able to quit.

In the second half of Day 1's coaching session, participants will learn how stress contributes to the persistence of smoking and functions as a trigger for smoking behavior. They will be taught to exercise different stress management techniques by using the three stress

management apps: urge surfing (Mood Surfing), cognitive reframing (Thought Shake-Up), and attentional control (Headspace) (see section 2.9.1).

2.5. Day 4 (Quit Day) Visit.

After several days of wearing the devices and using the smartphone app to answer EMAs, participants will return to the lab on Day 4 for their second in-person study visit. The Day 4 session lasts for 1.5 – 2.0 hours and involves smoking cessation coaching and additional training on how to use the stress management apps. The session will begin with a discussion of smoking “slips” and how to prepare for them in the days ahead. Then, a majority of the Day 4 session will facilitate experiential learning of the skills engaged by each of the three smartphone stress management apps: Mood Surfing, Thought Shakeup, and Headspace (see section 2.9.1 below). Participants will be advised that they may use the apps whenever they choose, and that they also will receive periodic prompts to use the apps throughout the remainder of the study. Next, coaches will facilitate a discussion about SMART goals³¹ and will help participants set a SMART goal to pursue between Days 4 and 5. Finally, participants will be given a brief overview of what to expect on the Day 5 and 8 coaching calls (see below) and will schedule these calls with their coach. Participants will be asked to quit smoking cigarettes from Day 4 through the remainder of the trial.

2.6. Day 5 and Day 8 Coaching Calls.

Participants will be contacted by their health coach on study Days 5 and 8, corresponding to 1 and 4 days post-quit date. Each coaching call will last approximately 10 minutes and will cover 3 topics: 1) progress toward the SMART goal set on the last visit/call; 2) benefits of quitting smoking experienced thus far; and 3) skills taught on Day 1 or 4 (e.g., avoid/substitute, tapering) that the participant utilized to overcome barriers to cessation. Coaches will leave time

at the end of each call to discuss any technical challenges with the smartphone applications or devices. The call will conclude with setting SMART goals about strategies to use to remain smoke-free and/or overcome barriers to cessation.

2.7. Day 15 Visit.

The final study visit will occur on Day 15: one day after the end of the 14-day study protocol.

At this visit, participants will complete a survey about device use, and return all study devices.

2.8. Sensors and Devices

2.8.1. AutoSense chestband sensors.²⁷

The AutoSense chestband sensors include: 1) a two-lead electrocardiogram (ECG) measuring electrical activity from the heart, 2) a respiratory inductive plethysmography band to measure relative lung volume and breathing rate at the rib cage, and 3) a three-axis accelerometer to assess motion artifacts in the data and provide inferences about physical activity. The unit is small (1'' x 2.5'') and includes a rechargeable battery to stream sensor measurements via wireless radio to a smartphone in real-time. The battery can last 10+ days without recharging. The unit hooks onto the respiratory band via a small hook, and two sets of wires extend from it. One set corresponds to two ECG leads for which electrodes are fastened to the center and left side of the chest. The second set of wires from the Autosense unit box plug into the respiratory band. The band and unit box will be worn underneath participants' clothing and across the chest, just underneath the armpits. Participants will be asked to wear these during all waking hours, and to take them off for sleep at night. They may charge the devices overnight if they wish and will be given enough ECG leads to change them every day. The AutoSense chestband has been worn by hundreds of participants in various lab and field studies of smoking cessation, stress, and drug use.^{18,30,31}

2.8.2. MotionSense wristband sensors.

MotionSense wristbands to be worn on each wrist sample 3-axis accelerometry at 16 Hz and 3-axis gyroscope at 32 Hz via continuously streaming Bluetooth LE. The device is powered by a small rechargeable battery that lasts for 24+ hours on a single charge. Participants will be advised to wear the devices all day during waking hours, and to charge them overnight using wireless JetTech charging pads.

2.9. Smartphone and Study Applications

mCerebrum smartphone software²⁷ consists of a suite of apps including a Sense2Stop app that triggers ecological momentary assessment (EMA)²⁷ surveys randomly on the dedicated study smartphone (Samsung Galaxy S4). The software signals participants via sound, vibration, and a visual notification to start their day. The time of the start-of-day notification is individually set per participant preference prior to the data collection period. Upon receiving the notification, the participant is required to press a button on the smartphone to acknowledge the start of the day and begin the first EMA block. Similarly, an end of day prompt, scheduled per participant preference, is pre-programmed to conclude the final EMA block and trigger the end-of-day EMA. The software is designed such that participants also may self-trigger the end of day EMA by pressing a smartphone button that becomes activated 3 hours prior to the pre-programmed end time. The study start and study end times only define the interval when EMAs can be triggered by the study software; sensor data will be collected at all times throughout the day. The software notification schema is designed to reduce burden by giving participants agency over the times when they may potentially be interrupted and asked to respond to EMAs or use study apps, while also collecting sufficient data to answer the study's research questions. Our experience with protocols of this type is that participants are very willing to engage with the

start-of-day notification in order to begin earning the micro-incentives associated with receiving and responding to EMAs.

The Sense2Stop protocol uses three types of EMAs: Random, Event-Triggered, and End-of-Day. These three EMA surveys can be found in Supplementary File 1. One Random EMA is triggered at a randomly selected time during each of three, 4-hour blocks that commence when the participant starts their day (e.g., 8am -12pm, 12pm-4pm, and 4pm-8pm), totaling 3 random EMAs triggered per day. Each random EMA includes 12-15 questions about smoking, craving, positive affect, negative affect and environmental cues or affordances for smoking since the last prompt. (3 questions appear or not depending on branching logic.) The validated single item craving score from the Minnesota Tobacco Withdrawal Scale was used to assess craving, and has correlated with sensor indicators of smoking in previous research.^{41,42,43} The affect items, adapted from the Perceived Stress Scale,⁴⁴ have been associated with sensor indications of stress in prior studies³⁸ using EMA designs.

Participants also use the app to self-report whenever they smoke. Each day, 3 or 4 randomly selected self-reported smoking episodes initiate an Event-Triggered EMA that asks the questions delivered in the random EMAs. Only a limited number of self-reported smoking episodes initiate Event-Triggered EMAs in order not to inadvertently incentivize smoking, since EMA responding is monetarily rewarded. Finally, the End of Day EMA asks participants to report the hour-long blocks of time in which they smoked.

