Machine-learning analysis identifies digital behavioral phenotypes for engagement and health outcome efficacy of mHealth interventions for obesity: post-hoc analyses of a randomized trial

Meelim Kim, Jaeyeong Yang, Woo-Young Ahn, Hyung Jin Choi

Submitted to: Journal of Medical Internet Research
on: January 16, 2021

Table of Contents

Machine-learning analysis identifies digital behavioral phenotypes for engagement and health outcome efficacy of mHealth interventions for obesity: post-hoc analyses of a randomized trial

Meelim Kim[1*] MA; Jaeyeong Yang[2*] MS; Woo-Young Ahn[2*] PhD; Hyung Jin Choi[3*] MD, PhD

[1]Seoul National University Seoul KR
[2]Seoul National University College of Medicine Seoul KR
[*]these authors contributed equally

Corresponding Author:
Hyung Jin Choi MD, PhD
Seoul National University College of Medicine
Room #216, Research Building, 103 Daehak-ro, Jongno-gu
Seoul
KR

Abstract

Background: The digital healthcare community has been urged to enhance engagement and clinical outcomes by analyzing multidimensional digital phenotypes.

Objective: This study aimed to investigate the performance of multivariate phenotypes predicting the engagement rate and health outcomes of digital cognitive behavioral therapy (dCBT) using a machine learning approach.

Methods: We leveraged both conventional phenotypes assessed by validated psychological questionnaires and multidimensional digital phenotypes within time-series data from a mobile app of 45 participants undergoing digital cognitive behavioral therapy (dCBT) for eight weeks. To discriminate the important characteristics, we conducted a machine-learning analysis.

Results: A higher engagement rate was associated with higher weight loss at 8 weeks (r = -0.59, p < 0001) and 24 weeks (r = -0.52, p = 0001). The machine learning approach revealed distinct multivariate profiles associated with varying impacts on the outcomes. Lower self-esteem on the conventional phenotype and higher in-app motivational measures on digital phenotypes commonly accounted for both engagement and health outcomes. In addition, eight types of digital phenotypes predicted engagement rates (mean R2 = 0416, SD = 0006). The prediction of short-term weight change (mean R2 = 0382, SD = 0015) was associated with six different digital phenotypes. Lastly, two behavioral measures of digital phenotypes were associated with a long-term weight change (mean R2 = 0590, SD = 0011).

Conclusions: Our findings successfully demonstrated how multiple psychological constructs, such as emotional, cognitive, behavioral, and motivational phenotypes, elucidate the mechanisms and clinical efficacy of digital intervention with the machine learning method. Our results also highlight the importance of assessing multiple aspects of motivation before and during the intervention to improve both engagement rate and clinical outcomes. This line of research may shed light on the development of advanced prevention and personalized digital therapeutics. Clinical Trial: ClinicalTrials.gov NCT03465306 (Retrieved September 18, 2017, https://register.clinicaltrials.gov/NCT03465306)

(JMIR Preprints 16/01/2021:27218)
DOI: https://doi.org/10.2196/preprints.27218

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✔ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Original Paper

Meelim Kim, MA
Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, Korea
BK21 Plus Biomedical Science Project Team, Seoul National University College of Medicine, Seoul, Korea
Address: 28 Yungun-Dong, Chongno-Gu, Seoul 110–799, Korea
Jaeyeong Yang, MS
Department of Psychology, Seoul National University, Seoul, Korea
Address: 1 Gwanak-ro, Gwanak-gu, Seoul, Korea, 08826
Woo-Young Ahn, PhD
Department of Psychology, Seoul National University, Seoul, Korea
Address: 1 Gwanak-ro, Gwanak-gu, Seoul, Korea, 08826
Hyung Jin Choi, MD
Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, Korea
BK21 Plus Biomedical Science Project Team, Seoul National University College of Medicine, Seoul, Korea
Address: 28 Yungun-Dong, Chongno-Gu, Seoul 110–799, Korea

Address for correspondence: Woo-Young Ahn

Department of Psychology, Seoul National University

1 Gwanak-ro, Gwanak-gu, Seoul, Korea, 08826

Tel: +82-2-880-2538; Email: wahn55@snu.ac.kr

Address for correspondence: Hyung Jin Choi

Department of Biomedical Sciences, Seoul National University College of Medicine

28 Yungun-Dong, Chongno-Gu, Seoul 110–799, Korea

Tel: +82-2-740-8204; Fax: +82-2-745-9528; E-mail: hjchoi@snu.ac.kr

Machine-learning analysis identifies digital behavioral phenotypes for engagement and health outcome efficacy of mHealth interventions for obesity: post-hoc analyses of a randomized trial

Abstract

Background - The digital healthcare community has been urged to enhance engagement and clinical outcomes by analyzing multidimensional digital phenotypes.

Objectives - This study aimed to investigate the performance of multivariate phenotypes predicting the engagement rate and health outcomes of digital cognitive behavioral therapy (dCBT) using a machine learning approach.

Methods - We leveraged both conventional phenotypes assessed by validated psychological questionnaires and multidimensional digital phenotypes within time-series data from a mobile app of 45 participants undergoing digital cognitive behavioral therapy (dCBT) for eight weeks. To discriminate the important characteristics, we conducted a machine-learning analysis.

Results - A higher engagement rate was associated with higher weight loss at 8 weeks (r = -0.59, p < 0001) and 24 weeks (r = -0.52, p = 0001). The machine learning approach revealed distinct multivariate profiles associated with varying impacts on the outcomes. Lower self-esteem on the conventional phenotype and higher in-app motivational measures on digital phenotypes commonly accounted for both engagement and health outcomes. In addition, eight types of digital phenotypes predicted engagement rates (mean R^2 = 0416, SD = 0006). The prediction of short-term weight change (mean R^2 = 0382, SD = 0015) was associated with six different digital phenotypes. Lastly, two behavioral measures of digital phenotypes were associated with a long-term weight change (mean R^2 = 0590, SD = 0011).

Conclusions - Our findings successfully demonstrated how multiple psychological constructs, such as emotional, cognitive, behavioral, and motivational phenotypes, elucidate the mechanisms and clinical efficacy of digital intervention with the machine learning method. Our results also highlight the importance of assessing multiple aspects of motivation before and during the intervention to improve both engagement rate and clinical outcomes. This line of research may shed light on the development of advanced prevention and personalized digital therapeutics.

Trial Registration - ClinicalTrials.gov NCT03465306 (Retrieved September 18, 2017, https://register.clinicaltrials.gov/NCT03465306).

Keywords: digital phenotype, clinical efficacy, in-app engagement, machine-learning analysis

Research in context

Evidence before the study

Much effort has been made to enhance the digital engagement and clinical efficacy of digital interventions, and digital phenotyping is known to improve the prediction of both the engagement and outcomes of the interventions. We searched PubMed for publications in English from the database inception until August 31, 2020, using the terms “digital phenotypes,” “machine learning,” “prediction models,” “engagement in-app,” “clinical efficacy,” and “mobile intervention.” Some studies have used machine learning approaches and reported the prediction model's performance, but with few interpretable implications regarding specific clinical phenotypes. There is a strong need to successfully integrate multidimensional digital phenotypes and reveal the psychological mechanisms underlying digital engagement and clinical outcomes.

Added value of the study

To our knowledge, this is the first study to predict both digital engagement and clinical efficacy in a digital intervention

trial by applying a machine learning approach to multidimensional phenotypes, including emotion, cognition, behavior, and motivation, using conventional and digital measures. Among the multidimensional phenotypes, several behavioral (carbohydrate intake and restrained eating style) and motivational (confidence and conventional motivation) phenotypes, as well as digital engagement (in-app activity), significantly predicted the target health outcome (weight change). The conventional phenotypes from the baseline and digital phenotypes also predicted digital engagement during the intervention period. Our model achieves not only excellent out-of-sample prediction accuracy but also provides multivariate profiles of digital phenotypes that contribute to digital engagement and clinical outcomes.

Implications of all the available evidence

This proof-of-concept study demonstrated that it is possible to develop an interpretable digital phenotype model that predicts digital engagement in a mobile app and a digital intervention's clinical efficacy. This work line might open a new era of applicable and personalized digital therapeutics, which assist clinicians in making clinical decisions during an intervention.

Introduction

The use of mobile tools, such as smartphones, to assist healthcare systems is rapidly growing in the current era. As the interactions between individuals and digital communities via mobile devices are progressively embedded in human lives, understanding the concept of a “digital phenotype” is also important. A digital phenotype is a collected set of data in a digital system by humans demonstrating intentionally or as a secondary outcome of other activities, influencing human behavior. Specifically, the expanding body of health-related data from mobile devices allows us to address human illness real-world life events. For example, data related to the timing and periods of one’s digital footprint can be examined as part of a patient's features with insomnia [1]. Similarly, while data from Google search can recognize suicidal ideation [2]. To date, digital technologies such as smartphone applications (apps) afford moment-by-moment perceptible measurements of a person’s behavior regarding preventive and predictive ways to manage health.

Obtaining app users' attention is a critical issue related to the app’s potential efficacy for behavior change. The association between intervention exposure and efficacy emphasizes the need for a detailed understanding of user engagement [3]. When we deliver an intervention via a mobile app, the users must be actively and frequently engaged with mobile apps to succeed within the treatment. Thus, identifying predictive markers that can inform engagement in mHealth interventions could potentially strengthen its effectiveness. Previous studies have found that social and gamified components' involvement or offering personalized feedback from human factors effectively enhances user engagement for app-based interventions [4,5]. In fact, identifying the major principles that can predict users’ engagement and health outcomes is important for exploring systemic elements to strengthen user engagement in digital intervention. Engagement with digital technology is intricate because it is not stationary but a progressive process [6]. It is also multifaceted in its environment, reflecting the quality of the user’s practice, their communication features, and their willingness to use the app over time or repeatedly [7]. Of special interest to this present issue, it is noted that intrinsic motivation is a significant precursor for engagement [8]. Moreover, a wide range of cognitive and emotional states such as self-interest and self-efficacy are closely related to the user’s engagement [7]. It is important to measure and estimate the user’s engagement with multiple components, intensifying the treatment's efficacy, and finding good responders for precision medicine. However, finding the major indicator that predicts who will benefit the most from digital intervention is insufficient. This resulted in only a minor portion of users obtaining advantages from the digital healthcare system [9,10]. Thus, it is necessary to explore how comprehensive and multidimensional digital phenotypes detect individual differences and determine the user’s engagement in the digital intervention. A conceptual framework of mHealth components, including examples of digital phenotypes, is presented in Figure 1.

