

# Using a complexity science approach to evaluate the effectiveness of just-in-time adaptive interventions: A meta-analysis

DIGITAL HEALTH  
Volume 9: 1–13  
© The Author(s) 2023  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/20552076231183543  
journals.sagepub.com/home/dhj



Zhan Xu<sup>1</sup>  and Eline Smit<sup>2</sup> 

## Abstract

**Objective:** Just-in-time adaptive interventions (JITAs), which allow individuals to receive the right amount of tailored support at the right time and place, hold enormous potential for promoting behavior change. However, research on JITAs' implementation and evaluation is still in its early stages, and more empirical evidence is needed. This meta-analysis took a complexity science approach to evaluate the effectiveness of JITAs that promote healthy behaviors and assess whether key design principles can increase JITAs' impacts.

**Methods:** We searched five databases for English-language papers. Study eligibility required that interventions objectively measured health outcomes, had a control condition or pre-post-test design, and were conducted in the real-world setting. We included randomized and non-randomized trials. Data extraction encompassed interventions' features, methodologies, theoretical foundations, and delivery modes. RoB 2 and ROBINS-I were used to assess risk of bias.

**Results:** The final analysis included 21 effect sizes with 592 participants. All included studies used pre- and post-test design. A three-level random meta-analytic model revealed a medium effect of JITAs on objective behavior change ( $g = 0.77$  (95% confidence interval (CI); 0.32 to 1.22),  $p < 0.001$ ). The summary effect was robust to bias. Moderator analysis indicated that design principles, such as theoretical foundations, targeted behaviors, and passive or active assessments, did not moderate JITAs' effects. Passive assessments were more likely than a combination of passive and active assessments to relate to higher intervention retention rates.

**Conclusions:** This review demonstrated some evidence for the efficacy of JITAs. However, high-quality randomized trials and data on non-adherence are needed.

## Keywords

Complexity science, complexity theory, just-in-time adaptive interventions, JITAs, meta-analysis

Submission date: 25 November 2022; Acceptance date: 5 June 2023

Prior research confirms abundant variation in using health interventions to promote behavior change and demonstrates the challenges in translating lab findings into the real-world setting.<sup>1,2</sup> This is illustrated by many practice-level interventions showing only limited success in behavior change, as based on recent systematic reviews.<sup>3,4</sup> According to the complexity science approach (complexity theory), interventions effectively addressing barriers and encouraging desired behaviors in one setting may fail to generate similar results elsewhere since “opportunities for change vary at each practice.”<sup>2</sup> Thus, interventions that

are individually tailored and are developed to be adaptive to changing contexts may be more effective in promoting healthy behaviors than non-adaptive interventions.<sup>3,5,6</sup> In this vein, just-in-time adaptive interventions (JITAs),

<sup>1</sup>School of Communication, Northern Arizona University, Flagstaff, AZ, USA  
<sup>2</sup>University of Amsterdam, Amsterdam, The Netherlands

### Corresponding author:

Zhan Xu, School of Communication, Northern Arizona University, 700 S. Knoles Dr, Flagstaff, AZ 86011, USA.  
Email: zhan.xu@nau.edu



which allow individuals to receive the right amount of tailored support at the right time and place, hold enormous potential for promoting the desired behavior change.

A JITAI is an emerging intervention design that adjusts the timing and intensity of message delivery based on an individual's evolving internal and contextual state with the goal to provide support at the moment when individuals need it most.<sup>7</sup> For example, a JITAI for preventing alcohol use disorders sends individuals tailored messages based on the risk level of alcohol use at a particular moment. If an individual is close to a high-risk location, such as a bar that is frequently visited in the past, the risk level is considered high and the person may receive warning messages. If the risk level is deemed low, individuals may not receive messages.<sup>8</sup> So far, JITAIs have mainly been used to change behaviors related to physical activity,<sup>9</sup> substance use,<sup>8</sup> and obesity.<sup>10</sup> However, research on JITAIs' implementation and evaluation is still in its early stages, and more empirical evidence regarding the use of JITAIs is needed.<sup>11</sup> In this study, we will take a complexity science approach to evaluate the effectiveness of JITAIs that promote healthy behaviors and assess whether key design principles and components can increase JITAIs' impacts.

## Literature review

Complexity science focuses on dynamic issues and problems that are often unpredictable and consist of interconnected relationships. It is a collection of theories and conceptual tools from different disciplines.<sup>12</sup> Two aspects of the complexity science approach that are closely related to health interventions include mathematical complexity and aggregate complexity. Both explain the promise of JITAIs for health behavior change and provide insights on health intervention design.<sup>2</sup>

### Mathematical complexity and aggregate complexity

Mathematical complexity refers to various complex systems that exist in behavior change.<sup>13</sup> As opposed to complex systems, linear or deterministic systems assume that the same results will occur over and over again, as long as the intervention starts with the same condition and applies the same equation.<sup>2</sup> For instance, if a lab study demonstrates the effectiveness of an anti-smoking campaign in reducing tobacco use among a certain population, then, when recruiting a similar population, using the same campaign, and following the same procedure, another study should generate a similar effect. In contrast, complex or non-linear systems refer to systems that often follow random and unpredictable trajectories.<sup>2</sup> As a result, an intervention that works in a lab study may not work in another setting. For example, in the natural setting, smoking decisions may be based not only on the anti-

smoking campaign but also—and perhaps primarily—on people's changing internal and contextual state.

