

OnTrack: development and feasibility of a smartphone app designed to predict and prevent dietary lapses

Evan M. Forman,¹ Stephanie P. Goldstein,¹ Fengqing Zhang,¹ Brittney C. Evans,¹ Stephanie M. Manasse,¹ Meghan L. Butryn,¹ Adrienne S. Juarascio,¹ Pramod Abichandani,¹ Gerald J. Martin,¹ Gary D. Foster^{2,3}

Abstract

Given that the overarching goal of weight loss programs is to remain adherent to a dietary prescription, specific moments of nonadherence known as “dietary lapses” can threaten weight control via the excess energy intake they represent and by provoking future lapses. Just-in-time adaptive interventions could be particularly useful in preventing dietary lapses because they use real-time data to generate interventions that are tailored and delivered at a moment computed to be of high risk for a lapse. To this end, we developed a smartphone application (app) called *OnTrack* that utilizes machine learning to predict dietary lapses and deliver a targeted intervention designed to prevent the lapse from occurring. This study evaluated the feasibility, acceptability, and preliminary effectiveness of *OnTrack* among weight loss program participants. An open trial was conducted to investigate subjective satisfaction, objective usage, algorithm performance, and changes in lapse frequency and weight loss among individuals ($N = 43$; 86% female; body mass index = 35.6 kg/m²) attempting to follow a structured online weight management plan for 8 weeks. Participants were adherent with app prompts to submit data, engaged with interventions, and reported high levels of satisfaction. Over the course of the study, participants averaged a 3.13% weight loss and experienced a reduction in unplanned lapses. *OnTrack*, the first just-in-time adaptive intervention for dietary lapses was shown to be feasible and acceptable, and *OnTrack* users experienced weight loss and lapse reduction over the study period. These data provide the basis for further development and evaluation.

Keyword

Weight, Lapses, Smartphone app, Diet, Digital

An estimated 68% of Americans are overweight or obese [1] and 49% of those who are overweight/obese are currently attempting to follow a lower-calorie diet in order to lose weight [2], with many others attempting to prevent weight gain [3]. Substantial weight loss (and weight loss maintenance), while difficult to achieve, occurs when people consistently adhere to intake and activity patterns designed to facilitate weight loss or weight loss maintenance [4–6]. However, specific moments of nonadherence (i.e., “lapses”) threaten weight control because each lapse represents episodes of excess calorie intake that may each be relatively small, but that can accumulate such that overall energy balance is positive. Each individual lapse also increases the likelihood

Implications

Practice: Smartphone apps are particularly useful for predicting and preventing individual health behaviors and apps, such as *OnTrack*, could be utilized to facilitate greater success during self-directed weight loss attempts.

Policy: Given the pervasiveness of obesity and associated health concerns, policymakers may wish to continue devoting resources to innovative methods for predicting and preventing problematic eating behaviors.

Research: Additional research is necessary to confirm the efficacy of *OnTrack* as a weight loss tool and further refine the process of real-time prediction and prevention of lapse behavior.

of future lapses as lapses are associated with feelings of hopelessness, which may lead to a complete abandonment of weight control efforts [7–10].

Many programs exist to facilitate adherence to a dietary prescription, ranging from smartphone applications (apps) to self-directed attempts to commercial weight management programs to gold-standard behavioral weight loss treatments [11]. However, these interventions: (a) do not adequately target the prevention of dietary lapses (e.g., most smartphone apps simply provide methods of tracking eating, physical activity and weight); (b) are often delivered out-of-context (e.g., in a clinician’s office); and (c) are delivered days, weeks, or even months before (or after) the challenge they address (e.g., temptations from a social event) [12–15]. Moreover, these interventions tend to deliver the same core intervention components to everyone. For example, gold-standard behavioral weight loss treatment delivers the same 25 or so treatment sessions to all participants. Virtually none of the most popular weight loss apps available use technology to tailor their interventions to the user, other than simple adaptations of

¹Center for Weight, Eating, and Lifestyle Science (WELL Center), Drexel University, Philadelphia, PA, USA

²Weight Watchers International, New York, NY, USA

³Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA

Correspondence to: EM Forman, emf27@drexel.edu

Cite this as: *TBM* 2019;9:236–245
doi: 10.1093/tbm/iby016

© Society of Behavioral Medicine 2018. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.

nutritional/physical activity recommendations and reinforcement and encouragement messages and gamifications [15, 16]. Just-in-time, adaptive interventions (JITAI) are a potential solution to many of the challenges of traditional weight management interventions. JITAI use real-time data to provide interventions that are tailored and delivered in a moment of need. Behavior change is more likely the closer the intervention is to the time and place of the target activity [17], suggesting that JITAI could be more effective than traditional health behavior change interventions.

JITAI, which have varied in intensity from simply providing in-the-moment feedback to facilitating in-the-moment employment of a skill, have been effective for substance abuse, physical activity [18], sedentary behavior [19], smoking [20], macronutrient adjustments [21, 22], and decreasing fat intake [23]. One randomized controlled trial evaluated a JITAI based on self-monitoring entries using a personal digital assistant that provided ongoing, adaptive dietary feedback throughout a 2-year behavioral weight loss program [24]. Results revealed that personal digital assistant-delivered tailored feedback, compared with electronic and paper versions of self-monitoring without feedback, produced a significant reduction in energy intake but no differences in weight losses at 24 months [24, 25]. A more recent trial found weight losses to be equivalent among participants using a popular weight loss app (LoseIt!) compared with those using LoseIt! plus a companion research app that provided real-time tailored feedback based on food entries [26]. Taken together, these studies indicate that real-time, tailored interventions are feasible; however, providing feedback on dietary behaviors after they occur may not influence weight loss.

An advanced form of JITAI can use machine learning approaches to analyze historical and real-time data to predict problematic behaviors before they occur. Machine learning algorithms can be a model for a general population and specific individuals, and improve in their performance as data accumulate on predictor and outcome variables.

Specific predictors of lapse behavior have been identified through ecological momentary assessment (EMA) studies of individuals attempting to follow a dietary intake for weight loss. Lapses are more likely during specific environmental (e.g., watching television, the presence of food cues in the environment) and internal contexts (e.g., hunger, negative or positive mood, stress, several, boredom and feelings of deprivation) [7, 8, 10, 27–33]. A number of other variables (including sleep deprivation, alcohol intake, and cognitive load) have also been indirectly associated with dietary nonadherence [34–37]. As such, it is theoretically possible that a JITAI could predict the likelihood of lapsing based on known triggers, warn the person when the risk level is high,

and deliver a targeted intervention aimed at preventing the lapse from occurring.