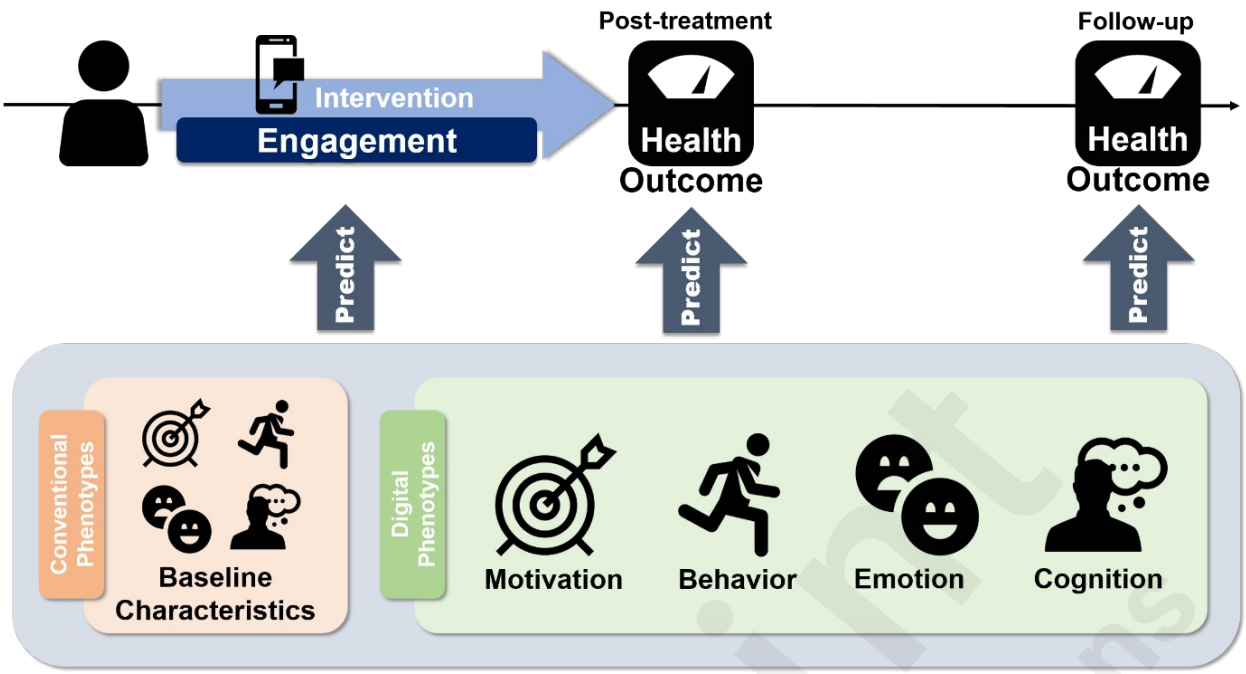


Figure 1. A conceptual framework of mHealth components and examples of digital phenotypes.

Another major issue in the digital era is the interpretation and filtering of data for clinical decisions. Although the rapid growth of digital technologies has led to comprehensive and abundant information about one's health status, analytical methods to clarify and simplify it have not advanced at a compatible pace [11]. This could be addressed as the main bottleneck in current digital phenotyping studies. Some pioneering research has demonstrated statistical methods to derive insights (which predict outcome) from various digital phenotypes [12–14]. However, the data are mostly heterogeneous and mixed with structured and unstructured frames containing random sampling, artifacts, and inconsistent completion, making traditional statistical models difficult. This can lead to limited or biased results from the data and a lack of replicability of the conclusions. Compared to conventional analytical methods, machine-learning analysis can obtain information from scattered and intricate data, offering insights to promote clinical decision-making. A recent study has shown that mortality prediction models using intensive care unit (ICU) data based on a machine learning approach were superior to conventional methods [15]. The algorithms supporting individual-specific predictions might enhance the usability of machine-learning prediction models. This could aid in the adaptation of machine learning models as clinical decision-support tools.

Although previous mHealth intervention studies have shown that user engagement is critical, little effort has been made to conceptualize and estimate it. The major reason is that only a few mHealth programs predominantly use the applicable data to investigate participants' engagement or to examine its correlation with primary outcomes. Utilizing digital phenotypes and enhancing our insight into them to promote management will involve refined approaches for choosing and investigating diverse digital health data streams in a definite manner. Here, we investigated multidimensional information at different time points using various assessment methods to monitor and predict the primary outcome's engagement and efficacy. This study plays a significant role in establishing the most practical and effective mHealth intervention paradigm.

Methods

Study design and participants

We performed a post-hoc analysis based on data from a previously reported open-label, eight-week, active comparator randomized controlled trial (RCT) in the dCBT study. The trial was registered with ClinicalTrials.gov NCT03465306 in March 2018. Methods of recruitment, inclusion, and exclusion criteria, and demographics have been published elsewhere [16]. A total of 70 female participants between 18 and 39 years of age, body mass index of 25 to 40 kg, eligibility for smartphone usage, and the scores in the highest 40% on the Situational Motivation Scale (SIMS; scores above 68 out of 112 in total) were enrolled. All study participants provided written informed consent before enrollment in the study. The Institutional Review Board of Seoul National University Hospital approved the study (approval number H-1707-122-872). The study protocol was registered at ClinicalTrials.gov NCT03465306 on January 15, 2018. This study was conducted to validate the obesity dCBT model's clinical efficacy and identify factors related to its efficacy.

Seventy participants were enrolled, but we focused on forty-five participants who underwent the CBT and responded to all the digital phenotypes' items throughout the clinical period. Additionally, we excluded seven participants due to dropout and one participant due to lack of participation (less than 15 days). Therefore, the data we analyzed included 37 participants. Further, all the digital phenotypes were averaged for each participant to predict their engagement during the intervention and their health outcomes for both short-term (8 weeks) and long-term (24 weeks). The mean age was 22.59 years, and the SD was 369.

Randomization and masking

Participants were randomly assigned to a control group or dCBT group with a ratio of 1 to 2. Research participants and research staff were aware of group assignment, but the group assignments were blinded to technicians and clinical staff.

Procedures

The aim of the current analysis was a post-hoc analysis of the effects of a dCBT intervention on obesity. The detailed design and procedures have been described elsewhere [16]. Briefly, the participants in the dCBT group (App+human CBT) consisted of daily individualized feedback and assignments from a clinical psychologist based on the CBT modules for eight weeks. All participants were asked to visit at baseline and 8 and 24 weeks. Anthropometric and self-administered questionnaires were collected at each study visit. The Noom app was mainly used for logging food diaries and delivering messages between the therapist and participants.

Measures

The statistical information for the baseline characteristics and in-app measures is shown in Table S1. There are two main structures: conventional and digital phenotypes, which are classified based on different algorithms. Conventional phenotypes were composed of previously developed and validated surveys. Digital phenotypes are generated by a newly devised scoring system consisting of a combination of active and passive digital features gathered from digital devices.

These phenotypes are categorized into four different dimensions: behavioral, cognitive, emotional, and motivational. For conventional phenotypes, four indices for each behavioral and emotional dimension, and one index for cognitive and motivational dimensions were assessed. Seventeen indices for behavioral, one for cognitive, five for emotional, and four for motivational dimensions were assessed regarding digital phenotypes.

Participants' situational motivation toward the weight loss program was assessed using an adapted version of the SIMS. SIMS typically measures four types of motivation to engage in a task (herein, the weight loss program) at a specific point in time, with four items per subscale: intrinsic motivation, identified regulation, external regulation, and motivation. SIMS has demonstrated acceptable levels of reliability and validity in previous research. The Body Shape Questionnaire-8C (BSQ-8C) is a brief form of the BSQ and consists of eight items extracted from the full version measuring the extent of the psychopathology of concerns about body shape. Higher BSQ values indicated greater body dissatisfaction. Depression was assessed using the Korean version of the Beck Depression Inventory (K-BDI-II) scoring system. A total score from 0 to 9 indicated no depression, 10 to 15 indicated mild depression, 16 to 23 indicate moderate depression, and 24 to 63 indicated severe depression. Anxiety was measured using the 20-item Trait Anxiety Scale (TAI) of the State-Trait Anxiety Inventory, with greater scores indicating more trait anxiety. The Rosenberg Self-Esteem Scale (RSES) measure of self-esteem was used in this research with a 10-item scale with all negatively worded items. Thus, higher scores implied lower self-esteem. Eating behavior was measured using the Dutch Eating Behavior Questionnaire (DEBQ), which has three different psychologically based eating behaviors: restrained eating, emotional eating, and external eating. It contains 33 items, with higher scores indicating a greater tendency to present subscale behavior. The frequency of automatic negative thoughts associated with depression was assessed using the Automatic Thoughts Questionnaire (ATQ-30). The scores ranged from 30 to 150, and higher scores implied that the participants experienced automatic negative thoughts more often. All psychological questionnaires were presented in Korean.