2.9.1. Stress Management Smartphone Interventions

Participants will have access to three stress management apps at any time during the study's 10-day post-quit period. They will also occasionally be prompted to use one of the three randomly chosen apps. Once the software randomly selects the app, it triggers a message

recommending this specific app to appear on the phone (e.g., “Time to surf your mood”). The message also includes 2 response options: (a) press ‘OK’ to launch the corresponding app; or (b) press ‘Cancel’ to make the prompt go away.

Mood Surfing. The Mood Surfing app is custom-designed in accordance with principles of Acceptance and Commitment Therapy (ACT).⁴⁵ The goal is for participants to utilize present-focused awareness of their thoughts and emotions without self-judgment. The app includes three activities: Use Your Imagination, Notice and Accept, and Surf the Mood. The Use Your Imagination activity⁴⁵ asks participants to visualize unpleasant thoughts or emotions as tangible objects placed on a table, and to describe what these objects looks like (e.g., color, size, shape). This process is intended to help participants distance themselves from negative thoughts that could impede smoking cessation. The Notice and Accept activity⁴⁶ asks participants to notice physical sensations caused by stress (e.g., pain, craving) and to describe the sensations (e.g., their strength, whether they change or remain the same). Deliberatively focusing on and accepting sensations without judgment can strip them of arousing emotions, lowering their perceived intensity.⁴⁶ Finally, the Surf the Mood activity is presented as an audio recording of waves crashing on the shore. While listening, participants are asked to imagine their moods as “waves” of the ocean. Metaphorically “surfing” or “riding out” the mood allows it to subside, just as the wave eventually calms as it reaches the shore.⁶

Thought Shakeup. Another custom-designed app, Thought Shakeup,⁴⁵ applies a Cognitive-Behavioral Therapy (CBT)⁴⁷ approach by asking participants to reframe or rephrase, negative thoughts that could interfere with smoking cessation. After entering a thought into the app, participants answer a series of questions about it, such as how true it seems, and why it may be correct or incorrect. Participants are then asked to reframe the negative thought in a more

realistic manner. Having gone through this exercise, they re-rate the apparent truth of both the original thought and the reframed thought to learn whether the original, distorted thought has lost some veracity.

Headspace. The commercial Headspace app includes a variety of guided meditation exercises that range from brief (e.g., 10 minutes of guided breathing) to extended and complex. Participants will have access to the entire Headspace suite and be able to select meditations that are useful for their current state (e.g., to quell anxiety; to promote sleep). They will use a study-provided email address and password to register for an account, ensuring that the manufacturer receives no identifiable participant information. Before using the application, participants will be advised that Headspace will receive information about their data usage, but not any of their identifiable data. Participants will be asked to sign a Privacy Acknowledgement that Northwestern University will retain.

2.10. Financial Incentives

Study incentives were designed to foster protocol adherence without confounding interpretation of intervention effects. Study participants will be offered an honorarium of \$25 for completing the Day 1 laboratory visit, \$20 for completing the Day 4 visit, and \$30 for completing the Day 15 visit. Responding to EMAs thoroughly and in a timely manner while wearing the study sensors will be financially incentivized. The base payment for completing Random and Event-Triggered EMAs will be \$0.50 per survey. To earn an incentive, participants are required to have worn the sensors for >70% of the 4-hour epoch during which the EMA occurred. Over the course of the study, each participant will be able to earn an incentive for a total of 42 random and 42 event-triggered EMAs. The number of Event-Triggered EMAs that participants are able to elicit by self-reporting smoking will be capped at 3-4 per day so as not to

reinforce smoking inadvertently. A bonus of \$0.25 will be given for each EMA survey completed within 5 minutes of receiving the EMA prompt.⁴⁸ A further bonus of \$13.50 will be given for completing $\geq 90\%$ of all random EMAs, and another bonus of \$70 will be given for wearing the sensors for $\geq 70\%$ of each 16-hour day. When participants signal the end of their day, they will be prompted to report the time of each cigarette smoked in a diary log. Completion of each End-of-Day log will be incentivized at \$1.00 per log. The maximum possible compensation for full adherence to all incentive contingencies will be \$250.

2.11. Measures of Stress and Smoking

2.11.1. cStress Detection Algorithm

Physiological stress will be detected continuously using cStress, a machine-learned model that uses Autosense's sensor-derived electrocardiogram (ECG) and respiration data to identify states of physiologic stress. cStress was developed and validated against laboratory gold standards (i.e., cold pressor test, public speaking, mental arithmetic), and then further validated against EMA self-reports in the field.³⁰ At one-minute intervals, cStress produces a stress likelihood (probability) measure scaled to be between 0 and 1. The cStress machine learning model, based on a support vector machine (SVM) that produces stress likelihood, is applied to each minute. In one study the cStress model showed an F1 score of 0.81 for differentiating in lab data whether participants were in a stress session or at baseline, and in an independent lab study, the cStress model obtained an F1 score of 0.9. When compared with self-reported stress on a weeklong field study, cStress showed a median F1 score of 0.71 and when tested in two other field studies, it produced median F1 scores of 0.72 and 0.65 when compared with self-reported stress. The cStress model produces hundreds of outputs each day, which are inherently noisy due to imperfections in the trained model.⁴⁹ Therefore, from these minute-level stress probability

measures, a “stress episode” detection method is used to link each minute with the previous minutes, forming a sequence of minutes (a time series) that can be evaluated together.⁴⁹ Incorporating data about previous minutes provides temporal context (“What is going on *before* a given minute?”), renders the minute-level cStress detection algorithm robust to short-term fluctuations and data aberrancies, and increases its accuracy. As Figure 2 indicates, the “stress episode” detection method also filters out episodes for which more than 50% of minutes show evidence of poor sensor connection (e.g., poor adherence to the body, or unstable connection to the smartphone) or contain physical activity (because physical activity confounds sensor-based detection of stress). The physiological signs of physical activity mimic those of stress because physical activity activates many of the same physiologic responses that would be activated by stress. The final output of this “stress episode” detection method is a real-time determination of whether each minute (accounting for previous minutes) is more likely to be part of a “probably stressed” episode or a “probably not stressed” episode.¹⁷ (Additional details about the algorithms used in this stress episode detection method can be found in Sarker et al (2017).¹⁷)