To date, one app has been developed to predict eating behaviors. SlipBuddy is a smartphone app-based JITAI designed to predict and prevent episodes of general overeating (i.e., a subjective sense that one has eaten more than one should have) among individuals with overweight hoping to lose weight [38]. The machine learning algorithm embedded in the app predicted such episodes with a 71% accuracy, delivered brief tailored interventions as needed and achieved a 1.05% 1-month weight loss. One limitation of this JITAI is that it required manually building individual algorithms for each participant by training and testing 16 separate models. As such, participants were unable to begin receiving interventions for 3 months, and SlipBuddy's scalability is limited due to this resource intensive design [38].

To target lapses from a weight control plan using methods that are more disseminable and scalable, we developed *OnTrack*. The app works in conjunction with an existing behavioral weight loss program such that the weight loss program recommends a dietary approach and *OnTrack* assists users with adhering to that approach. *OnTrack* utilizes a machine learning algorithm to automatically build models of lapse behavior, predict lapses before they occur, and delivers microinterventions when lapse risk is high. To develop *OnTrack*, we identified a host of predictive variables for lapsing (Table 1), used smartphone-based EMA to collect these data among individuals with overweight following a dietary plan for weight loss, and identified a machine learning model capable of accurately predicting dietary lapses [39]. The model now informs app-delivered just-in-time interventions. This study evaluated the feasibility, acceptability and preliminary effectiveness of *OnTrack* among a group of overweight individuals assigned to a structured commercial weight loss plan for 8 weeks. This open trial investigated subjective satisfaction, objective usage, algorithm performance, and changes in lapse frequency and weight loss. Additionally, we evaluated whether lapse reductions correlated to weight loss.

METHODS

Participants

Participants were 44 adults (18–65 years old) with overweight or obesity (body mass index 25–50 kg/m²) who owned an iPhone. Individuals were excluded if they were enrolled in another structured weight loss program, were pregnant or planning to become pregnant, reported disordered eating symptoms, had a medical condition that contraindicated weight loss, had begun or changed dosage of a weight or appetite-affecting medication in the past 3 months, or had a history of bariatric surgery. One participant dropped out of the study prematurely due to

Table 1 Variables assessed by *OnTrack*, how often and when assessment occurred, and response choices displayed to participant

Variable name	Question frequency	Time of day rules	Response type
Affect	~3–4 per day	All available times	5-point Likert Scale
Boredom	~3–4 per day	All available times	Yes/No
Hunger	~3–4 per day	All available times	Yes/No
Cravings	~3–4 per day	All available times	
Tiredness	~3–4 per day	All available times	Yes/No
Unhealthy food availability	~3–4 per day	All available times	Yes/No
Temptations	~3–4 per day	All available times	Yes/No
Missed meals/Snacks	~3–4 per day	All available times	Yes/No/Unsure
Self-efficacy (confidence)	~1–2 per day	No night time	5-point Likert Scale
Motivation	~1–2 per day	All available times	5-point Likert Scale
Socializing (with or without food present)	~1–2 per day	Afternoons and evenings	Yes with food/Yes without food/No
Watching TV	~1–2 per day	Afternoons and evenings	Yes/No
Negative interpersonal interactions	~1–2 per day	All available times	Yes/No
Healthy food presence	~1–2 per day	All available times	Yes/No
Cognitive load	~1–2 per day	All available times	5-point Likert Scale
Food cues (advertisements)	~1–2 per day	All available times	Yes/No
Hours of sleep	Once	Morning	Continuous
Exercise	Once	Evenings	Yes/No
Alcohol consumption	Once	Afternoons and evenings	Yes/No
Planning food intake	Once	Morning and afternoon	Yes/No
Time of day	Continuous Measurement	All available times	—

Adapted from Goldstein et al. (2017).

unforeseen circumstances (e.g., leaving the country), and so the data below reflect 43 participants.

Design

This study utilized an open trial design. All participants were assigned to the Weight Watchers (WW) online weight loss program (more information below) and to *OnTrack* for a period of 8 weeks. Assessments were conducted at pre- and postintervention.

Procedures

Participants were recruited through print and online advertisements, and screened via a brief screening call. Participants recorded dietary intake using the WW mobile app, and entered data related to lapses and lapse triggers in *OnTrack*. (Apps were provided at no cost.) Assessments took place in person.

Intervention descriptions

Weight watchers intervention

All participants received access to WW Online, accessible via website or mobile app. The dietary component of WW is the SmartPoints plan, which assigns each food and beverage a SmartPoints value based on protein, fat, carbohydrates, and fiber. The plan is designed to achieve weight control by encouraging consumption of satiating foods low in

energy density, such as fruit, vegetables, and low-fat proteins. As part of the dietary plan, WW assigns participants a SmartPoints budget, personalized based on their age, gender, weight, and height. As part of their budget, each participant was given a Daily SmartPoints goal (range from 30 to 93), and a Weekly SmartPoints bank (range from 14 to 42).

In order to allow participants to standardize the definition of a lapse (see below), we adapted the daily SmartPoints goals to specific meal and snack point targets throughout the day. In particular, participants were allotted 15% of their daily SmartPoints for breakfast, 25% of their Daily SmartPoints for lunch, 40% of their Daily SmartPoints for dinner, and two snacks of 10%. They were instructed to closely adhere to these goals with minimal adjustments and were allowed to adjust point goals during the follow-up appointment (2–5 days after the baseline). Participants were instructed to use their Weekly SmartPoints as desired. Participants were instructed to track their daily food intake via the WW app or website.