There are six types of behavioral phenotypes assessed in apps: food restriction, overeating and binge eating, late-night meals, snacking, food choice, and activity rate. Food restriction was evaluated using kcal per meal per day. Overeating and binge eating were assessed by kcal per meal per day, and the speed per meal—the late-night meal was investigated using the dinner kcal and the time per meal. Snacking was estimated using snack kcal. Food choice was examined based on the type of food per meal, total amount of sodium and sugar, number of food types per meal, and percentage of nutritional types (carb, protein, and fat). The activity rate was measured as the number of steps and the total hours of exercise. Automatic thoughts were grouped into six categories: selective abstraction, arbitrary inference, overgeneralization, magnification/minimization, personalization, and absolutism. There were 20 automatic thoughts, and participants could add thoughts related to food or eating behaviors. Example statements for automatic thoughts are listed in Table S2. We assessed five negative emotions closely related to problematic eating habits: irritation, loneliness, nervousness, boredom, and depression. The participants were asked to report each type of negative emotion scores using a visual analog scale (VAS) between 0 and 100. The motivation was assessed using four dimensions: will, rank of importance, confidence, and satisfaction. These different types of motivation were scored using a 10-point Likert scale (1–10).

Outcomes

The primary outcomes were changes in body weight and the number logged into the app. Body weight was assessed by

InBody H20B (InBody Co., Ltd., Seoul, South Korea) at baseline, 8, and 24 weeks in light street clothing and without socks and shoes. The number logged into the app was examined by tracking the actions such as responses to the daily assessment (responses per day), meals logged (meals per week), green foods defined by Noom logged (logged per week), exercise logged (times per week), exercise time registered (minutes per week), steps recorded (steps per week), weigh-ins logged (times per week), articles read (articles per week), group posts (posts per week), group comments (comments per week), messages sent to coaches (messages per week), and group likes (likes per week). The engagement rate was assessed using these objective indices for each participant.

Statistical analysis

We analyzed the data to predict three target outcomes: (a) the number of mobile activities during the experiment session, (b) the weight change rate between pre-session (week 0) and post-session (week 8), and (c) the weight change rate between post-session and follow-up. The weight change rates were calculated as the ratio of the weight difference to the baseline weight as $\frac{weight_{before} - weight_{after}}{weight_{before}}$. Correlations between the number of logs and weight change rates were analyzed to determine the relationship between engagement and health outcomes.

A machine-learning approach using an elastic net [17] was conducted. The elastic net is a penalized regression method that automatically selects significant variables by reducing the regression coefficients of unimportant features to zero. Using 41 behavioral, cognitive, motivational, and emotional measures, we tried to reveal which measure contributes to predicting behavioral changes before and after treatment.

The analysis procedure for out-of-sample regressions is similar to that in a previous study [18,19]. To conduct out-of-sample regression, we use leave-one-out cross-validation (LOOCV), which trains a model with data except for a single point and then evaluates the point's prediction. Root mean squared errors (RMSE) computed for all possible train test splits are averaged to the leave-one-out cross-validation error, which is the measure for evaluating the model fit.

To acquire generalizable coefficients, we conducted model fitting 1,000 times for each possible alpha value (α), which is the ratio between the ridge and lasso penalty terms. Figure S1 shows the RMSE with 100 alpha values (from 0.01 to 1 with an interval of 0.01), and we chose the alpha value that minimizes RMSE across all participants. After choosing the model with the best fit, we analyzed the regression coefficients.

Results

Relationship between the number of logs and weight changes

Figure 2 shows the correlations between the number of logs (engagement) and weight change (health outcomes). For the weight change during the 8-week intervention, two variables were highly correlated ($r = -0.59$, $t = -4 - 32$, $df = 35$, $p < 0.0001$; Figure 2A), which indicates that participants who had engaged in the in-app activity more actively lost weight. This result was the same for the weight change between baseline and follow-up ($r = -0.52$, $t = -3 - 59$, $df = 35$, $p = 0.00099$; Figure 2B). These short-term and long-term health outcomes were highly correlated with each other ($r = 074$, $t = 6.60$, $df = 35$, $p < 0.0001$; Figure 2C).

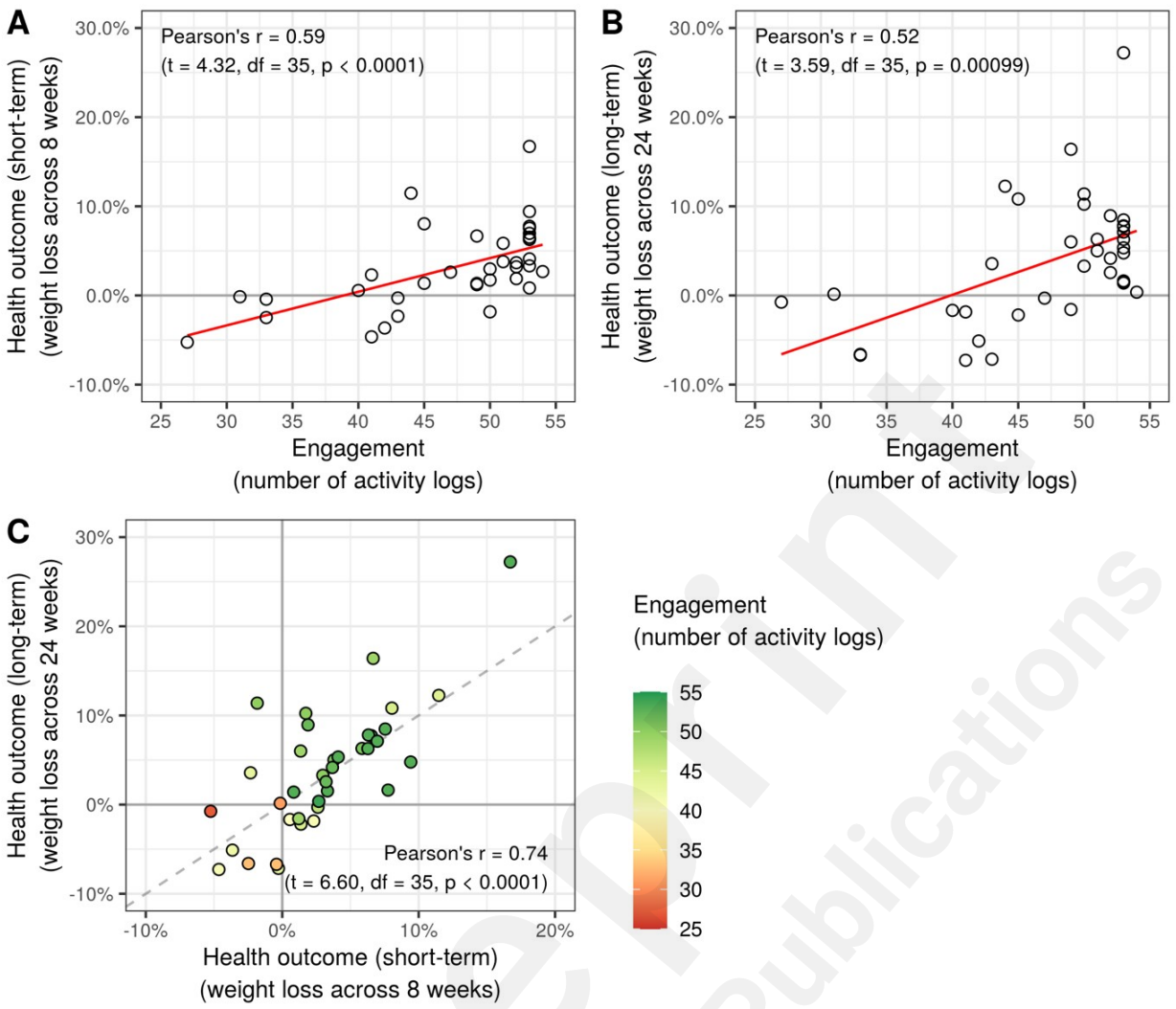


Figure 2. Relationships between engagement and health outcomes. The health outcome larger than zero indicates weight loss compared to baseline.

Elastic net results

Through the leave-one-out cross-validations with different values for the mixing parameter (alpha), we chose the best value for each model that showed the minimum RMSE between the data and predicted outcomes. The estimated mixing parameters, alphas, were 008, 015, and 0.53 for predicting engagement, short-term health outcome, and long-term health outcome, respectively (see Figure S1). The alpha estimate for the long-term health outcome was much higher than that in the other two models, suggesting that the multivariate pattern is more parsimonious. Its coefficients are prone to shrink to zero while predicting long-term weight change.

Figure 3 illustrates the multivariate profiles of conventional and digital phenotypes to predict in-app engagement and the health outcomes of digital healthcare. In-app engagement, computed as the number of daily activity logs, was significantly associated with lower self-esteem, lower body satisfaction, and higher external eating behaviors, measured as conventional phenotypes. For digital phenotypes, engagement was predicted by lower intake of food with a high-calorie density index (CDI), higher food intake in the morning (breakfast, morning snack), lower food intake after that (lunch, dinner, evening snack), higher sugar intake, higher intake of moderate or low CDI food, and higher frequency of interactions with the therapist. Higher emotional and motivational measures in digital phenotypes were also involved,

such as irritation, boredom, depression, satisfaction, will, and confidence.

For short-term health outcomes, lower emotional eating behavior, lower self-esteem, lower anxiety, higher external eating behavior, and higher motivation predicted the weight change rate for eight weeks. The 8-week weight change was also predicted by lower intake of high-CDI food, lower carb, lower sodium, lower fat intake, higher afternoon snack intake, lower dinner intake, higher intake of low CDI food, and higher frequency interactions with a healthcare mentor. Furthermore, short-term health outcomes were positively associated with emotional and motivational features in digital phenotypes, such as boredom, irritation, will, satisfaction, and confidence.

In contrast, fewer phenotypes are involved in the prediction of long-term health outcomes. Lower self-esteem, lower food addiction, lower body satisfaction, higher motivation, and higher restricting eating behavior in conventional phenotypes predicted the 24-week weight change. For digital phenotypes, the long-term health outcome was predicted by lower carb intake, lower lunch and evening snack intake, lower fat intake, lower steps in a day, higher satisfaction, higher will, and higher confidence.