2.11.2. puffMarker Smoking Detection Algorithm

The previously validated machine-learned puffMarker algorithm¹⁸ will be applied continuously to wrist-worn inertial sensor and chest-worn respiration data to identify smoking episodes. In brief, puffMarker first identifies candidate wrist-to-mouth smoking movements. Next, false positive wrist-to-mouth movements are screened and filtered based on features (e.g., hand orientation, duration, degree of movement, etc.) that suggest a confounding movement (i.e., yawning, eating, etc.) rather than a cigarette puff. Then, puffMarker selects respiration cycles that correspond to the wrist-to-mouth movements. Several features that are used by the machine learned model to detect smoking puffs are extracted from the breath cycle corresponding to the

detected hand-to-mouth gesture and from micro patterns in wrist motion from wrist-worn accelerometer and gyroscope sensors. Post-processing occurs next, during which isolated puffs (single puffs that are not proximal to other puffs) are removed, and puffs from the non-dominant hand are discarded. (Non-dominant hand data are discarded because prior research using puffMarker found few instances when participants used the nondominant hand for smoking.)¹⁸ Finally, smoking episodes are identified if four or more adjacent puffs are detected by puffMarker¹⁸. The puffMarker model was reported to achieve a recall rate of 96.9% and a false positive rate of 1.1% on training data. When tested in a field study of newly abstinent smokers, the model achieved a recall rate of 87.5% and one false positive in every 6 days of sensor wearing.

2.12. Micro-Randomization

From the participant's quit day (day 4) onward through day 14, each minute is evaluated for potential micro-randomization to receive either (1) a stress management prompt (i.e., opening of one of the three randomly determined stress management smartphone apps) or (2) no prompt. Randomization will only occur if: (1) according to the cStress detection algorithm, the participant is in the middle of a "probably stressed" episode or a "probably not stressed" episode; (2) the participant is available (i.e., not driving; the phone battery is > 10%); (3) the participant is not overburdened, that is: (a) has not received an EMA (random or event-triggered) in the last 10 minutes; (b) has not received a stress management intervention prompt within the last hour. This decision rule for randomization availability is diagrammed in Figure 2. (All programming source code for the Sense2Stop study can be found at <https://github.com/MD2Korg>. The specific configuration for availability is at <https://github.com/MD2Korg/mCerebrum-Configuration/tree/master/1.0/> /Northwestern/STU00201566/mCerebrum/org.md2k.ema_scheduler

Figure 2. Availability for Randomization to Intervention Prompt versus no Prompt.

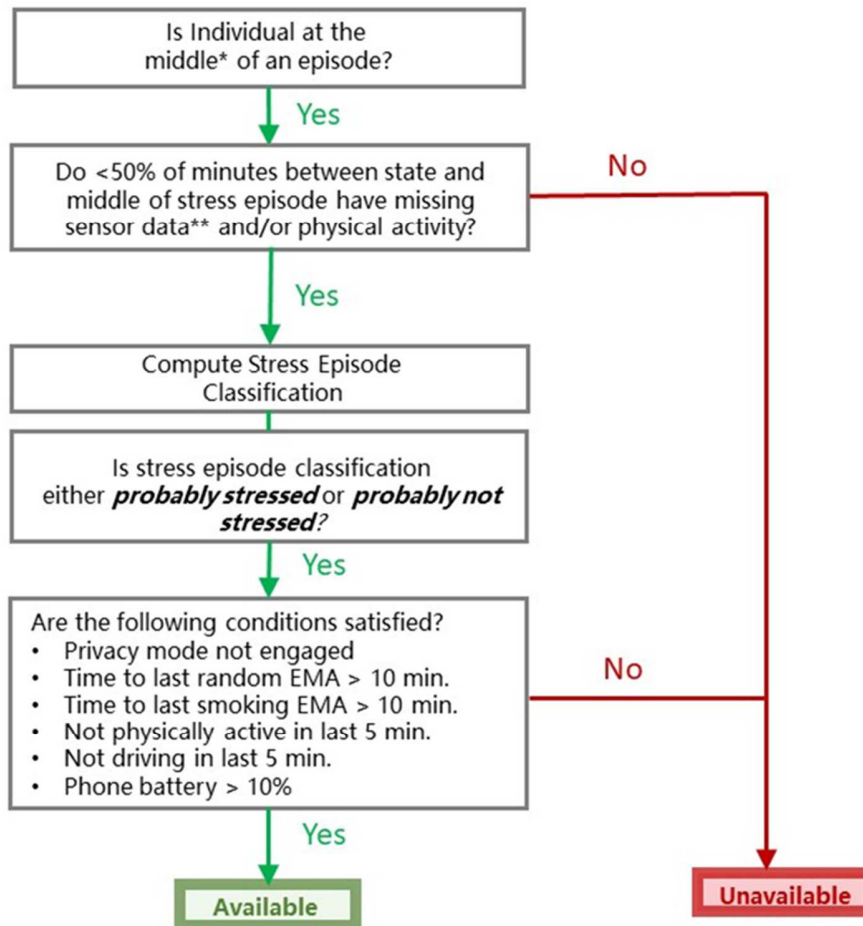


Figure 2 caption: Sense2Stop algorithms running on smartphone detect episodes of stress or no stress when fewer than 50% of the minutes between the start and midpoint of an episode have missing data or physical activity. Such minutes are available to be randomized in real time for the mobile phone to deliver an intervention prompt or no prompt if criteria are met for low burden, safety, and battery power.