OnTrack intervention

OnTrack assessed potential lapse triggers by prompting users to complete semi-random surveys six times a day during waking hours. The app also allowed users to enter data at will when a lapse occurred. (See [Figure 1](#) for *OnTrack*'s main

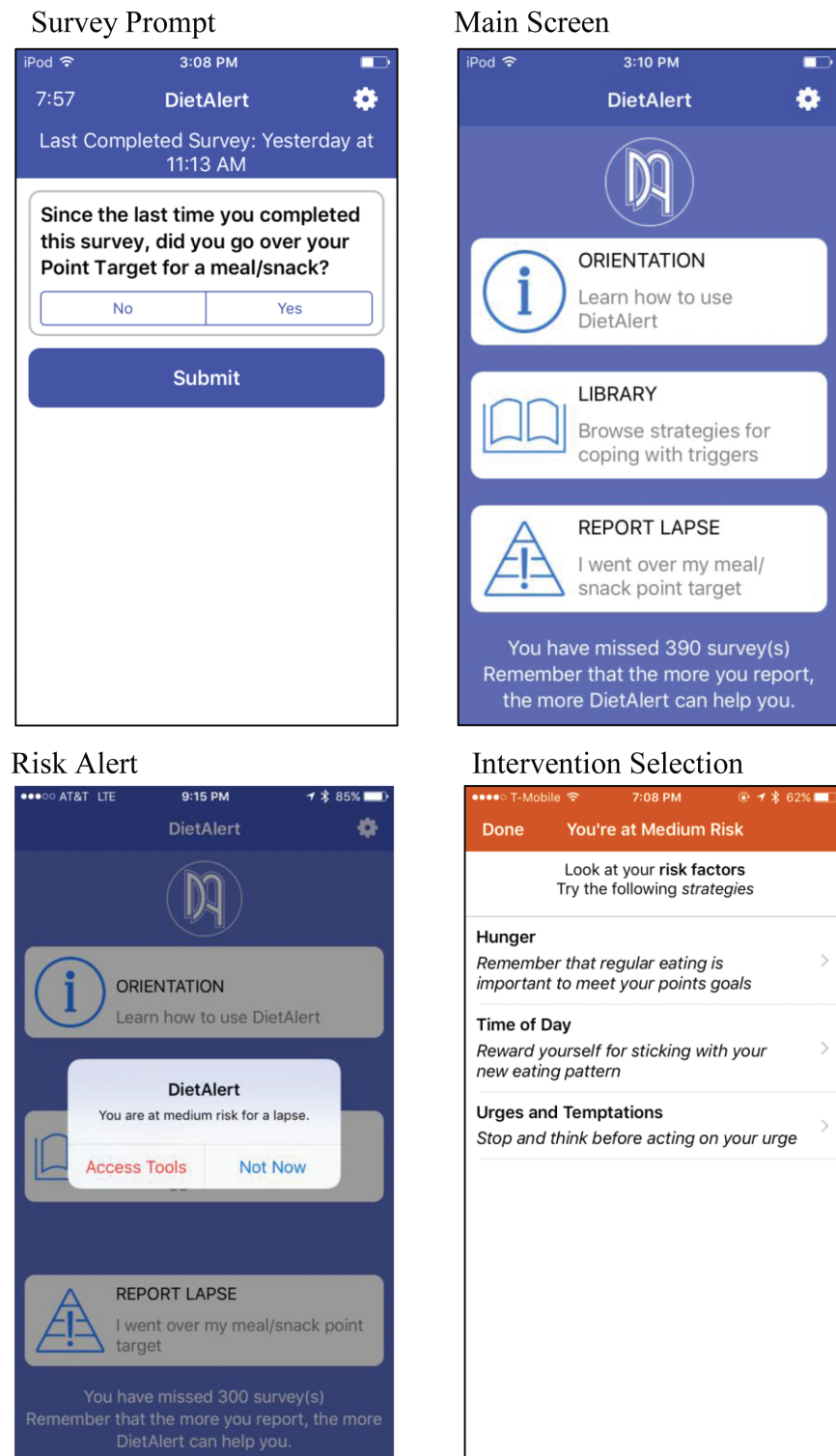


Fig 1 | Smartphone app Screenshots (as seen by participants). The app was previously called *DietAlert* and was changed to *OnTrack* based on participant feedback.

screen and an example of a data entry prompt.) Each time new information was entered into the app, the machine learning algorithm determined associations between risk factors (listed in Table 1) and lapse behavior (going over the assigned point target for a meal or snack). *OnTrack* employs a cost-sensitive ensemble model utilizing logit boost

[40], bagging [41], random subspace [42], random forest [43], and Bayes net [44] (for detail, see [39]). Participants provided data for 2 weeks before lapse risk alerts were enabled. *OnTrack*'s algorithm was created from group-level data (collected during the algorithm development phase) and this initial model was continuously updated with all available

individual data in real-time to personalize future dietary lapse predictions and tailor interventions (for detail, see [45]).

When risk for lapse was identified, an alert was issued communicating that the participant was at risk and displaying up to three factor(s) substantially contributing to level of risk (extracted from the machine learning algorithm). In response to each risk factor, participants saw one-sentence interventions describing possible strategies to cope with each specific risk factor, which they could select to view a more detailed intervention (approximately 1–2 app screens). The interventions were selected from a bank of 157 (7–10 for each risk factor), which were developed using an empirically based health behavior change taxonomy [46], and typically contained some form of user interaction such as write-in text or checkboxes. Further details on intervention development and deployment of the app (originally titled DietAlert) have been previously described [39], and screenshots of the risk alert and intervention system are displayed in Figure 1.

Measures

Weight

Participant weights (to the nearest 0.01 lb) were obtained by a calibrated scale at all assessments, and a stadiometer was used to measure participants' height to the nearest 1.0 cm at baseline. Participants were asked to wear light clothing and remove shoes for assessment.

Dietary lapses

Participants used *OnTrack* to record lapses from the WW dietary plan throughout the duration of the study. A dietary lapse was defined as, “any instance in which you exceeded your SmartPoints goal for a meal or snack.” In this sense, we included both the conventional definition of an unintended slip from one's planned meal/snack as well as intentional overages that the participant planned to remedy. Participants were asked at each semi-random survey prompt if they had experienced a dietary lapse since the previous survey. Participants were also encouraged to enter lapses into *OnTrack* as soon as they occurred. At each lapse entry, participants were asked to report the exact time and date of the lapse. When participants reported a lapse, they were also asked to describe if the lapse fell into the following categories: (1) consumed a high-point food item that they had intended to avoid, (2) consumed more of a food than planned, (3) ate at a time they had not intended, (4) did not know the point values of a food, and (5) planned to go over the meal/snack target. Categories were not mutually exclusive. *OnTrack* predicted all lapses (regardless of type); however, we distinguished between planned (i.e., where the participant selected category 5 above) and unplanned

lapses when evaluating the app's effectiveness over time, as these categories could represent clinically distinct behaviors (implications are discussed below) [32, 47, 48].