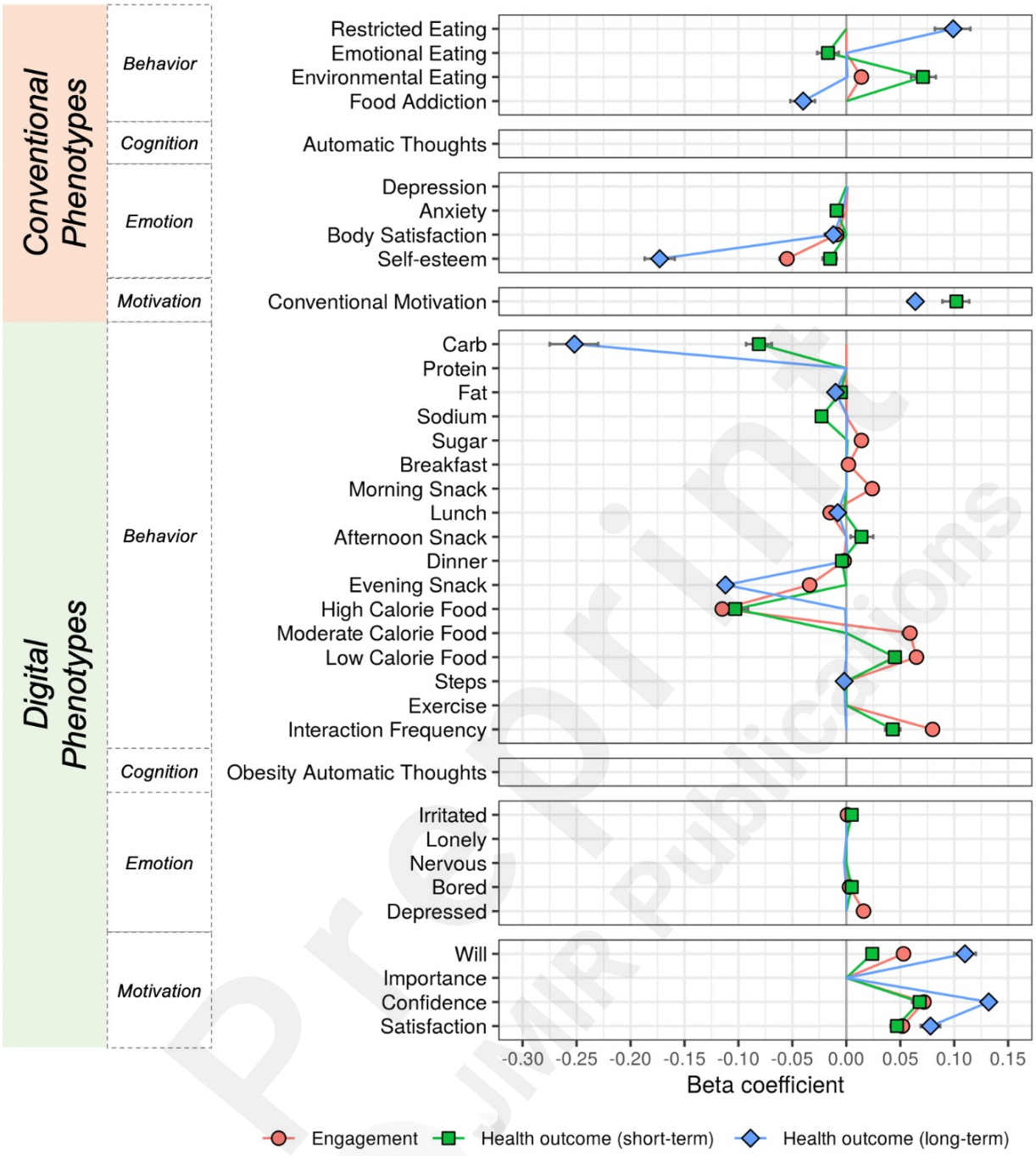


Figure 3. Multivariate patterns of conventional and digital phenotypes for predicting engagement (red) as well as short-term (green) and long-term (blue) health outcomes. Points indicate the averaged beta coefficients across 100 repetitions of net elastic analysis (see Methods for details). A positive beta estimate of a phenotype indicates an association between the phenotype and higher in-app activities (engagement) or more weight loss (health outcomes). The points, which contain zero in the simulated 95% ranges, are omitted.

Common predictors across dependent variables were associated with different phenotypes (Figure 4 and Table 1). Engagement and health outcomes were commonly affected by lower self-esteem in conventional phenotypes and higher in-app motivational measures in digital phenotypes. In other words, decreased self-esteem before the intervention and inclined motivation during the intervention highly predicted more in-app activities and more weight loss following the intervention. Furthermore, common predictors between engagement and short-term health outcomes include the behavioral dimension of digital phenotypes, such as the frequency of coach interaction and low/high-calorie food intake.

For predicting short-term and long-term health outcomes, carb intake was the most commonly influential predictor. Conversely, conventional and digital phenotypes' motivational measures were positively associated with health outcomes (see Figure 5).

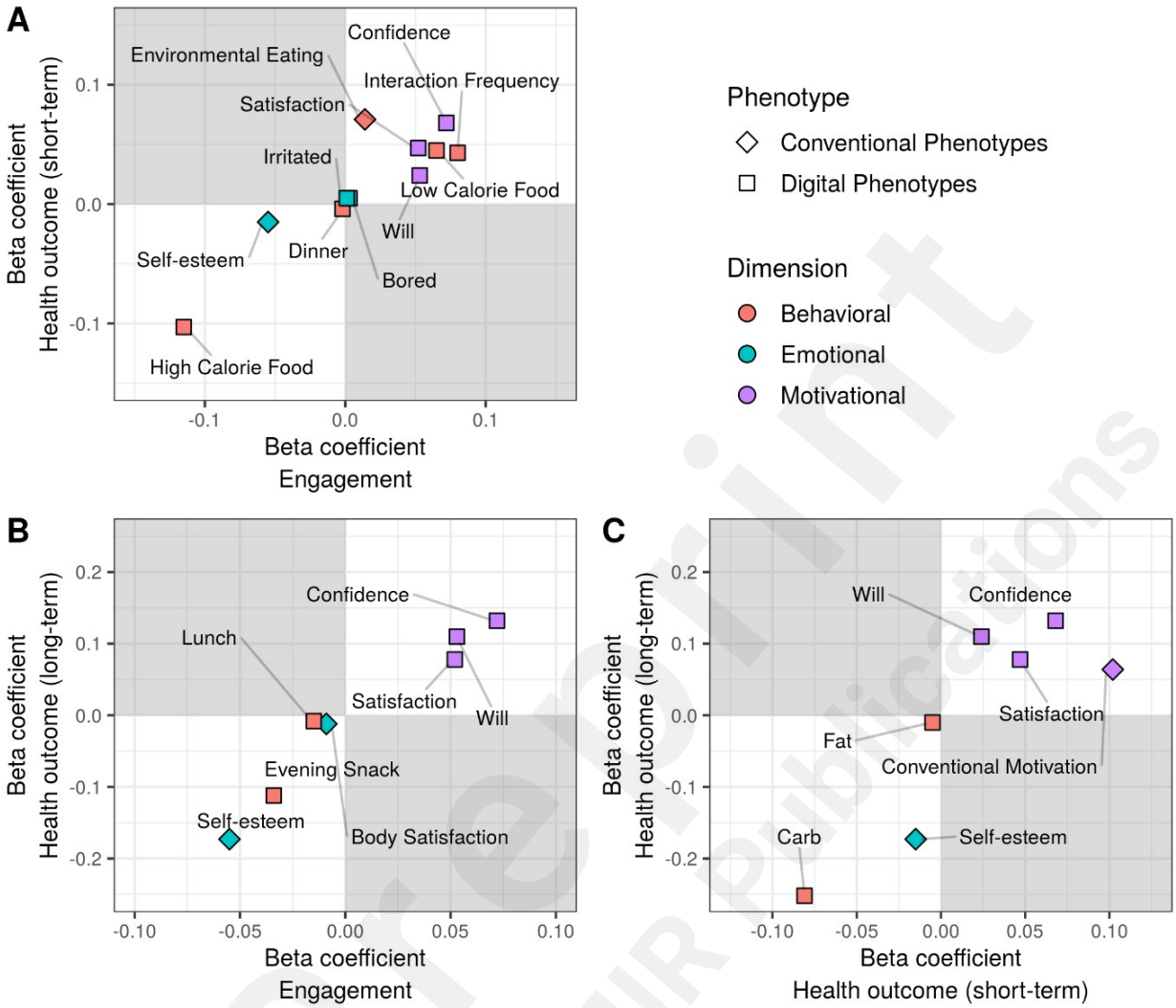


Figure 4. Common predictors between engagement and health outcomes. Each axis indicates the beta estimate for predicting engagement and health outcomes. A positive beta coefficient indicates a positive association with engagement but negative associations with health outcomes (weight changes).

Table 1. Common and specific predictors of conventional and digital phenotypes for predicting engagement and health outcomes.[a]

		Common predictors	Predictors specific to each dependent variable		
			Engagement	Health outcome (short-term)	Health outcome (long-term)
Conventional phenotypes		Self-esteem ↓	Body Satisfaction ↓ External Eating ↑	Emotional Eating ↓ Anxiety ↓ External Eating ↑ Conventional Motivation ↑	Food Addiction ↓ Body Satisfaction ↓ Conventional Motivation ↑ Restrictive Eating ↑
Digital	Behavioral	-	High-Calorie Food ↓	High-Calorie Food ↓	Carb ↓

phenotypes					
			Night Snack ↓ Lunch ↓ Dinner ↓ Breakfast ↑ Sugar ↑ Morning Snack ↑ Moderate Calorie Food ↑ Low-Calorie Food ↑ Interaction Frequency ↑	Carb ↓ Sodium ↓ Fat ↓ Afternoon Snack ↓ Low-Calorie Food ↑ Interaction Frequency ↑	Night Snack ↓ Lunch ↓ Fat ↓ Steps ↓
	Emotional	-	Irritated ↑ Bored ↑ Depressed ↑	Irritated ↑ Bored ↑	-
	Motivational	Satisfaction ↑ Will ↑ Confidence ↑	-	-	-

[a]Common predictors in the first column were involved in all models. The cognitive dimension of digital phenotypes is omitted due to a lack of significance.