*The middle of an episode (also called the episode's peak) is determined using a time-series method called Moving Average Convergence Divergence.^{49, 50}

**Causes of missing sensor data include sensor detachment and dropped WiFi connection.^{51,52}

The probability of a minute being randomized to deliver an intervention prompt versus no prompt will be balanced according to two factors: (1) whether the participant has already experienced a post-quit smoking lapse or not and (2) whether the current episode is classified as

“probably stressed” or “probably not stressed.” Specifically, probabilities will be weighted such that participants who have *not* yet lapsed will receive 0-6 prompts per day with an average of 3 prompts per day (a daily average of 1.5 during episodes classified as probably stressed and a daily average of 1.5 during episodes classified as probably not stressed). Because physiological stress episodes decrease after a smoking lapse,⁵³ participants who have already lapsed will receive an average of 2.5 prompts per day (a daily average of 1 prompt during episodes classified as probably stressed and a daily average of 1.5 during episodes classified as probably not stressed). Randomization is performed using a micro-randomization algorithm; see the Supplementary File 2, section 3 for details.⁵⁴

2.13. Measurement of Study Outcomes,

2.13.1. Primary Outcome Measure

The primary study outcome, assessed during the 2-hour window following each micro-randomization is the participant’s probability of experiencing a stress event. This outcome will be measured longitudinally over the 120 minutes in every post-randomization window. Because both physical activity and stress activate cardiovascular and autonomic nervous system responses,^{55,56} the quality of stress detection from wearable sensors is poor during periods of physical activity. Therefore, primary outcome assessment needs to appropriately disambiguate periods of physiological activation due to stress from those due to confounding physical activity. Accordingly, Sense2Stop’s primary outcome will be trichotomous, reflecting the probability that the cStress algorithm has classified each period during the 120 minutes immediately post-randomization as “probably stressed,” “probably not stressed,” or “physically active.” Note that use of the “physically active” outcome category does not impair the ability to discriminate periods of probable stress from periods of probable nonstress: its use in the trichotomous

outcome simply parses out measurement error arising from confusability between sensor signals of physical activity and those of stress.

The statistical methodology for use with the primary trichotomous outcome is outlined in the Supplementary File 2 and will be described in greater detail in the primary outcome publication. If there are no missing data minutes during this interval, the number of probably stressed, probably not stressed, and physically active post-randomization minutes that are evaluated to determine our trichotomous outcome will equal 120. Each minute will be labelled as probably stressed, probably not stressed or physically active using the stress episode detection method (described above in Section 2.11.1). Minutes that do not correspond to one of these three values (e.g., due to poor sensor adherence to the body, or unstable connection to the smartphone) will be treated as missing minutes. The statistical method for the *primary* analysis (described below) will evaluate the 120-minute measurements following intervention versus no intervention in order to determine whether there is an overall treatment effect.

2.13.2. Secondary Outcome Measure.

A deterministic rule that incorporates both PuffMarker and self-report of smoking will be used to identify true smoking events for study analyses. First, positive smoking events identified by PuffMarker and confirmed by self-report data will be treated as actual smoking events. PuffMarker smoking events that are not confirmed by self-report will be discarded. Then, self-reports of smoking provided in response to random EMAs or spontaneously reported as event-triggered EMAs will be treated as actual smoking events, even in the absence of PuffMarker corroboration. The rationale is twofold: a) that PuffMarker, although promising, has not yet been validated against gold standard biomarkers of smoking outside of the laboratory, and b) that bias in self-reporting smoking is toward underreporting rather than overreporting.^{57,58}

A diagram of the study design, assessments and outcomes appears in Figure 3. Note that the primary and secondary outcomes examined in Sense2Stop are proximal outcomes – short term effects of the intervention that are posited to fall along the mechanistic causal pathway by which it influences the distal outcome. Effects of the intervention on the distal outcome – the probability of full smoking relapse during the 10-day period post-quit date – cannot be measured in an optimization trial. Impact on the distal outcome will be measured once decision rules for the EMI are derived from this MRT; that algorithmic decision rule will then be tested in an RCT against a control, which might be a fixed intervention or usual care.

Figure 3. Design of the Sense2Stop Micro-Randomized Trial.

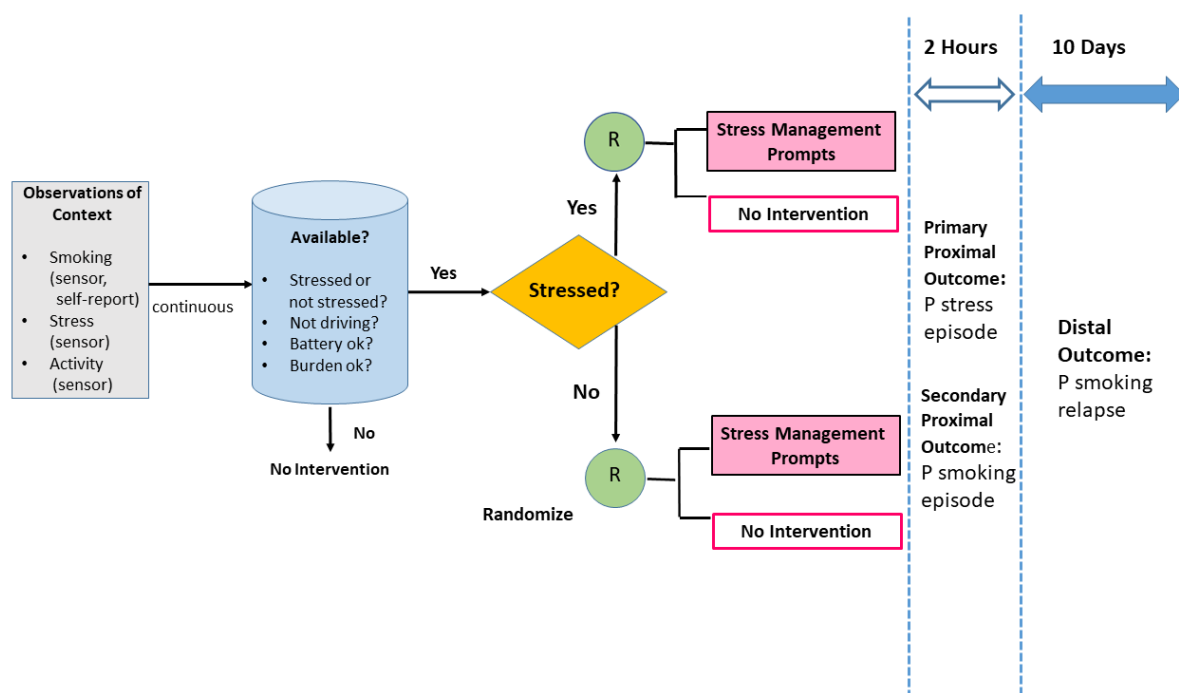


Figure 3 caption: Depicts assessment modalities (sensor, self-report) used to observe Sense2Stop proximal outcomes (smoking, stress) and confounders (physical activity). Algorithms operating on the mobile phone transmit episodes of continuous digital data for micro-randomization (R) to intervention or no intervention only when the participant is available and the episode is able to be classified as probably stressed or probably not stressed. Proximal outcomes (probabilities of stress and smoking) are assessed in the 2-hour window after each micro-randomization. Note that the 10-day distal outcome is not being directly assessed in either the primary or secondary analytic plans outlined in the current manuscript.