Lapse triggers

Each *OnTrack* survey prompt contained eight questions that assessed a subset of internal and external lapse triggers. Given that there were 20 possible self-reported triggers for lapse, not all triggers were assessed at each prompt to reduce participant burden [49]. See Table 1 for a complete list of lapse triggers assessed and frequency of assessment. Lapse triggers were measured on 5-point Likert scales or dichotomous (yes/no) questions.

Utilization

OnTrack automatically recorded the number of survey prompts delivered and completed, and the number of risk alerts received and viewed.

Perceived helpfulness

Once risk alerts were enabled, users received an end-of-day survey question about that day's risk alert utility and accuracy. Users were also instructed to rate the helpfulness of each intervention delivered on a provided 1- to 5-star scale.

Technology acceptance

At the conclusion of the study, participants completed the Technology Acceptance Model Scales [TAMS] [50] to assess satisfaction, perceived usefulness and usability, and technical problems with the *OnTrack* app. The TAMS consists of a Likert rating scale (from 1 to 7), and two text-entry questions for participant positive and constructive feedback. The TAMS has demonstrated adequate reliability and validity in measurement [50].

RESULTS

Participant characteristics

Participants' (86% female; $n_{female} = 37$; $n_{male} = 6$) average age was 50.98 years ($SD = 12.72$ years) and body mass index (BMI) was 35.6 kg/m^2 ($SD = 5.88 \text{ kg/m}^2$). Race/ethnicity was self-identified as follows: 74.4% White, 18.3% Black, 2.3% Latino/Latina, and 4.6% as other. Approximately 80% of participants had used an app to track their eating and/or physical activity previously and the majority of participants (72.1%) self-reported using apps on his or her phone more than once per day at the time of study enrollment.

OnTrack app use

To quantify *OnTrack* app usage, percent of surveys completed (“adherence”) and opened risk alerts over time were calculated. Participants received six surveys each day. Participants were adherent

with daily prompts to enter information into the app, responding to an average of 85.1% of the 336 daily surveys ($SD = 14.6\%$, range = 44.6–98.8%). However, a repeated measures ANOVA revealed that adherence to surveys declined over time, Wilk's $\Lambda = 0.19$, $F(7,36) = 20.82$, $p < .001$. Average adherence with surveys during the first week of study participation was 92.14% ($SD = 7.13\%$); during the final study week, average compliance dropped to 76.08% ($SD = 2.70\%$).

When participants were identified to be at risk for a lapse by the machine learning algorithm operating within the app, they were provided with a notification called a “risk alert.” Participants received an average of 7.29 risk alerts per week ($SD = 2.40$ risk alerts, range = 2.75–12.00 risk alerts). Participants opened an average of 5.14 risk alerts ($SD = 2.43$ risk alerts, range = 1–11 risk alerts), that is, 70.15% of these alerts, such that they could view the personalized factors placing them at risk and see a list of suggested microinterventions (e.g., one-sentence suggestions to combat a particular risk factor). Of the 21 possible triggers, 29% of intervention triggers were based on time of day, 16.7% based on low motivation, and 10% based on fatigue. The remaining 18 triggers were identified as risk factors less than 10% of the time. A repeated measures ANOVA revealed that viewing alerts decreased over the course of the study period from week 3 ($M = 77.9\%$, $SD = 24.3$) to week 8 ($M = 68.3\%$, $SD = 27.2$), Wilk's $\Lambda = 0.59$, $F(5,21) = 1.90$, $p = .04$. Once an alert was opened, participants chose to open 44.7% of the microinterventions in order to read a “full” intervention (i.e., two to three pages of app text).

Satisfaction and acceptability

To examine satisfaction and acceptability, descriptive statistics were calculated for participants' overall *OnTrack* app ratings at the end of the study as well as satisfaction ratings with app interventions provided throughout the study.

Per responses on the TAMS (possible range from Strongly Disagree (1) to Strongly Agree (7)) at post-treatment, participants indicated that the app was easy to use ($M = 6.14$, $SD = 1.58$) and that they had minimal technical issues ($M = 2.91$ out of 7, $SD = 1.24$). Participants rated the app as moderately useful ($M = 4.64$, $SD = 1.58$) and enjoyable ($M = 4.37$, $SD = 1.62$), with a somewhat positive behavioral intention to use ($M = 4.48$, $SD = 1.86$). Participants who provided ratings on the library interventions (53.5% of the total sample) indicated that the interventions were moderately helpful ($M = 3.94$ out of 5, $SD = 0.69$). Of interventions delivered via risk alert, 81.02% were rated as 3 stars or greater (out of 5 stars). The highest-rated interventions addressed time of day, fatigue, low motivation, and boredom triggers, whereas the lowest-rated addressed watching TV, bad mood, and lack of available healthy foods. Participants reported that the app may be more

acceptable and useful if it were integrated with WW (to increase overall ease of monitoring), was even further tailored to adapt the daily EMA questions to suit each individual's problem areas, was less reliant on wireless connection (to facilitate improved data entry and intervention access), and contained features to block prompts/interventions during certain times of the day (e.g., periods of driving, meetings).

Algorithm accuracy

During the development phase of *OnTrack* (a data-collection-only period), we were able to use ensemble classifiers described above to achieve an algorithm accuracy of 72%, sensitivity of 70%, and specificity of 72% [51]. The algorithm was implemented in this study with the understanding that more individualized data would enhance algorithm performance.

To evaluate algorithm performance during this study, algorithm-generated predictions were compared to actual lapse occurrences (as self-reported by participants). Table 2 illustrates a confusion matrix that compares all lapses that were predicted to those that occurred during weeks 3–8 (when algorithm-generated interventions were triggered). Given that the app was both attempting to predict and intervene on lapses, the metrics of sensitivity of the machine learning algorithm and overall prediction accuracy are insufficient to objectively interpret the algorithm performance. For example, an incorrect lapse prediction could be a false positive, or it could be a case in which the intervention effectively prevented a lapse. Conversely, if there was an “accurate” lapse prediction, the interventions may not have adequately prevented the lapse.

As such, the best representation of model performance is the negative predictive value (e.g., the proportion of true non-lapses among all the predicted non-lapses) because this statistic is not biased to the provision of intervention on lapses. The negative predictive value in this study was 80%, indicating good model performance on this criterion. The majority of participants (72.9%) reported that interventions received for the day were accurate and/or helpful (e.g., “I received alerts that were helpful”, “I received alerts and a portion of them were helpful”, “I received no risk alerts and that seemed right”).