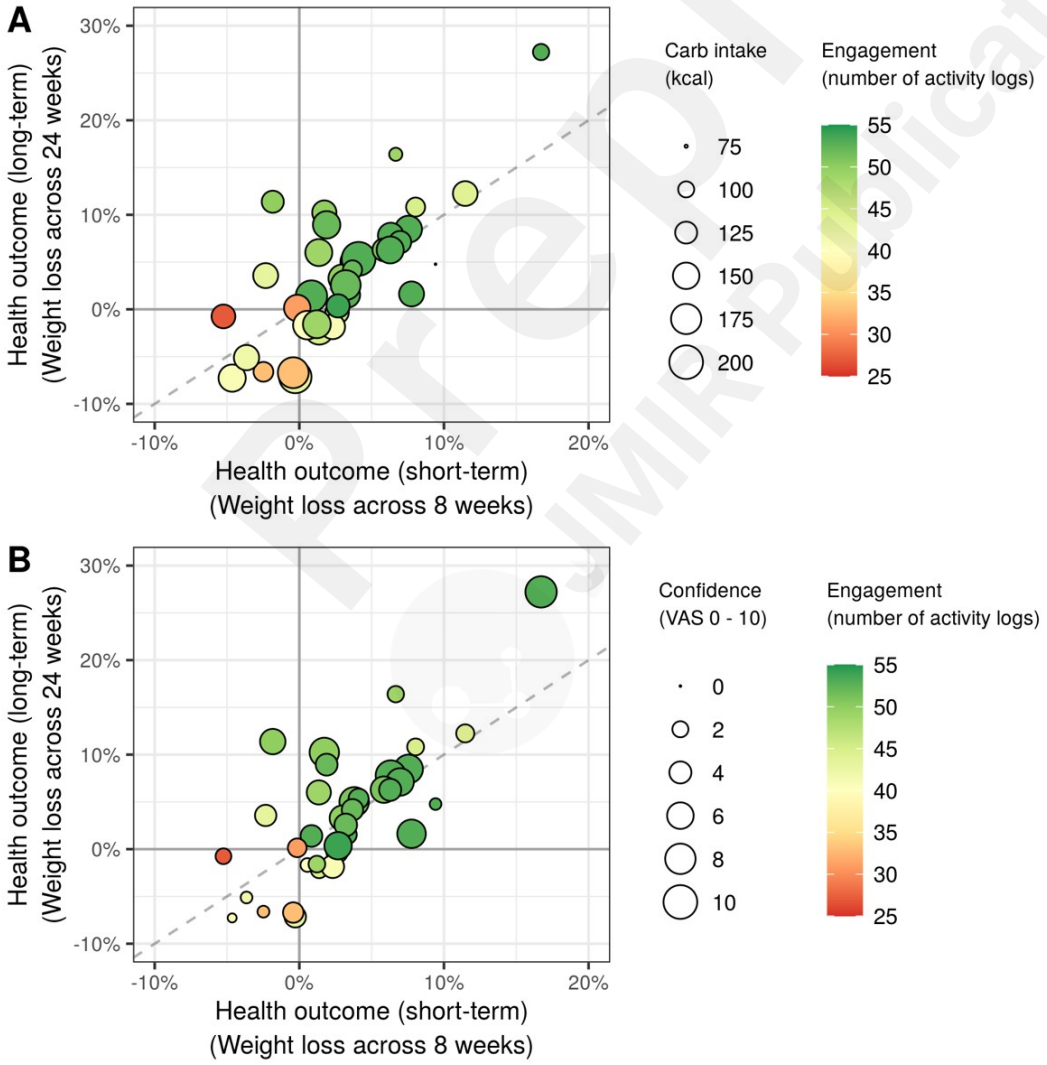


Figure 5. Two examples of common predictors between short-term and long-term health outcomes: carb intake and

confidence in digital phenotypes.

Regarding the model performance of the three prediction models, the machine-learning approaches successfully predicted the engagement rate (mean R^2 = 0416, SD = 0006), short-term weight change (mean R^2 = 0382, SD = 0.015), and long-term weight change (mean R^2 = 0590, SD = 0011). Especially in predicting long-term weight change, approximately 59% of the outcome variance is explained by the prediction model. In sum, these model performances suggest that the multivariate profiles in conventional and digital phenotypes provide the phenotypes that are significantly associated with engagement and health outcomes.

Discussion

Using a machine-learning approach based on elastic net regression, we successfully demonstrated the conceptual paradigm's applicability with complex dimensions of how in-app engagement is formed and affects health outcomes. This study showed that mobile applications' engagement was significantly associated with health outcomes, even four months after the cessation of digital interventions. We also found that both conventional motivation (before the intervention) and in-app motivation (during the intervention) were closely related to both engagement and clinical outcomes. Multiple aspects of motivation before and during the intervention could be used to predict engagement and health outcomes. Furthermore, both engagement and health outcomes are associated with multivariate psychological indices patterns, such as behavioral, cognitive, emotional, and motivational components, driven by regularized multivariate profiles obtained with the machine learning approach. From the results, we conclude that individuals' psychological states are the primary elements that influence engagement and health outcomes.

This study makes a clear implication on how engagement with apps influences clinical outcomes. Our finding that a higher frequency of logging into an app drives more significant improvements in health outcomes during the active intervention period is consistent with previous studies [20,21]. A notable finding in this study, however, is that those who logged into the app more frequently also showed more favorable health outcomes after the cessation of the active intervention period. These results indicate that engagement is paramount to the app's potential effectiveness for behavior change, leading to a change in symptomatology. Thus, it is feasible for clinicians and users to predict their health outcomes according to the intensity of their participation in apps.

Digital interventions via apps are not the only realm in which engagement is an issue. Both face-to-face and digital interfaces encounter difficult problems in maintaining adherence and engagement with monitoring, medications, and psychotherapies [22]. Since digital therapeutics are beneficial to monitor and analyze real-time data and reach out to users without barriers in space and time; however, they are more applicable to offer immediate feedback and prevent attrition than face-to-face clinics. From this perspective, a previous meta-analysis has claimed that integrating a human factor into the treatment is an actionable strategy to alleviate the dropout rates in the digital intervention [23]. Our result is also supportive in that the number of messages (interaction frequency between the user and therapist) showed the highest positive standardized coefficient with the engagement with the app. Taken together, we suggest that human feedback is involved in the development of digital therapeutics to strengthen the engagement rate, leading to greater clinical efficacy.

For the first time, this study evaluated the multiple dimensions of motivation at two different periods: before (conventional motivation) and during the intervention (in-app motivation). The common predictors of both engagement

rate and health outcomes were in-app motivational phenotypes, referred to as satisfaction with the intervention, desire to improve health outcomes, and self-confidence. The level of self-esteem at baseline was also a common predictor of both engagement and health outcomes. Moreover, before implementing the intervention, the level of motivation was strongly related to health outcomes in both the short- and long-term courses. Altogether, these results suggest that motivation is the main component that determines engagement and health outcomes.

Previously, pragmatic qualities, systematic flow, satisfaction, usability, and aesthetics were known as the major contributors to digital therapeutics engagement [7,20,22]. These prior results only serve as a basis for preliminary hypotheses on what may force engagement with apps. Few studies have examined engagement based on individuals' interactions with various intervention elements such as frequency of access, an average of steps, article views, message views, and so on [21,24,25]. However, it is still challenging to establish a standardized approach to assess these phenotypes' engagement because of various factors such as diverse technological aspects, different intervention exposure times, and individual characteristics. Thus, we suggest measuring the multiple aspects of motivation directly before and during the intervention to predict dropout and give each participant individualized attention.

This is the first study to categorize diverse digital phenotypes into four different constructs: behavior, cognition, emotion, and motivation. This allows a comprehensive understanding of the nature of behavior change, which is closely related to the engagement and clinical outcomes of digital interventions. We suggest that the behavioral phenotypes (calorie density of food, snack time of the day, amount of food intake per meal, and frequency of message interactions with the therapists), emotional phenotypes (irritated, bored, and depressed), and motivational phenotypes (satisfaction, will, and confidence). However, none of the cognitive phenotypes was capable of the engagement rate in the app. The phenotypes predicting the health outcomes were similar but not identical to the engagement because the amount of nutritional intake was included instead of the amount of food intake per meal for the behavioral phenotypes, and depressive moods were excluded from the emotional phenotypes. These findings imply that not only users' physical participation in a specific target behavior (e.g., logging food diary, number of steps) and behavior in digital spaces (e.g., number of accesses) but also the user's psychological conditions (e.g., emotion and motivation) are relevant to engagement and clinical outcomes.

To the best of our knowledge, this is the first study to apply a machine-learning approach to provide relevant insights into improving both the adherence and clinical outcomes of digital interventions. We demonstrated the whole framework of how different types of phenotypes at baseline and during the intervention, carry out in-app engagement and health outcomes. We used machine-learning strategies with digital phenotypes to find an applicable model to predict intervention adherence for the first time. This is also the first study to examine the determinants of significant weight changes from digital interventions. Additionally, our first attempt to explore the phenotypes in two different periods (at baseline and during the intervention) and categorized them into four distinctive dimensions (behavior, cognition, emotion, and motivation) present more comprehensive perceptions of engagement mechanisms clinical outcomes. Finally, this study applied two specific methods, in-app and an online survey, for the first time, to collect sufficient data, which led us to explore various components attaining favorable solutions for the issue of engagement and clinical efficacy in digital therapeutics.