2.14. Analytic Plan

2.14.1. Primary aim.

The primary hypothesis is that administration of a prompt to perform a stress management exercise, as compared to no prompt, will reduce the likelihood of being stressed in the subsequent two hours, and this effect will be stronger when the prompt is administered when the individual is stressed. For each micro-randomization, we will compute the minute level proximal outcome (described above in Section 2.13.1) of whether the individual is probably stressed during each of the 120 minutes in the post-randomization window. Using a generalization of Boruvka et al. (2018), we will use log-linear regression to model the primary outcome (e.g., probably stressed), and will include control covariates that are most highly correlated with stress to reduce noise and increase our power to detect a causal treatment effect¹. To test the primary hypothesis, we will include two additional covariates: an indicator of stress episode type (probably stressed vs. probably not stressed) and an indicator of the randomized treatment condition (intervention vs. no intervention). Support for the hypothesis will be provided by finding a significant interaction between these two covariates, such that an intervention prompt has a greater likelihood of preventing the occurrence of proximal (short-term) stress when it is delivered during stress, as compared to no stress.

To account for missing data in the outcomes, we will investigate which observed variables can be used to explain the missing data. These variables will be included in the analyses to adjust for missingness. Additionally, to increase confidence in the accuracy of our results, we plan to conduct sensitivity analyses. In the case of the primary outcome, we will

label the missing minutes in two extreme ways: that is, as all probably stressed or as all probably not stressed to assess the dependence of the results on the missingness.

2.14.2. Secondary Aim.

The secondary hypothesis is that administration of a prompt to perform a relaxation exercise will reduce the odds of smoking for the next two hours, and this effect will be stronger if the prompt is administered when the individual is stressed. To test the secondary hypothesis, we will use a generalization of Boruvka et al., (2018) for use with binary outcomes.⁵⁹ At each available decision point, we will compute the binary proximal outcome of whether the individual smoked in the subsequent 120 minutes based on the smoking event times. As in Boruvka et al. (2018), we will use “robust standard errors,” combined with weighting, to accommodate the dependence in outcomes across decision points within each individual, yet avoid causal bias. As with the primary analysis, we will include baseline demographic and descriptive covariates to reduce noise and increase our power to detect a treatment effect. To assess moderation, we will include the output of the cStress algorithm as a binary moderator in the analysis.

To account for missing data in the outcomes, we will investigate which observed variables can be used to explain the missing data and will include these variables in the analyses to adjust for missingness. Sensitivity analyses for the secondary outcome will use exploratory data analysis methods to inform measurement-error models, i.e., specify the probability of the set of measurements including self-report (end-of-day log and EMA), and puffMarker given a true, “latent” smoking trajectory. A Markov Chain Monte Carlo algorithm will generate a set of “complete smoking event datasets” (i.e., 1000). We will then perform the analysis on each complete, synthetic dataset and take the average across these results. The sensitivity analysis will more carefully account for the noisy nature of the self-report and sensor data.

As outlined in the ‘Secondary Outcome’ section above, we will evaluate two complementary approaches to operationalizing the smoking outcome (the presence of a smoking event during each two-hour window). First, we will compute models that identify smoking using the deterministic rule (i.e., considering a smoking event present only when both puffmarker and self-report detect it). The deterministic rule will only use puffMarker to pinpoint the timing of smoking given a matched self-report of a smoking event, i.e., we rely on puffMarker as a more objective measure of smoking timing when we can match a self-report of smoking to that time. If the deterministic rule is considered untenable—that is, sensitivity analysis demonstrates large uncertainty in smoking times based on multiple measurements—then we will employ a sensitivity analysis using multiple imputation. Specifically, we will take the 1000 “complete smoking event datasets” and analyze them separately using the above analytic method and combine them using standard multiple imputation methods.

2.14.3. Sample size and Power

When this project was developed, given the novelty of its design, there were no validated methods to determine the necessary sample size for stratified MRTs. Hence, we planned sample size based on laboratory and field studies that performed continuous stress or smoking assessment from wearable sensors. These studies indicate good detection accuracy with sample sizes ranging between 20³⁰ and 72.⁶⁰ A sample size of 75 was planned based on the upper range of these sample sizes. The planned sample size of 75 did not account for additional considerations such as frequency of decision points or availability, nor potential attrition due to data quality. During this project, our team developed sample size resources for planning stratified MRTs to detect main effects of mHealth prompts.⁶¹ Extension of the developed sample size calculators is required when planning stratified MRTs to detect interaction between mHealth

prompts and a time-varying covariate (e.g., stress). Specifically, extending the simulation-based calculator to incorporate the complexity of the Sense2Stop study design is highly non-trivial, and therefore a post-hoc power analysis is not considered.

3. Discussion

Despite substantial research and public policy effort, smoking remains one of the costliest public health issues facing the world today.⁶² Smoking cessation interventions that can reach modern smokers to produce sustained abstinence are needed urgently. Existing smoking cessation interventions do little to address dynamic, in-the-moment risk processes (e.g., stress) that are known to contribute to smoking relapse. Consider, for example that smoking cessation and⁶³ stress management interventions are generally offered on a fixed schedule rather than in synchrony with when stress occurs or need for relapse prevention is greatest.^{15,16} Advances in mobile technologies (e.g., worn sensors wirelessly connected to one's smartphone) and real-time predictive analytics (e.g., smartphone-based machine learned models) allow for objective in-the-wild risk assessments and just-in-time remote interventions that may be well-suited to fill a current gap in smoking cessation treatment. These tools offer the potential to objectively monitor a person's dynamically fluctuating state of risk of smoking and need for intervention based on internal states (e.g., stress, craving) and contextual triggers (e.g., location, time of day). By passively monitoring internal and external risks, digital technologies afford the potential to push out intervention when need and receptivity are optimal, rather than requiring that a person have insight into their risk dynamics and pull for support when it is needed to prevent relapse.