Table 2 Confusion matrix of predictions generated by *OnTrack* and the subsequent lapse occurrences (as self-reported by participants)

	Outcome reported by participant		
	No lapse	Lapse	
Outcome predicted by app	No lapse	4,903	1,228
	Lapse	1,253	324

Lapse frequency and weight change

To evaluate lapses frequency and its relationship to weight change, lapses were categorized into “planned” and “unplanned” depending on participant self-report. Participants averaged approximately 18.26 ($SD = 18.48$) planned lapses and 37.74 ($SD = 28.47$) unplanned lapses across the study period. A repeated measures ANOVA was used to evaluate change in number of self-reported lapses over time when controlling for compliance with completing survey prompts (the primary method for reporting lapses). Results indicated that there was no significant relationship between time and lapses when controlling for compliance for both planned lapses, Wilk’s Lamda = 0.81, $F(7,34) = 1.10$, $p = .38$, and unplanned lapses, Wilk’s Lamda = 0.86, $F(7,34) = 0.79$, $p = .61$. Figure 2 illustrates changes in unplanned and planned lapses by week. Paired samples t -tests were used to directly compare the number of lapses during week 1 to the number to lapses during week 8. Unplanned ($M = 2.09$, $SD = 5.59$), $t(41) = 2.43$, $p = .02$, $d = 0.38$) but not planned lapses ($M = 0.26$, $SD = 2.45$), $t(41) = 0.69$, $p = .29$, $d = 0.10$) decreased significantly over the study period.

With regard to evaluating weight change, we calculated percentage of body weight lost across participants, compared pre- and postintervention BMI, and examined the percentage of individuals who achieved clinically significant weight loss using mobile intervention standards (i.e., 3% or greater) and standards for appreciable health benefits (i.e., 5% or greater) [52, 53]. Final weight data were missing for four participants, as they were unable to attend the final assessment. As such, intent-to-treat analyses were conducted using the last-observation-carried-forward method. Participants demonstrated an average of 3.13% weight loss ($SD = 2.98\%$, range = -9.00% to 7.00%). A paired samples t -test revealed a significant difference between pre- ($M = 35.65$) and postintervention BMI ($M = 34.78$), $t(42) = 3.61$, $p = .001$, $d = 0.55$. 43.6% of participants achieved weight losses of at least 3% and 35.9% of participants achieved weight loss of 5% or greater.

Lastly, the relationship between lapses and weight was examined using Pearson’s correlations. In particular, we were interested in examining behaviors of individuals who were “successful” weight losers, which we defined as those who lost at least 3% of their body weight. This 3% threshold is a benchmark chosen based on the primary intervention being solely app-based, the 8-week period and findings from similar studies [54–57]. An independent samples t -test was used to examine differences in lapse frequency between individuals who met this threshold and those who did not. Pearson’s correlations revealed a nonsignificant association between overall weight loss and unplanned ($r = 0.10$, $p = .52$) as well as planned ($r = -.04$, $p = .79$) lapses. When comparing individuals who did and did not reach the $>3\%$ weight loss threshold, the differences in lapse frequency were small and nonsignificant for planned lapses ($M_{\text{difference}} = 4.19$, $SE = 5.78$), $t(41) = 0.73$, $p = .47$, $d = .22$, but medium-large and significant at a trend level for unplanned lapses ($M_{\text{difference}} = -23.36$, $SE = 12.25$), $t(41) = -1.91$, $p = .06$, $d = .58$.

DISCUSSION

This study is the first to develop a JITAI for preventing dietary lapses among participants with overweight and obesity. While the design did not allow us to draw firm conclusions about the efficacy of *OnTrack*, the app demonstrated the feasibility and acceptability of a JITAI method for preventing dietary lapses. *OnTrack* also demonstrated preliminary effectiveness in reducing unplanned dietary lapses and facilitating weight loss.

Subjectively, participants indicated that *OnTrack* was easy to use and helped them achieve weight control. Of note, their basis of comparison may have been commercial-grade apps, as 80% of participants had previously used commercially available weight loss/health promotion apps. Participants showed excellent adherence with prompts to enter surveys. However, we did observe a modest fall-off in response rates through time, which could suggest that compliance would be problematic over the longer

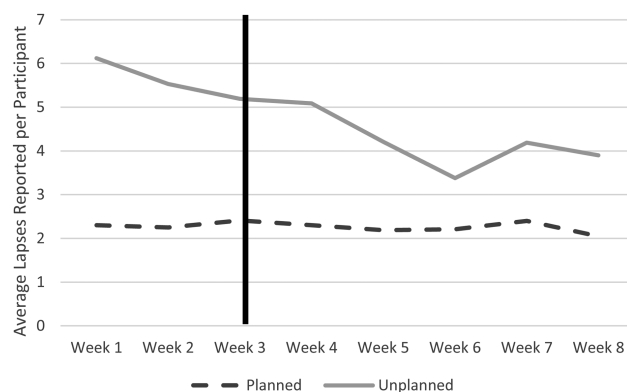


Fig 2 | Average reported lapses by study week. Line represents when risk alerts began.

term. Even at the end of the 8-week study, the rate of survey responding (76.0%) was high and as good or better than reported by other EMA studies even after only 2 weeks [58–60]. Devising a system that facilitates continued compliance over many months is critical for this type of intervention, given that the algorithm uses these data to inform the timing and content of intervention delivery.

One possible motivator of survey compliance may have been that survey responses were used for the purposes of generating interventions, and participants believed these to be helpful. Participants perceived the algorithm-determined lapse risk alerts to come at helpful times and rated the delivered interventions favorably. Objectively measured engagement with the app was high throughout the 8 weeks of the study, with participants opening the majority (70%) of risk alert interventions and frequently choosing to view the most comprehensive version. While this rate decreased somewhat over time, this trend is to be expected as participants became increasingly familiar with the reasons they were lapsing.

Contrary to hypotheses, participants did not experience a decrease in overall dietary lapses over the course of the intervention when controlling for survey compliance. Examining changes from week 1 to week 8 revealed that unplanned lapses decreased over time. This decreasing trend was not evident in planned lapses. The differences between planned and unplanned lapses are not surprising given that *OnTrack* was designed to help prevent overeating episodes triggered by internal and external eating cues (vs. a deliberate decision to eat more on one occasion and less at another). One limitation in interpreting these results is that the observed reductions in lapse reports could be conflated with reductions in app usage over time. Additional studies comparing *OnTrack* to an EMA-only version are warranted to explore these relationships further.