This study had several limitations. First, all participants received cognitive behavioral therapy, so it lacked a control group that did not receive any intervention. Second, the number of participants was relatively small, which might not be sufficient for reliable interpretation. However, as we extracted multivariate profiles to predict engagement and health

outcomes, we remedied the shortage by using a machine learning approach. Furthermore, as this study explores the challenging concept of digital interventions, a small number of participants are still tolerable to apply the machine learning analysis [26]. Third, the experiment did not track longitudinal changes in health outcomes in the app.

Using a machine learning approach, we successfully established and validated an intuitive analytic strategy and provided visualization with a multiplex components paradigm of causality underlying digital psychotherapy on health outcomes. Our results revealed a key mechanism of psychological features interacting with multiple dimensions of motivation, which induce engagement in the app, enhancing clinical efficacy. We expect that this study will play a significant role in establishing the most practical and effective mHealth intervention model, a vital insight for precision digital medicine.

Contributors

MK led and designed the conceptual framework of mHealth components. JY conducted the data analysis and performed statistical/machine learning analysis. HJC conceived the project and provided valuable insights for devising and visualizing the concept of this project. WYA supervised the data analysis and served as a technical advisor for this project. MK and JY wrote the manuscript and edited the manuscript. HJC and WYA contributed to manuscript writing. All authors approved the final version of the manuscript for submission.

Data sharing

The data that support the findings of this study are available from our Github repository, https://github.com/CCS-Lab/digital-phenotypes-mHealth_2020.

Acknowledgments

We would like to thank Youngin Kim, the chief executive manager of Noom, for supporting this study by sharing the company's advanced digital technologies.

Conflicts of Interest

None declared.

Funding – This study was supported by a grant from the National Research Foundation of Korea (NRF), funded by the Korean Government, Ministry of Science and ICT (MSIT) (No. NRF-2018R1A5A2025964 to HJC, and no. NRF-2018R1C1B3007313 and No. NRF-2018R1A4A1025891 to WYA), the Institute for Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korean government (MSIT) (No. 2019-0-01367, BabyMind, to WYA), and the Creative-Pioneering Researchers Program through Seoul National University (to WYA).

Abbreviations

dCBT = Digital Cognitive Behavioral Therapy
RCT = Randomised Controlled Trial
BMI = Body Mass Index
ICU = Intensive Care Unit
SIMS = Situational Motivation Scale
BSQ-8C = Body Shape Questionnaire
K-BDI-II = Beck Depression Inventory-II in Korean
TAI = Trait-Anxiety Inventory

RSES = Rosenberg Self Esteem Scale
DEBQ = Dutch Eating Behavior Questionnaire
ATQ-30 = Automatic Thoughts Questionnaire
VAS = Visual Analog Scale
LOOCV = Leave-one-out-cross-validation
RMSE = Root Mean Squared Errors

CDI = Calorie Density Index

References

1. Jain SH, Powers BW, Hawkins JB, Brownstein JS. The digital phenotype. Nat Biotechnol. 2015;33: 462-463. doi:10.1038/nbt.3223, PubMed: 25965751.
2. Gunn JF, 3rd, Lester D. Using google searches on the internet to monitor suicidal behavior. J Affect Disord. 2013;148: 411-412. doi:10.1016/j.jad.2012.11.004, PubMed: 23182592.
3. Laranjo L, Arguel A, Neves AL, Gallagher AM, Kaplan R, Mortimer N, et al. The influence of social networking sites on health behavior change: a systematic review and meta-analysis. J Am Med Inform Assoc. 2015;22: 243-256. doi:10.1136/amiajnl-2014-002841, PubMed: 25005606.
4. Edney S, Ryan JC, Olds T, Monroe C, Fraysse F, Vandelanotte C, et al. User Engagement and Attrition in an App-Based Physical Activity Intervention: Secondary Analysis of a Randomized Controlled Trial. J Med Internet Res. 2019;21: e14645. doi:10.2196/14645, PubMed: 31774402.
5. Tolou-Shams M, Yonek J, Galbraith K, Bath E. Text Messaging to Enhance Behavioral Health Treatment Engagement Among Justice-Involved Youth: Qualitative and User Testing Study. JMIR MHealth UHealth. 2019;7: e10904. doi:10.2196/10904, PubMed: 30950808.
6. O'Brien HL, TE. What is user engagement? A conceptual framework for defining user engagement with technology. J Am Soc Inf Sci Technol. 2008;59: 938-955. doi:10.1002/asi.20801.
7. Wiebe EN, Lamb A, Hardy M, Sharek D. Measuring engagement in video game-based environments: Investigation of the User Engagement Scale. Comput Hum Behav. 2014;32: 123-132. doi:10.1016/j.chb.2013.12.001.
8. Przybylski AK, Ryan RC RM. A motivational model of video game engagement. Rev Gen Psychol. 2010;14: 154-166., doi, doi:10.1037%2Fa0019440.
9. Carpenter J, Crutchley P, Zilca RD, Schwartz HA, Smith LK, Cobb AM, et al. Seeing the "Big" Picture: Big Data Methods for Exploring Relationships Between Usage, Language, and Outcome in Internet Intervention Data. J Med Internet Res. 2016;18: e241. doi:10.2196/jmir.5725, PubMed: 27580524.
10. Al-Asadi AM, Klein B, Meyer D. Posttreatment attrition and its predictors, attrition bias, and treatment efficacy of the anxiety online programs. J Med Internet Res. 2014;16: e232. doi:10.2196/jmir.3513, PubMed: 25316533.
11. Marsch LA. Opportunities and needs in digital phenotyping. Neuropsychopharmacology.

2018;43: 1637-1638. doi:10.1038/s41386-018-0051-7, PubMed: 29703995.

12. Ben-Zeev D, Scherer EA, Brian RM, Mistler LA, Campbell AT, Wang R. Use of Multimodal Technology to Identify Digital Correlates of Violence Among Inpatients With Serious Mental Illness: A Pilot Study. Psychiatr Serv. 2017;68: 1088-1092. doi:10.1176/appi.ps.201700077, PubMed: 28669285.
13. Martinez-Martin N, Insel TR, Dagum P, Greely HT, Cho MK. Data mining for health: staking out the ethical territory of digital phenotyping. npj Digit Med. 2018;1: 1-5. doi:10.1038/s41746-018-0075-8, PubMed: 31211249.
14. Lutz W, Hofmann SB SG, Fisher AJ, Husen K, Rubel JA. Using network analysis to predict treatment dropout in patients with mood and anxiety disorders: A methodological proof-of-concept study. *Scientific reports*. **8**, 1-9., doi:https://doi.org/10.1038/s41598-018-25953-0 (2018).
15. Johnson AE, Ghassemi MM, Nemati S, Niehaus KE, Clifton DA, Clifford GD. Machine Learning and Decision Support in Critical Care. Proc IEEE Inst Electr Electron Eng. *Proc IEEE Inst Electr Electron Eng*. 2016;104: 444-466. doi:10.1109/JPROC.2015.2501978, PubMed: 27765959.
16. Kim M, Kim Y, Go Y, Lee S, Na M, Lee Y, et al. Multidimensional Cognitive Behavioral Therapy for Obesity Applied by Psychologists Using a Digital Platform: Open-Label Randomized Controlled Trial. JMIR MHealth UHealth. 2020;8: e14817. doi:10.2196/14817, PubMed: 32352391.
17. Zou H, Hastie T. Regularization and variable selection via the elastic net. J Royal Statistical Soc B. 2005;67 **2**: 301–320. doi:10.1111/j.1467-9868.2005.00503.x.
18. Ahn WY, Vassileva J. Machine-learning identifies substance-specific behavioral markers for opiate and stimulant dependence. Drug Alcohol Depend. 2016;161: 247-257. doi:10.1016/j.drugalcdep.2016.02.008, PubMed: 26905209.
19. Ahn WY, Kishida KT, Gu X, Lohrenz T, Harvey A, Alford JR, et al. Nonpolitical images evoke neural predictors of political ideology. Curr Biol. 2014;24: 2693-2699. doi:10.1016/j.cub.2014.09.050, PubMed: 25447997.
20. Graham AK, Lattie EG, Mohr DC. Experimental Therapeutics for Digital Mental Health. JAMA Psychiatry. 2019;Volume 76: 1223–1224. doi:10.1001/jamapsychiatry.2019.2075, PubMed: 31433448.
21. Taki S, Lymer S, Russell CG, Campbell K, Laws R, Ong KL, et al. Assessing User Engagement of an mHealth Intervention: Development and Implementation of the Growing