At the same time as digital technologies enable just-in-time adaptive intervention, their availability raises new questions about how to optimally configure interventions. For example, the new reality that it is feasible to deliver intervention at any time of day or night creates a need

to curtail burden vigilantly, lest participants become overwhelmed and disengage from the intervention. Accordingly, note that instead of delivering stress management intervention regularly in 30-50 minute in-person sessions or their 10 to 15-minute telehealth equivalent, Sense2Stop substitutes a very light touch format: the opening of an app on a smartphone that prompts the participant to engage in stress management exercises.

With digital clinical trials becoming a reality,⁶⁴ it behooves us not to re-enact prior practices that impeded maximizing the efficiency and effectiveness of behavioral clinical trials. One prior error, pinpointed by the multiphase optimization framework (MOST), involves making arbitrary assumptions about which treatment components should be included in a bundled treatment package, and the temporal and contextual circumstances under which that intervention should be delivered.⁶⁵ By analogy, although it may seem intuitive to assume that the optimal time to deliver stress management intervention is when the individual is experiencing stress, an alternative plausible hypothesis is that stress causes cognitive interference that diminishes receptivity to intervention. If so, it would be optimal to deliver intervention when the person is not stressed and is able to practice stress regulation exercises without distraction by stress-related cognitive interference. The important point is that we do not know. Hence, we require experimentation via a micro-randomized trial to learn the decision rule that will optimize how the JITAI adapts treatment to the patient's dynamically fluctuating states.

The long-term goal of this research is to develop a JITAI to prevent smoking relapse after a quit attempt by helping participants regulate stress. To progress toward that goal, after the Sense2Stop study, it may prove necessary to conduct a larger scale MRT with improved sensor technology. That MRT would be fully powered to test hypotheses about whether and under what conditions performing a stress-regulation exercise reduces the likelihood of a stress episode

occurring shortly after intervention and also reduces the likelihood of a smoking episode occurring during the same interval. Results from this full-scale MRT may inform an optimized JITAI whose efficacy for long-term smoking relapse prevention will be evaluated relative to a suitable comparison condition (e.g., a static gold standard smoking cessation treatment or treatment as usual) via a randomized controlled trial (RCT).

The micro-randomized trial (MRT), which Sense2Stop illustrates, is gateway to an extension of personalized behavioral medicine that goes beyond the treatment matching that is feasible based on knowing a participant's baseline characteristics. The ability of digital tools to randomize minutes and to adapt treatment in real time recognizes the reality that not all health vulnerabilities are static; some fluctuate as episodic states.⁶⁶ Dynamic interventions that adapt in response to individual changes in need and context illustrate the type of contextual tailoring and personalization that will become feasible in the new era of digital sensing and digital clinical trials.⁶⁴

Despite their promise, mobile technologies and real-time analytics are still young, and their suitability to facilitate smoking cessation has yet to be thoroughly studied. Consequently, the Sense2Stop trial addresses unanswered research questions regarding the feasibility as well as the optimal timing of a stress management JITAI to prevent relapse in recently quit smokers. We anticipate that findings from the Sense2Stop study will have important limitations that will warrant consideration. Primarily, the study technologies and statistical methods are cutting edge, and so unforeseen challenges may occur. Pilot work from our group and others reveals a risk for unavoidable missing data resulting from sensor detachment, dropped wifi connection, hardware malfunction, privacy filtering, and sensor non-wear. Additionally, it is possible that wearing a sensor suite could be burdensome or in some way unduly influence the participant's behavior.

Part of our contribution to the field of digital clinical trials will be to characterize these complications and to evaluate methods to compensate for them in future research (e.g., statistical models of missing data). Furthermore, although the machine learned models used here to detect stress and smoking have been extensively validated in various samples of smokers in both lab and field studies, machine learned models are constantly evolving, and validity can vary slightly across samples. Finally, the planned sample size was driven by accuracy of these machine learned detection algorithms. Since binary and trichotomous proximal outcomes require additional consideration in MRT sample size calculations, it is possible that the study is not fully powered for its aims.

Despite these limitations, the Sense2Stop study will address a critical gap in smoking cessation research by using a micro-randomized trial design to optimize a just-in-time adaptive intervention for smoking relapse prevention. JITAIs may be particularly well-suited to address dynamic, time- and context-bound episodic influences that trigger the occurrence of health risk behaviors. Altogether, this study will lay a foundation for future mHealth interventions and digital trials that incorporate real-time sensing and machine-learned analytics to facilitate dynamic, personalized mobile intervention.

ACKNOWLEDGEMENTS

The authors express appreciation to S.M. Hossain, T. Hnat, N. Salaheen and S. Chatterjee for developing study software; D. Wetter for assessment advice; K. Witkiewitz, I. Yovel and M. Sharmin for input about stress intervention; and S. Samiei, E. Daly, G. Ledford, M. DeZelar, S. Cukier, and S. Hoffman for implementing the human studies protocol. This research is supported by NIH grant U54EB020404 Center of Excellence for Mobile Sensor Data-to-Knowledge (MD2K) and also by NIH grants R01DK108678 and U01CA229437.