Participants evidenced a 3.13% weight loss over the 2-month study period, with drop-outs conservatively coded as having lost 0% weight. The additive efficacy of *OnTrack* cannot be directly assessed due to the lack of a control group. However, the rate of weight loss observed (1.6%/month) can be informally compared with a meta-analytic average of weight loss from a weight loss app alone (0.6%/month) [61]. A more distant, but perhaps still-useful benchmark, is the rate at which gold-standard in-person behavioral treatment achieves weight loss across the first 6 months (1.1%/month) [62]. Also of note is that 35.9% of *OnTrack* participants experienced a 5% weight loss by posttreatment, which is considerably higher than the proportion reported by a recent study of individuals using the Weight Watchers app [63]. All told, results support the promise of *OnTrack* for enhancing weight loss. Importantly, the variability of weight change in this study was high (range: –7.8 kg to +7.3 kg), which warrants further

investigation of potential moderators of outcomes. The association between lapse frequency and weight loss further bolsters the underlying notion that lapses are a viable target for intervention. The lack of an association between weight loss and *planned* lapses is most likely explained by the fact that these lapses were compensated for as planned, such that the total calorie intake for the day was not in excess. In fact, gold-standard behavioral weight loss programs encourage this type of planning or “banking” as a way of allowing for occasional eating more than typical while still maintaining an energy deficit [64–66]. As such, future iterations of *OnTrack* might focus on a specific lapse type (e.g., unplanned lapses) rather than all lapses that occur from an eating plan.

With regard to the performance of the app-based machine learning algorithm, participants perceived most risk alerts to be accurate and appropriate. However, objective interpretation of the algorithm performance is non-trivial because *OnTrack* was both attempting to predict and intervene on lapses. One option to explore in future iterations of the *OnTrack* risk algorithm is to assume that a lapse would have occurred had the app not intervened. In such a case, one could tag a non-lapse behavior being followed by a lapse risk alert prediction as a lapse. Recoding false negatives into true positives could restore balance to algorithm performance, as well as assist the algorithm in learning the correct trigger–lapse associations. Future research is needed to investigate the clinical and statistical impact of recoding false positives into true positives when a JITAI is aiming to predict and prevent a proximal outcome in real time. Another logical next step is to conduct a microrandomized trial (e.g., each opportunity for intervention is randomized to a condition) to evaluate the effectiveness of lapse prediction and subsequent intervention [67].

Strengths and limitations

Several limitations of the current project must be acknowledged. First, we conducted an open trial, which reduces our confidence that *OnTrack* was responsible for reductions in lapse frequency or for weight loss. Secondly, while we took care to define lapses as objectively as possible, participants likely became better over time with counting points, and it is possible that participant perception of lapses changed over time. Additionally, as is typically the case in weight loss studies, most participants were female; thus, app engagement and outcomes may not generalize to men. Lastly, the intervention continued for only 8 weeks, whereas a successful weight control intervention would presumably have to be engaging and impactful for considerably longer.

Future research should aim to address limitations of this study. First, a randomized controlled trial is a logical next step to determine whether *OnTrack* is the causal factor in the reduction of lapses and weight loss. There are several possibilities for a comparison condition, including a WW-only

condition (thus testing the impact of the full JITAI app) or a WW + EMA condition (thus removing the potential reactivity effects of repeated prompting). In addition, several participants provided qualitative data indicating that many of the repeated questions did not feel relevant to them; future iterations of *OnTrack* could better tailor data inputs, that is, only ask questions that are relevant to the participant. Similarly, *OnTrack* could incorporate and utilize more passive channels of data collection, such as geolocation, which could help increase the accuracy of the algorithm (e.g., provide a risk alert to a participant when he is near an ice cream shop, where he had lapsed before). In addition, more interactive features (e.g., videos, badges, other forms of positive reinforcement) could be incorporated and tested for their correlation to prevent the decrease in engagement over time. Recruiting more male participants would also improve the generalizability of a future study.

Strengths of this study include the development and provision of a machine learning algorithm that was able to predict and intervene on dietary lapses, thus applying modern statistical methods to behavior change research. *OnTrack* used cost-sensitive ensemble decision trees that showed promise during the development phase of the app, achieving 70% sensitivity and 72% specificity in a previous trial [51]. These standards meet established conventions for prediction of human behavior [38, 68]. Moreover, *OnTrack* is an exemplar for the deployment of these algorithms in real time and highlights important methodological hurdles for employing JITAI for lapses. In addition, we paired *OnTrack* with a popular weight loss app, making the findings applicable to more main-stream populations. Lastly, we measured outcomes across modalities including self-reported satisfaction, app compliance and feature usage, lapse frequency, and weight.

CONCLUSIONS

In sum, *OnTrack* achieved its preliminary aim in successfully deploying the first JITAI for dietary lapses. These findings are significant in that a positive correlation between *OnTrack* usage and weight loss was observed. Clinically significant weight losses were observed, and further research is necessary to determine the extent to which *OnTrack* is responsible for these outcomes. This weight loss in conjunction with the feasibility and acceptability suggests that, with further testing and research, *OnTrack* could prove to be a disseminable companion app to reduce dietary lapses and facilitate weight loss.

Acknowledgments: We would like to acknowledge and thank the efforts of our student programmers, Kyle Levin, Will Fligor, and Zack Haubach for their efforts on developing *OnTrack*.

Compliance with ethical standards: The findings reported here have not been previously published and the manuscript is not being simultaneously submitted elsewhere. The authors have full control of all primary data and agree to allow the journal to review the data if requested. This research was conducted under the supervision of Drexel University's Institutional Review Board and there are no ethical disclosures to be made.

Ethical Approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors.

Funding: This study was funded by the Weight Watchers Karen Miller-Kovach Research Grant from The Obesity Society and the Drexel Ventures Innovation Fund awarded to Dr. Forman.

Conflict of Interest: Gary D. Foster serves as the Chief Scientific Officer at Weight Watchers International, Inc. Remaining authors have no additional conflicts of interest to disclose.

Informed Consent: Informed consent was obtained from all individual participants included in the study.