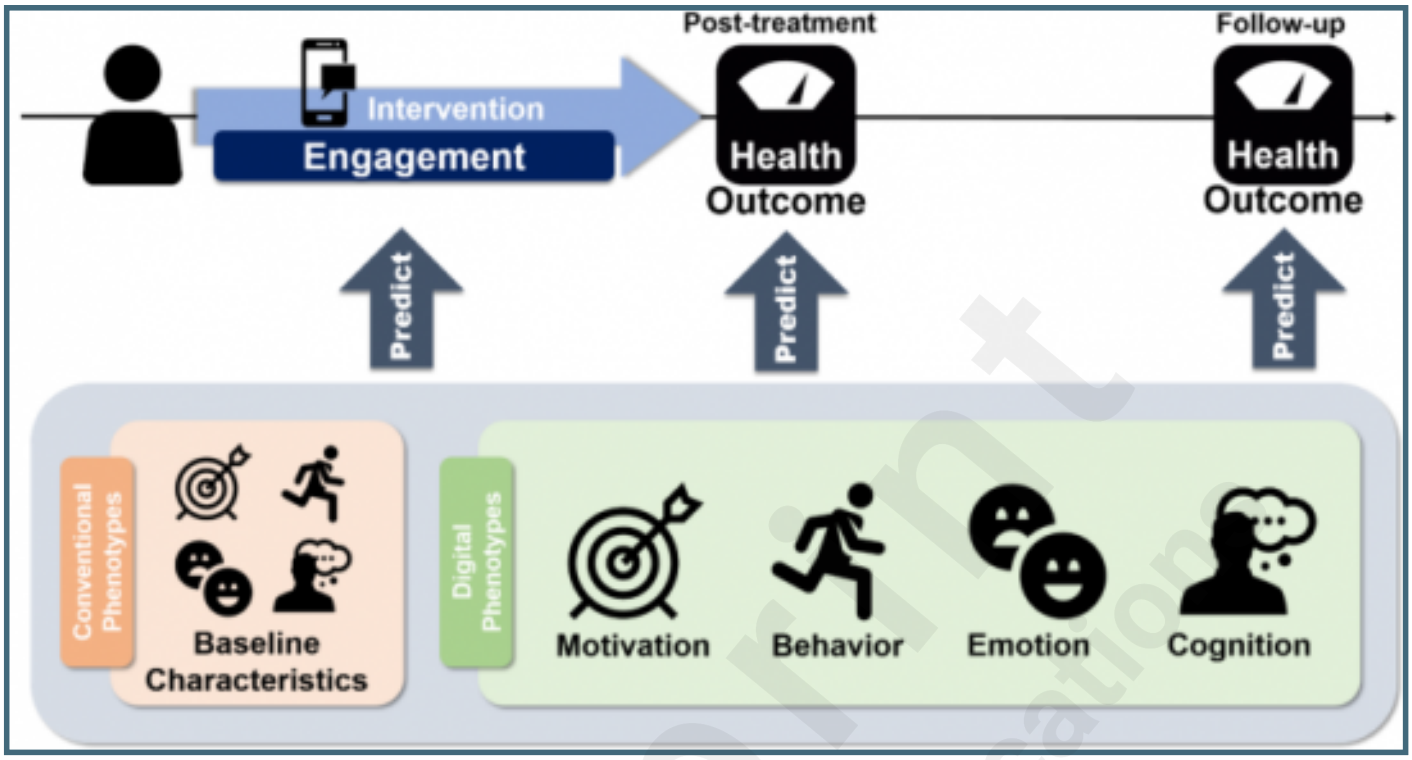
Healthy App Engagement Index. JMIR MHealth UHealth. 2017;5: e89. doi:10.2196/mhealth.7236, PubMed: 28663164.

22. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory, and improvements. Evid Based Ment Health. 2018;21: 116-119. doi:10.1136/eb-2018-102891, PubMed: 29871870.
23. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: A systematic review and meta-analysis. J Affect Disord. 2020;263: 413-419. doi:10.1016/j.jad.2019.11.167, PubMed: 31969272.
24. Baltierra NB, Muessig KE, Pike EC, LeGrand S, Bull SS, Hightow-Weidman LB. More than just tracking time: Complex measures of user engagement with an internet-based health promotion intervention. J Biomed Inform. 2016;59: 299-307. doi:10.1016/j.jbi.2015.12.015, PubMed: 26732997.
25. Davies C, Corry K, Van Itallie A, Vandelanotte C, Caperchione C, Mummery WK. Prospective associations between intervention components and website engagement in a publicly available physical activity website: the case of 10,000 Steps Australia. J Med Internet Res. 2012;14: e4. doi:10.2196/jmir.1792, PubMed: 22260810.
26. Tazawa Y, Liang KC, Yoshimura M, Kitazawa M, Kaise Y, Takamiya A, et al. Evaluating depression with multimodal wristband-type wearable device: screening and assessing patient severity utilizing machine-learning. Heliyon. 2020;6: e03274. doi:10.1016/j.heliyon.2020.e03274, PubMed: 32055728.

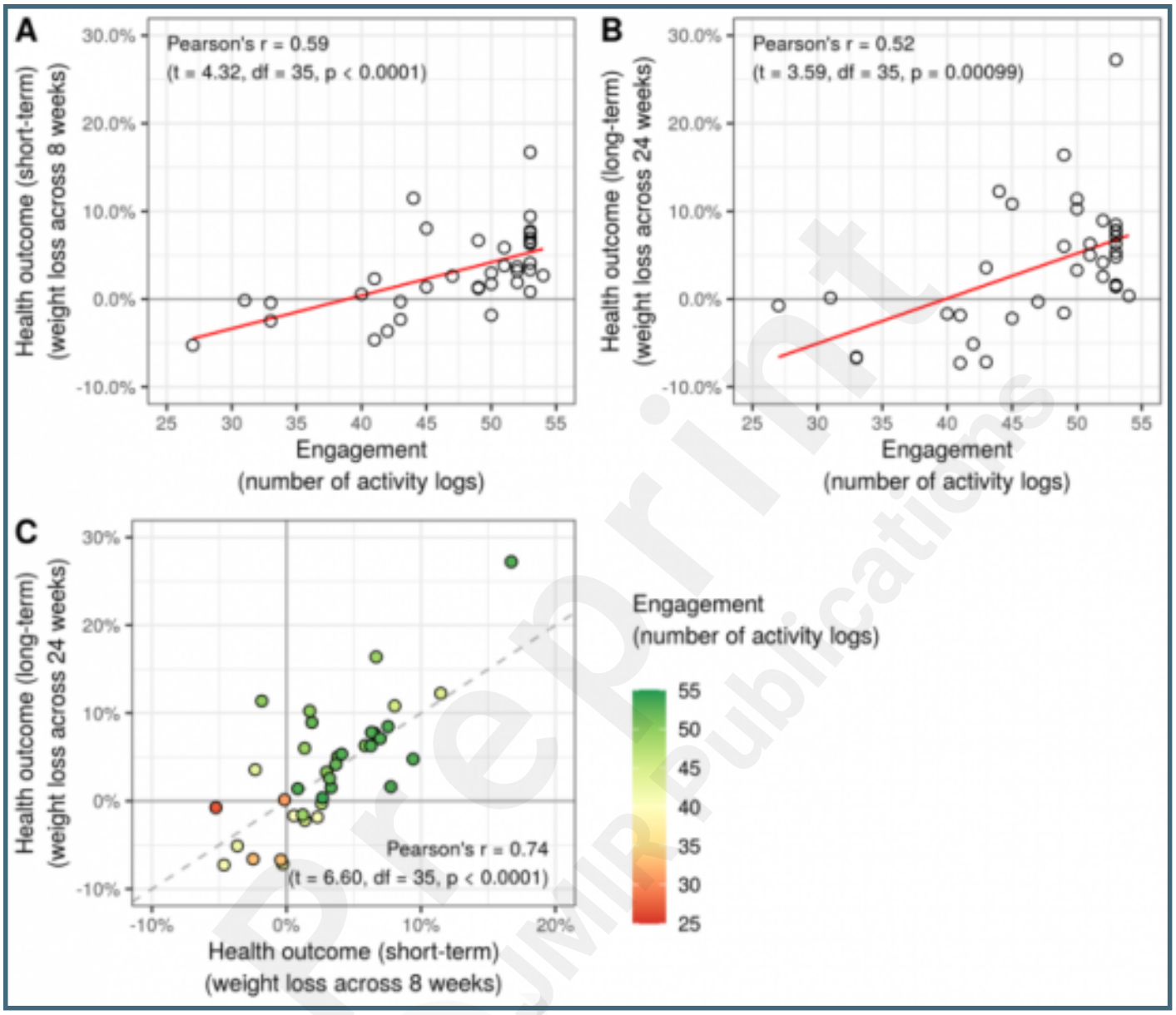
Supplementary Files

Figures

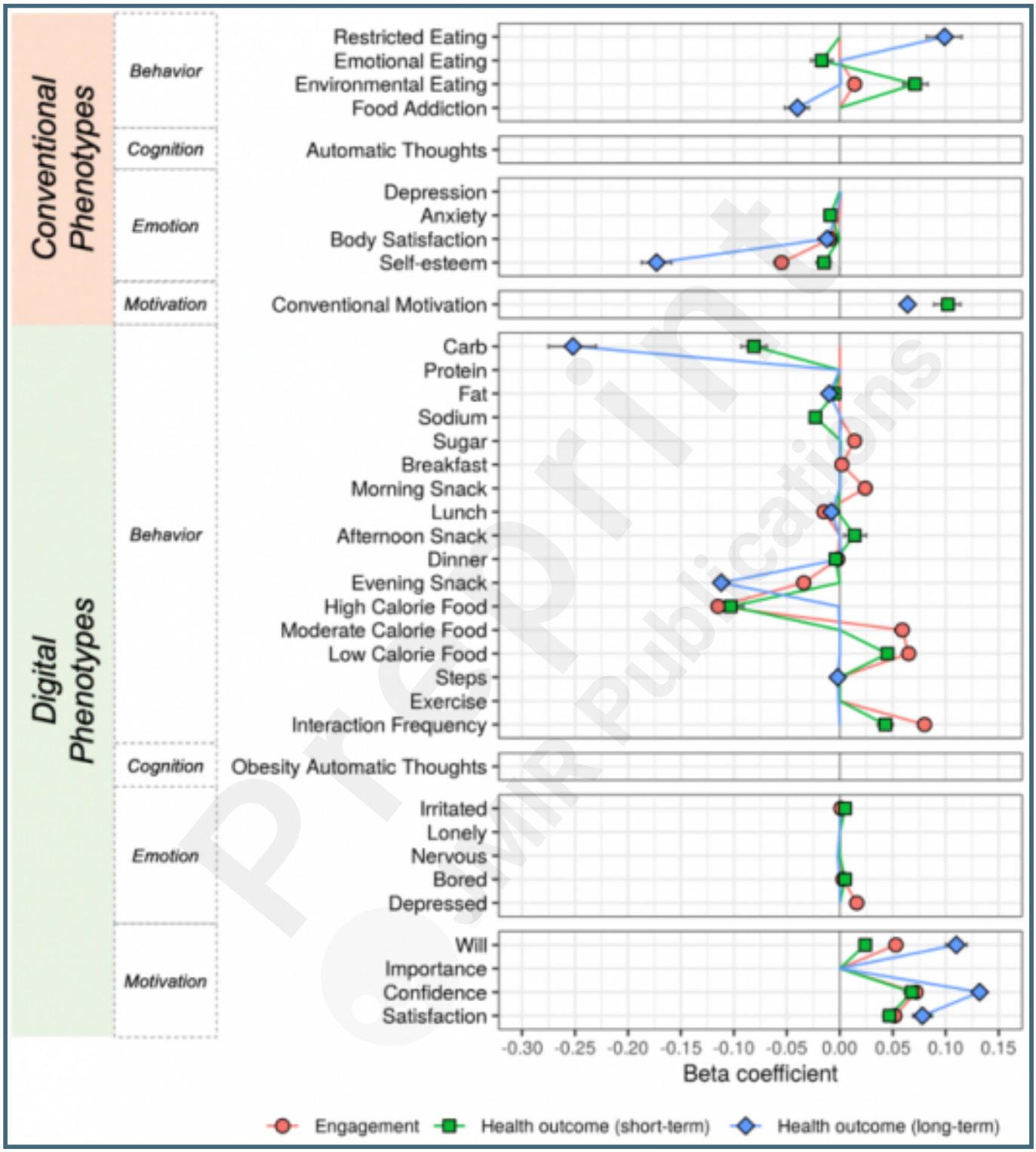
A conceptual framework of mHealth components and examples of digital phenotypes.