References

1. Smoking and Tobacco Fast Facts. CDC Fact Sheet Web site. http://www.cdc.gov/tobacco/data_statistics/fact_sheets/fast_facts/. Published 2019. Accessed 2021.
2. Jamal A, Phillips E, Gentzke AS, et al. Current Cigarette Smoking Among Adults - United States, 2016. *MMWR Morb Mortal Wkly Rep.* 2018;67(2):53-59.
3. Creamer MR, Wang TW, Babb S, et al. Tobacco Product Use and Cessation Indicators Among Adults - United States, 2018. *MMWR Morb Mortal Wkly Rep.* 2019;68(45):1013-1019.
4. Hughes JR, Keely J, Naud S. Shape of the relapse curve and long-term abstinence among untreated smokers. *Addiction.* 2004;99(1):29-38.
5. Baer JS, Kamarck T, Lichtenstein E, Ransom CC. Prediction of smoking relapse: analyses of temptations and transgressions after initial cessation. *J Consult Clin Psychol.* 1989;57(5):623-627.
6. Marlatt G, Gordon J. *Relapse prevention: Maintenance strategies in the treatment of addictive behaviors.* New York: Guilford Press; 1985.
7. Kassel JD, Stroud LR, Paronis CA. Smoking, stress, and negative affect: correlation, causation, and context across stages of smoking. *Psychol Bull.* 2003;129(2):270-304.
8. Khantzian EJ. The self-medication hypothesis of substance use disorders: a reconsideration and recent applications. *Harv Rev Psychiatry.* 1997;4(5):231-244.
9. Shiffman S, Wills, TA. *Coping and substance use.* New York, NY: Academic Press; 1985.
10. Yong HH, Borland R, Cooper J, Cummings KM. Postquitting experiences and expectations of adult smokers and their association with subsequent relapse: findings from the International Tobacco Control (ITC) Four Country Survey. *Nicotine & tobacco research : official journal of the Society for Research on Nicotine and Tobacco.* 2010;12 Suppl:S12-19.
11. McKee SA, Sinha R, Weinberger AH, et al. Stress decreases the ability to resist smoking and potentiates smoking intensity and reward. *J Psychopharmacol.* 2011;25(4):490-502.
12. Slopen N, Kontos EZ, Ryff CD, Ayanian JZ, Albert MA, Williams DR. Psychosocial stress and cigarette smoking persistence, cessation, and relapse over 9-10 years: a prospective study of middle-aged adults in the United States. *Cancer Causes Control.* 2013;24(10):1849-1863.
13. Shiffman S. Dynamic influences on smoking relapse process. *Journal of personality.* 2005 Dec;73(6):1715-48.
14. Livingstone-Banks J, Norris E, Hartmann-Boyce J, et al. Relapse prevention interventions for smoking cessation. *Cochrane Database Syst Rev.* 2019;2019(10).
15. Schell LK, Monsef I, Wöckel A, Skoetz N. Mindfulness-based stress reduction for women diagnosed with breast cancer. *Cochrane Database Syst Rev.* 2019;3:CD011518.
16. Tetrick L, Winslow C. Workplace Stress Management Interventions and Health Promotion. *Annual Review of Organizational Psychology and Organizational Behavior.* 2015;2(1):583-603.

17. Sarker H, Hovsepian K, Chatterjee S, et al. From Markers to Intervention: The Case of Just-in-Time Stress Intervention. In: Rehg J, Murphy S, Kumar S, eds. *Mobile Health*. Springer, Cham; 2017:411-433.
18. Saleheen N, Ali AA, Hossain SM, et al. puffMarker: A Multi-Sensor Approach for Pinpointing the Timing of First Lapse in Smoking Cessation. *Proc ACM Int Conf Ubiquitous Comput*. 2015;2015:999-1010.
19. Nahum-Shani I, Smith SN, Spring BJ, et al. Just-in-Time Adaptive Interventions (JITAs) in Mobile Health: Key Components and Design Principles for Ongoing Health Behavior Support. *Ann Behav Med*. 2018;52(6):446-462.
20. Liston C, McEwen BS, Casey BJ. Psychosocial stress reversibly disrupts prefrontal processing and attentional control. *Proceedings of the National Academy of Sciences of the United States of America*. 2009;106(3):912-917.
21. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. *Br J Addict*. 1991;86(9):1119-1127.
22. Sirota AD, Rohsenow DJ, Dolan SL, Martin RA, Kahler CW. Intolerance for discomfort among smokers: comparison of smoking-specific and non-specific measures to smoking history and patterns. *Addictive behaviors*. 2013;38(3):1782-1787.
23. Sirota AD, Rohsenow DJ, Mackinnon SV, et al. Intolerance for Smoking Abstinence Questionnaire: psychometric properties and relationship to tobacco dependence and abstinence. *Addictive behaviors*. 2010;35(7):686-693.
24. Patton JH, Stanford MS, Barratt ES. Factor structure of the Barratt impulsiveness scale. *Journal of clinical psychology*. 1995;51(6):768-774.
25. Cammann C, Fichman M, Jenkins G, J K. Michigan Organizational Assessment Questionnaire. In: Seashore S, Lawyer E, Mirvis P, Cammann C, eds. *Assessing Organizational Change: A Guide to Methods, Measures, and Practices*. New York, NY: Wiley-Interscience; 1983:71-138.
26. Vinokur A, Pierce P, Buck C. Work-family conflicts of women in the Air Force: Their influence on mental health and functioning. *Journal of Organizational Behavior*. 1999;20:865-878.
27. Radloff L. The CES-D scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement*. 1977;1:385-401.
28. Eaton W, Muntaner C, Smith C, Tien A, M Y. Center for Epidemiologic Studies Depression Scale: Review and revision (CESD and CESD-R). In: Maruish M, ed. *The Use of Psychological Testing for Treatment Planning and Outcomes Assessment*. 3rd ed. Mahwah, NJ: Lawrence Erlbaum; 2004:363-377.
29. Fydrich T, Dowdall D, Chambless D. Reliability and Validity of the Beck Anxiety Inventory. *Journal of Anxiety Disorders*. 1992;6:55-61.
30. Cohen S, R M, Kamarck T, Hoberman H. Measuring the functional components of social support. In: Sarason I, Sarason B, eds. *Social support: theory, research, and applications*. The Hague, Holland: Martinus Nijhoff; 1985:73-94.
31. Merz EL, Roesch SC, Malcarne VL, et al. Validation of interpersonal support evaluation list-12 (ISEL-12) scores among English- and Spanish-speaking Hispanics/Latinos from the HCHS/SOL Sociocultural Ancillary Study. *Psychol Assess*. 2014;26(2):384-394.