References

1. Yang L, Colditz G. Prevalence of overweight and obesity in the United States. *JAMA Intern Med.* 2015;175(8):1412–1413.
2. Yaemsiri S, Slining MM, Agarwal SK. Perceived weight status, overweight diagnosis, and weight control among US adults: the NHANES 2003–2008 Study. *Int J Obes (Lond).* 2011;35(8):1063–1070.
3. Snook KR, Hansen AR, Duke CH, Finch KC, Hackney AA, Zhang J. Change in percentages of adults with overweight or obesity trying to lose weight, 1988–2014. *Jama.* 2017;317(9):971–973.
4. Brownell KD, Jeffery RW. Improving long-term weight loss: pushing the limits of treatment. *Behav Ther.* 1987;18(4):353–374.
5. Klem ML, Wing RR, McGuire MT, Seagle HM, Hill JO. A descriptive study of individuals successful at long-term maintenance of substantial weight loss. *Am J Clin Nutr.* 1997;66(2):239–246.
6. Wilson GT. Behavioral treatment of obesity: thirty years and counting. *Behav Res Ther.* 1994;16(1):31–75.
7. Carels RA, Douglass OM, Cacciapaglia HM, O'Brien WH. An ecological momentary assessment of relapse crises in dieting. *J Consult Clin Psychol.* 2004;72(2):341–348.
8. Carels RA, Hoffman J, Collins A, Raber AC, Cacciapaglia H, O'Brien WH. Ecological momentary assessment of temptation and lapse in dieting. *Eat Behav.* 2001;2(4):307–321.
9. Schumacher LM, Martin GJ, Goldstein SP, et al. Brief report: ecological momentary assessment of cognitive-affective responses to dietary lapses. *Health Psychol.* 2018;37(2):148.
10. Forman EM, Schumacher LM, Crosby R, et al. Ecological momentary assessment of dietary lapses across behavioral weight loss treatment: characteristics, predictors, and relationships with weight change. *Ann Behav Med.* 2017;51(5):741–753.
11. Tsai AG, Wadden TA. Systematic review: an evaluation of major commercial weight loss programs in the United States. *Ann Intern Med.* 2005;142(1):56–66.
12. Aguilar-Martinez A, Solé-Sedeño JM, Mancebo-Moreno G, Medina FX, Carreras-Collado R, Saigó-Rubió F. Use of mobile phones as a tool for weight loss: a systematic review. *J Telemed Telecare.* 2014;20(6):339–349.
13. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. *J Med Internet Res.* 2013;15(4):e86.
14. Bodenheimer T, Handley MA. Goal-setting for behavior change in primary care: an exploration and status report. *Patient Educ Couns.* 2009;76(2):174–180.
15. Chen J, Cade JE, Allman-Farinelli M. The most popular smartphone apps for weight loss: a quality assessment. *Jmir Mhealth Uhealth.* 2015;3(4):e104.
16. Pagoto S, Schneider K, Jovic M, DeBrosse M, Mann J. Evidence-based strategies in weight-loss mobile apps. *Am J Prev Med.* 2013;45(5):576–582.
17. Morris ME. Motivating change with mobile: seven guidelines. *Interactions.* 2012;19(3):26–31.