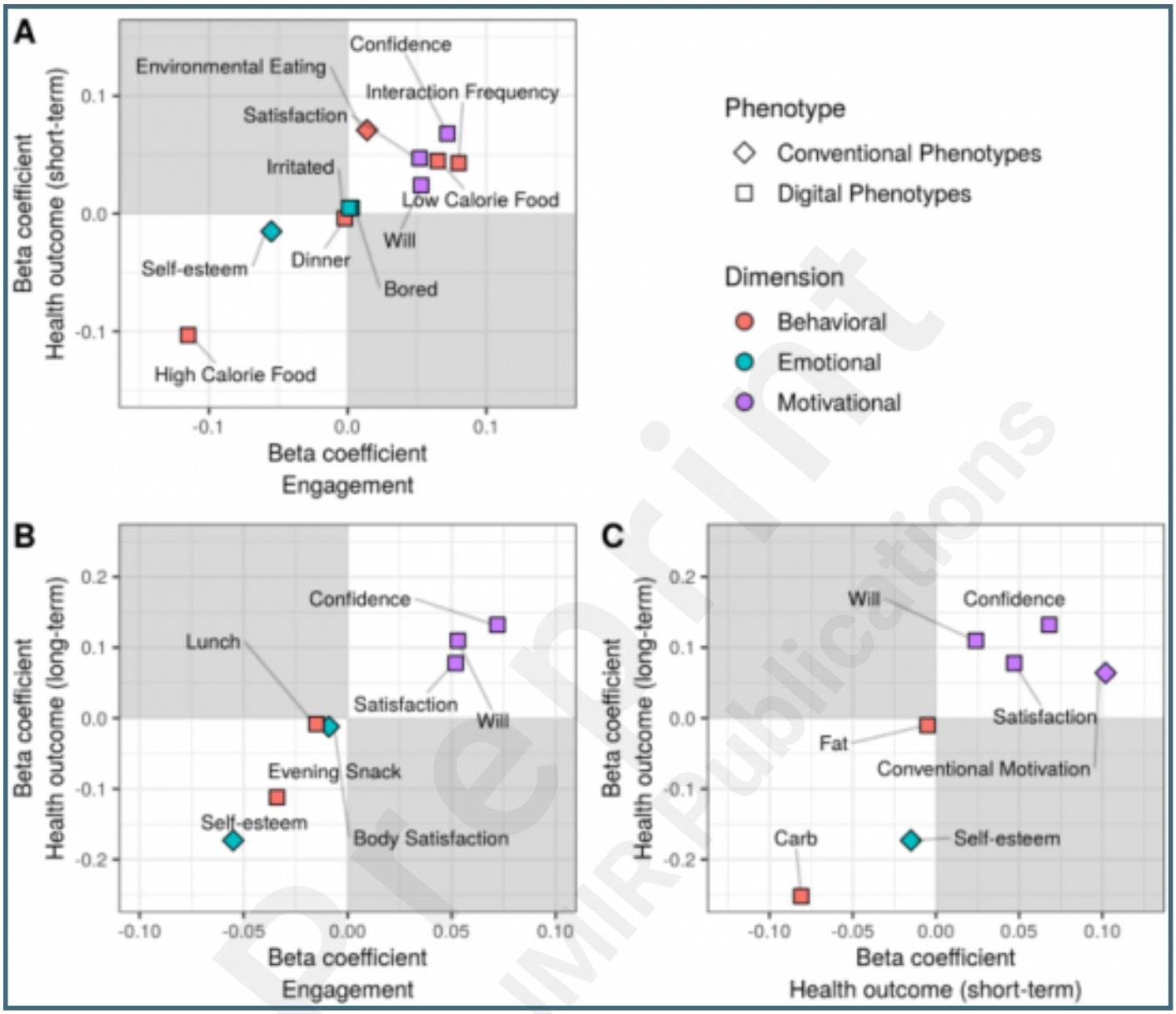
Relationships between engagement and health outcomes. The health outcome larger than zero indicates weight loss compared to baseline.



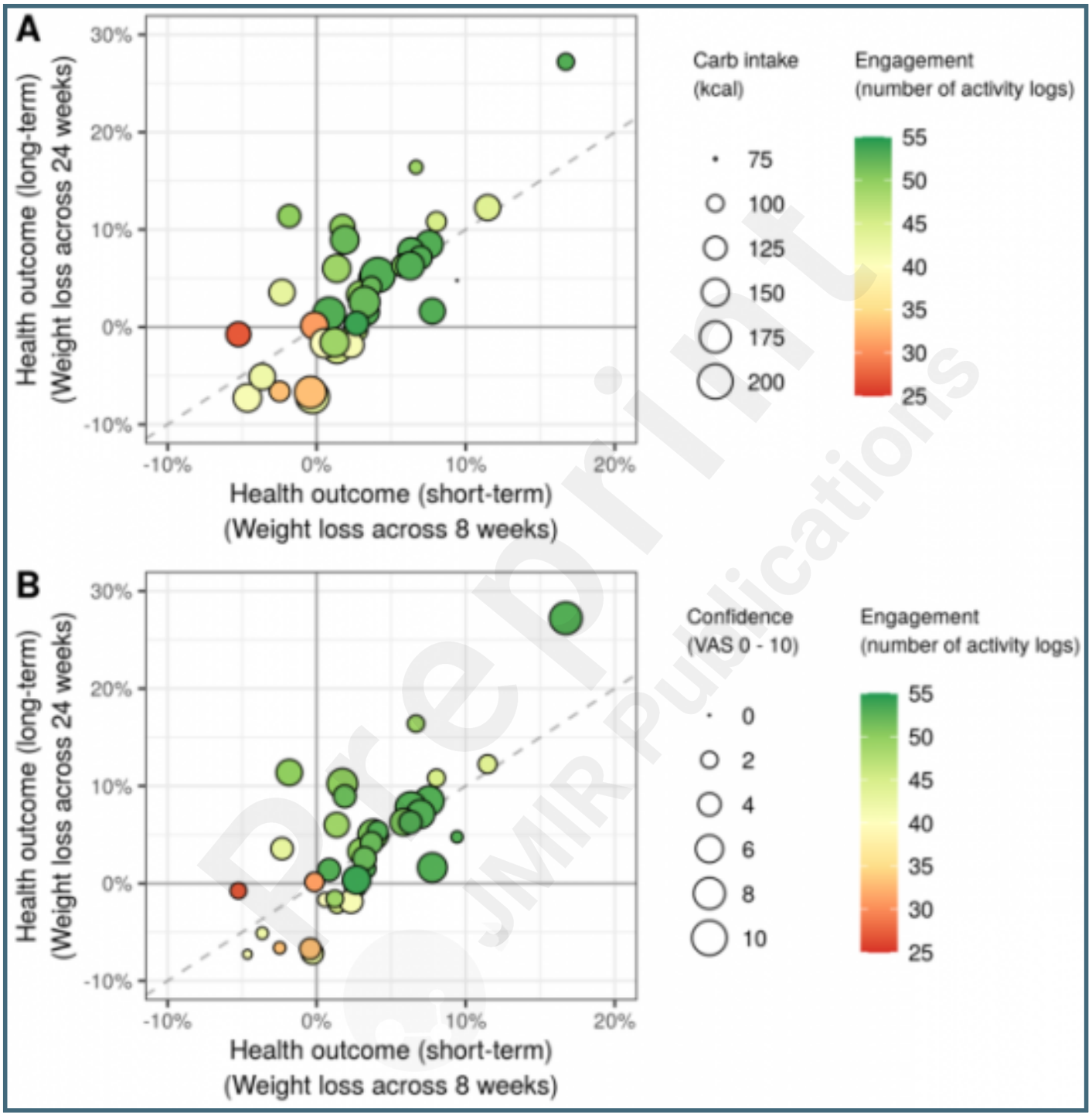
Multivariate patterns of conventional and digital phenotypes for predicting engagement (red) as well as short-term (green) and long-term (blue) health outcomes.



Common predictors between engagement and health outcomes.



Two examples of common predictors between short-term and long-term health outcomes: carb intake and confidence in digital phenotypes.



Multimedia Appendixes

Supplementary materials.
URL: https://asset.jmir.pub/assets/1dff44b24be3a92372e2a11ebd104cf9.docx

CONSORT (or other) checklists

CONSORT-EHEALTH (V 1.6.1) - Submission/Publication Form.
URL: https://asset.jmir.pub/assets/0bf2945d7004b9f1fa060d2eb95b3e84.pdf