32. Hossain SM, Hnat T, Saleheen N, et al. mCerebrum: A Mobile Sensing Software Platform for Development and Validation of Digital Biomarkers and Interventions. *Proc Int Conf Embed Netw Sens Syst.* 2017;2017.
33. U.S. Preventive Services Task Force. Final Update Summary: Tobacco Smoking Cessation in Adults, Including Pregnant Women: Behavioral and Pharmacotherapy Interventions, September 2015;
34. USDHHS. Smoking Cessation: A Report of the Surgeon General. USDHHS Public Health Service. Rockville, MD, 2020.
35. Spring B. *Smoking Cessation with Weight Gain Prevention: Workbook.* New York: Oxford University Press; 2008.
36. Doran G. There's a S.M.A.R.T. Way to Write Management's Goals and Objectives. *Management Review.* 1981;70(11):35-36.
37. Ertin E, Stohs N, Kumar S, Raji A, al'Absi M, Shah S. AutoSense: unobtrusively wearable sensor suite for inferring the onset, causality, and consequences of stress in the field. *Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems.* 2011:274-287.
38. Hovsepian K, al'Absi M, Ertin E, Kamarck T, Nakajima M, Kumar S. cStress: Towards a Gold Standard for Continuous Stress Assessment in the Mobile Environment. *Proceedings of the ACM International Conference on Ubiquitous Computing UbiComp (Conference).* 2015;2015:493-504.
39. Kennedy AP, Epstein DH, Jobes ML, et al. Continuous in-the-field measurement of heart rate: Correlates of drug use, craving, stress, and mood in polydrug users. *Drug Alcohol Depend.* 2015;151:159-166.
40. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annual review of clinical psychology.* 2008;4:1-32.
41. J.R. Hughes, D. Hatsukami Signs and symptoms of tobacco withdrawal *Arch. Gen. Psychiatry,* 43 (1986), pp. 289-294
42. Ivan Berlin, Edward G. Singleton, Stephen J. Heishman, Predicting smoking relapse with a multidimensional versus a single-item tobacco craving measure, *Drug and Alcohol Dependence,* Volume 132, Issue 3, 2013, Pages 513-520,
43. Chatterjee, Soujanya, et al. "mCrave: continuous estimation of craving during smoking cessation." *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing.* 2016.
44. A global measure of perceived stress. Cohen S, Kamarck T, Mermelstein R *J Health Soc Behav.* 1983 Dec; 24(4):385-96.
45. Yovel I, Mor N, Shkarov H. Examination of the core cognitive components of cognitive behavioral therapy and acceptance and commitment therapy: an analogue investigation. *Behav Ther.* 2014;45(4):482-494.
46. Witkiewitz K, Desai SA, Bowen S, Leigh BC, Kirouac M, Larimer ME. Development and evaluation of a mobile intervention for heavy drinking and smoking among college students. *Psychology of addictive behaviors : journal of the Society of Psychologists in Addictive Behaviors.* 2014;28(3):639-650.
47. Beck J. *Cognitive behavioral therapy: Basics and beyond.* 3rd ed. New York NY: Guilford Press; 2013.
48. Musthag M, Raji A, Ganesan D, Kumar S, Shiffman S. Exploring micro-incentive strategies for participant compensation in high-burden studies. *Proceedings of the*

- 13th international conference on Ubiquitous computing (ACM UbiComp)*. 2011:435-444.
49. Sarker H, Tyburski M, Rahman M, et al. Finding significant stress episodes in a discontinuous time series of rapidly varying mobile sensor data. *Proceedings of the 2016 SIGCHI conference on human factors in computing systems (ACM CHI)*. 2016:4489-4501.
 50. Hossain S, Ali A, Rahman M, Ertin E, Epstein D, Kennedy A, Preston K, Umbricht A, Chen Y, Kumar S. Identifying drug (cocaine) intake events from acute physiological response in the presence of free-living physical activity. *ACM IPSN*. 2014:71–82.
 51. Ni K, Ramanathan N, Chehade M, et al. Sensor network data fault types. *ACM Transactions on Sensor Networks*. 2009;5(3):25.
 52. Rahman M, Bari R, Ali AA, et al. Are We There Yet? Feasibility of Continuous Stress Assessment via Wireless Physiological Sensors. *ACM BCB*. 2014;2014:479-488.
 53. Nakajima M, Lemieux AM, Fiecas M, et al. Using novel mobile sensors to assess stress and smoking lapse. *Int J Psychophysiol*. 2020;158:411-418.
 54. Liao P, Dempsey W, Sarker H, et al. Just-in-Time but Not Too Much: Determining Treatment Timing in Mobile Health. *Proc ACM Interact Mob Wearable Ubiquitous Technol*. 2018;2(4).
 55. Kenney W, Wilmore J, Costill D. *Physiology of Sports and Exercise*. Champaign, IL: Human Kinetics; 2020.
 56. Wilhelm FH, Grossman P. Emotions beyond the laboratory: theoretical fundamentals, study design, and analytic strategies for advanced ambulatory assessment. *Biol Psychol*. 2010;84(3):552-569.
 57. Murray DM, O'Connell CM, Schmid LA, Perry CL. The validity of smoking self-reports by adolescents: a reexamination of the bogus pipeline procedure. *Addictive behaviors*. 1987;12(1):7-15.
 58. Luepker RV, Pallonen UE, Murray DM, Pirie PL. Validity of telephone surveys in assessing cigarette smoking in young adults. *American journal of public health*. 1989;79(2):202-204.
 59. Qian T, Yoo H, Klasnia P, Almirall D, Murphy S. Estimating Time-Varying Causal Excursion Effects in Mobile Health with Binary Outcomes. *Biometrika*. In Press.
 60. al'Absi M, Hatsukami D, Davis GL, Wittmers LE. Prospective examination of effects of smoking abstinence on cortisol and withdrawal symptoms as predictors of early smoking relapse. *Drug Alcohol Depend*. 2004;73(3):267-278.
 61. Dempsey W, Liao P, Kumar S, Murphy S. The Stratified Micro-Randomized Trial Design: SAmples Size Considerations for Testing Nested Causal Effects of Time-Varying Treatments. *Annals of Applied Statistics*. 2019;0(0):1-33.
 62. Goodchild M, Nargis N, Tursan d'Espaignet E. Global economic cost of smoking-attributable diseases. *Tob Control*. 2018;27(1):58-64.
 63. Shiffman S, Waters AJ. Negative affect and smoking lapses: a prospective analysis. *J Consult Clin Psychol*. 2004;72(2):192-201.
 64. Inan OT, Tenaerts P, Prindiville SA, et al. Digitizing clinical trials. *NPJ Digit Med*. 2020;3:101.

65. Collins L. *Optimization of Behavioral, Biobehavioral, and Biomedical Interventions: The Multiphase Optimization Strategy (MOST)*. New York: Springer; 2018.
66. Zubin J, Spring B. Vulnerability--a new view of schizophrenia. *J Abnorm Psychol.* 1977;86(2):103-126.