18. Adams MA, Sallis JF, Norman GJ, Hovell MF, Hekler EB, Perata E. An adaptive physical activity intervention for overweight adults: a randomized controlled trial. *Plos One*. 2013;8(12):e82901.
19. Thomas JG, Bond DS. Behavioral response to a just-in-time adaptive intervention (JITAI) to reduce sedentary behavior in obese adults: Implications for JITAI optimization. *Health Psychol*. 2015;34S:1261–1267.
20. Kumar S, Abowd GD, Abraham WT, et al. Center of excellence for mobile sensor data-to-knowledge (MD2K). *J Am Med Inform Assoc*. 2015;22(6):1137–1142.
21. Haapala I, Barengo NC, Biggs S, Surakka L, Manninen P. Weight loss by mobile phone: a 1-year effectiveness study. *Public Health Nutr*. 2009;12(12):2382–2391.
22. Burke LE, Styn MA, Glanz K, et al. SMART trial: a randomized clinical trial of self-monitoring in behavioral weight management—design and baseline findings. *Contemp Clin Trials*. 2009;30(6):540–551.
23. Beasley JM, Riley WT, Davis A, Singh J. Evaluation of a PDA-based dietary assessment and intervention program: a randomized controlled trial. *J Am Coll Nutr*. 2008;27(2):280–286.
24. Burke LE, Styn MA, Sereika SM, et al. Using mHealth technology to enhance self-monitoring for weight loss: a randomized trial. *Am J Prev Med*. 2012;43(1):20–26.
25. Ambeba EJ, Ye L, Sereika SM, et al. The use of mHealth to deliver tailored messages reduces reported energy and fat intake. *J Cardiovasc Nurs*. 2015;30(1):35–43.
26. Burke LE, Zheng Y, Ma Q, et al. The SMARTER pilot study: testing feasibility of real-time feedback for dietary self-monitoring. *Prev Med Rep*. 2017;6:278–285.
27. Forman EM, Butryn ML. A new look at the science of weight control: how acceptance and commitment strategies can address the challenge of self-regulation. *Appetite*. 2015;84:171–180.
28. Lowe MR, Annunziato R, Riddell LJ, et al. Reduced energy density eating and weight loss maintenance: 18-month follow-up results from a randomized controlled trial. *Obes Res*. 2003. In *Obesity research* (Vol. 11, pp. A22). North Amer Assoc Study Obesity, Silver Spring, MD.
29. Stroebe W, Aarts H, Schut H, Schut H, Kruglanski AW. Why dieters fail: testing the goal conflict model of eating. *J Exp Soc Psychol*. 2008;44(1):26–36.
30. McKee HC, Ntoumanis N, Taylor IM. An ecological momentary assessment of lapse occurrences in dieters. *Ann Behav Med*. 2014;48(3):300–310.
31. Thomas JG. *Toward a Better Understanding of the Development of Overweight: A Study of Eating Behavior in the Natural Environment Using Ecological Momentary Assessment* [PhD thesis]. Philadelphia, PA: Drexel University; 2009.
32. Drapkin RG, Wing RR, Shiffman S. Responses to hypothetical high risk situations: do they predict weight loss in a behavioral treatment program or the context of dietary lapses? *Health Psychol*. 1995;14(5):427–434.
33. Smyth JM, Wonderlich SA, Sliwinski MJ, et al. Ecological momentary assessment of affect, stress, and binge-purge behaviors: day of week and time of day effects in the natural environment. *Int J Eat Disord*. 2009;42(5):429–436.
34. Hofmann W, Friese M. Impulses got the better of me: alcohol moderates the influence of implicit attitudes toward food cues on eating behavior. *J Abnorm Psychol*. 2008;117(2):420–427.
35. Muraven M, Collins RL, Shiffman S, Paty JA. Daily fluctuations in self-control demands and alcohol intake. *Psychol Addict Behav*. 2005;19(2):140–147.
36. Ward A, Mann T. Don't mind if I do: disinhibited eating under cognitive load. *J Pers Soc Psychol*. 2000;78(4):753–763.
37. Baumeister RF. Yielding to temptation: self-control failure, impulsive purchasing, and consumer behavior. *J Consum Res*. 2002;28(4):670–676.
38. Tulu B, Ruiz C, Allard J, et al. SlipBuddy: a mobile health intervention to prevent overeating. Proceedings of the 50th Hawaii International Conference on System Sciences, 2017.
39. Goldstein SP, Evans BC, Flack D, 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.
40. Kotsiantis SB. Logitboost of simple bayesian classifier. *Informatica*. 2005;29(1):43–49.
41. Quinlan JR. *Bagging, boosting, and C4. 5*. Vol. 1. AAAI/IAAI; 1996. pp. 725–730.
42. Skurichina M, Duin RP. Bagging, boosting and the random subspace method for linear classifiers. *Pattern Anal Appl*. 2002;5(2):121–135.
43. Chen C, Liaw A, Breiman L. *Using Random Forest to Learn Imbalanced Data*. Vol. 110. Berkeley: University of California; 2004.
44. Friedman N, Geiger D, Goldszmidt M. Bayesian network classifiers. *Mach Learn*. 1997;29(2–3):131–163.
45. Goldstein SP, Zhang F, Thomas JG, Butryn ML, Herbert JD, Forman EM. Application of machine learning to develop a predictive algorithm for weight control lapses. under review
46. Michie S, Ashford S, Snihotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. *Psychol Health*. 2011;26(11):1479–1498.
47. Latner JD, McLeod G, O'Brien KS, Johnston L. The role of self-efficacy, coping, and lapses in weight maintenance. *Eat Weight Disord*. 2013;18(4):359–366.
48. Grilo CM, Shiffman S, Wing RR. Relapse crises and coping among dieters. *J Consult Clin Psychol*. 1989;57(4):488–495.
49. Goldstein SP, Forman EM, Butryn ML, Herbert J. Web-based versus in-person physical activity promotion programs for college students: differential preferences for programming content. *Health Commun*. 2017:1–7.
50. Lee Y, Kozar KA, Larsen KR. The technology acceptance model: past, present, and future. *Commun Assoc Inform Syst*. 2003;12(1):50.
51. Goldstein SP. *A Preliminary Investigation of a Personalized Risk Alert System for Weight Control Lapses*. Philadelphia, PA: Drexel University; 2016.
52. Van Gaal LF, Wauters MA, De Leeuw IH. The beneficial effects of modest weight loss on cardiovascular risk factors. *Int J Obes Relat Metab Disord*. 1997;21(Suppl 1):S5–S9.
53. Wing RR, Lang W, Wadden TA, et al.; Look AHEAD Research Group. Benefits of modest weight loss in improving cardiovascular risk factors in overweight and obese individuals with type 2 diabetes. *Diabetes Care*. 2011;34(7):1481–1486.
54. Patrick K, Raab F, Adams MA, et al. A text message-based intervention for weight loss: randomized controlled trial. *J Med Internet Res*. 2009;11(1):e1.
55. Flores Mateo G, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. *J Med Internet Res*. 2015;17(11):e253.
56. Laing BY, Mangione CM, Tseng CH, et al. Effectiveness of a smartphone application for weight loss compared with usual care in overweight primary care patients: a randomized, controlled trial. *Ann Intern Med*. 2014;161(10 Suppl):S5–12.
57. Lyzwinski LN. A systematic review and meta-analysis of mobile devices and weight loss with an intervention content analysis. *J Pers Med*. 2014;4(3):311–385.
58. Thomas JG, Doshi S, Crosby RD, Lowe MR. Ecological momentary assessment of obesogenic eating behavior: combining person-specific and environmental predictors. *Obesity (Silver Spring)*. 2011;19(8):1574–1579.
59. Haedt-Matt AA, Keel PK. Revisiting the affect regulation model of binge eating: a meta-analysis of studies using ecological momentary assessment. *Psychol Bull*. 2011;137(4):660–681.
60. Haedt-Matt AA, Keel PK. Hunger and binge eating: a meta-analysis of studies using ecological momentary assessment. *Int J Eat Dis*. 2011;44(7):573–578.
61. Schippers M, Adam PC, Smolenski DJ, Wong HT, de Wit JB. A meta-analysis of overall effects of weight loss interventions delivered via mobile phones and effect size differences according to delivery mode, personal contact, and intervention intensity and duration. *Obes Rev*. 2017;18(4):450–459.
62. Group DPPR. Achieving weight and activity goals among diabetes prevention program lifestyle participants. *Obes Res*. 2004;12(9):1426.
63. Thomas JG, Raynor HA, Bond DS, et al. Weight loss and frequency of body weight self-monitoring in an online commercial weight management program with and without a cellular connected “smart” scale: a randomized pilot study. *Obesity science & practice*. 2017; 3(4), 365–372.
64. Group LAR. The Look AHEAD study: a description of the lifestyle intervention and the evidence supporting it. *Obesity (Silver Spring, Md)*. 2006;14(5):737.
65. Group DPPR. The diabetes prevention program (DPP). *Diabetes Care*. 2002;25(12):2165–2171.
66. Forman EM, Butryn ML, Juarascio AS, et al. The mind your health project: a randomized controlled trial of an innovative behavioral treatment for obesity. *Obesity (Silver Spring)*. 2013;21(6):1119–1126.
67. Klasnja P, Hekler EB, Shiffman S, et al. Microrandomized trials: an experimental design for developing just-in-time adaptive interventions. *Health Psychol*. 2015;34S:1220–1228.
68. Chih MY, Patton T, McTavish FM, et al. Predictive modeling of addiction lapses in a mobile health application. *J Subst Abuse Treat*. 2014; 46(1): 29–35.