Although it is impractical, if not impossible, to fully characterize all internal and external factors that drive behavior change, one can capture opportunities and moments that are most likely to influence behavioral decisions. For instance, it is challenging to predict all aspects that affect decisions in smoking cessation, yet it is feasible to detect high-risk locations that may lead individuals to smoke again<sup>14</sup> or identify hand gestures that suggest these individuals are about to smoke.<sup>15</sup> Both high-risk locations and hand gestures are tailoring variables and can be used to determine when and which intervention messages are delivered.<sup>5</sup>

Aggregate complexity refers to the interplay among all the various factors that affect behaviors.<sup>13</sup> Dynamic interactions among internal and external factors may lead to unexpected outcomes. For example, at the beginning of a real-world intervention, differences in initial conditions, such as “the larger health care system to which each belongs and the social and economic conditions within the local community,”<sup>2</sup> may result in a disproportionately large effect on forthcoming events. During the intervention, unplanned occurrences (e.g. distractions) may take place and disrupt the decision-making process. Moreover, individuals may modify their behaviors based on the feedback they receive from or other interactions with the intervention.<sup>13</sup> Since the interplay among various factors cannot be fully foreseen, adaptation is needed in intervention design. Adaptation refers to the ongoing use of information to alter the content, timing, and frequency of messages and is emphasized in JITAI design.<sup>16</sup> It requires monitoring (a) whether individuals are “in a state that requires support,” (b) what type (or amount) of support is needed given the individual's state, and (c) whether providing this support has the potential to disrupt the desired process.<sup>5</sup>

### Passive and active assessments of tailoring variables

Based on the complexity science approach, JITAIs aim to adjust intervention timing and content to individuals' changing internal and contextual states by targeting certain tailoring variables to capture the opportunities for behavior change. Tailoring variables allow the timing of intervention delivery to be modified. Timing “reflects the particular condition(s) that someone or something is in at a particular point or period of time.”<sup>5</sup> It is based on the events or conditions, such as when individuals approximate a high-risk location, which may be unexpected. JITAIs will deliver intervention messages when “states of opportunity” or the “right time” occurs.<sup>5</sup> For instance, if the time when individuals approximate a high-risk location is deemed to be the tailoring variable and indicative of the right timing to send out a risk warning message, participants will receive instant messages when they approach these locations. If individuals stay

in low-risk locations, they will not receive any messages or will receive different types of messages.

To ensure timely and frequent adaption, JITAIs often use mobile devices, which allow people to access interventions anytime and anywhere.<sup>17</sup> Another advantage of using mobile devices to deliver JITAIs is that the sound or vibration associated with mobile messages can capture an individual's attention immediately.<sup>18</sup> Besides mobile devices, some JITAIs use sensors (e.g. accelerometer, camera) either with mobile devices or separately to capture raw data (i.e. heartbeat, walking distance, etc.) and utilize machine learning algorithms to translate sensor-based raw data to classify health outcomes, such as stress episodes,<sup>19</sup> smoking puffs,<sup>20</sup> and physical activities.<sup>21</sup>

Sensors and machine learning algorithms make passive assessments of tailoring variables feasible. Passive assessments refer to assessments of tailoring variables passively without the active engagement of participants. For example, participants' locations can be passively monitored and used to determine when to send an alert.<sup>8</sup> Besides passive assessments, some JITAIs also use active assessments to measure tailoring variables. Active assessments require participants to frequently self-report their thoughts and behaviors. For example, participants self-report three times a day about their difficulties performing tasks to regulate their moods, improve their sleep, etc.<sup>22</sup> However, active assessments are more prone to subjective interpretation and may cause fatigue when participants need to repeatedly answer similar surveys multiple times a day.<sup>5</sup> A recent study showed that a combination of passive and active assessments may be more persuasive than passive assessments alone.<sup>11</sup> This meta-analysis will examine the moderating effects of passive and active assessments of tailoring variables on JITAIs' effectiveness.

### *JITAI with or without a theoretical basis*

Besides being tailored in terms of delivery timing and frequency, the content of JITAIs is also often adapted to the dynamic contexts an individual might find themselves in. For example, JITAIs' messages have been tailored to participants' stages of behavioral change<sup>14</sup> or possibilities for performing a risky behavior.<sup>23</sup> Thus, several recent JITAIs have been developed according to motivational and/or socio-cognitive theories, such as utilizing the concept of self-efficacy that is emphasized in socio-cognitive theories.<sup>24,25</sup> These theories explain how to achieve positive health behavior change. However, to date, many JITAIs have been developed with little use of theories.<sup>5</sup>

Theoretically grounded JITAIs have the potential to generate better outcomes than those lacking a theoretical basis.<sup>26</sup> However, most existing theoretical frameworks are not dynamic and only treat underlying mechanisms that drive behavioral change as linear or deterministic

systems instead of as complex systems.<sup>27</sup> Several acknowledge the existence of complex systems. For example, behavioral motivation theories highlight the importance of addressing "states of opportunity" at the right time<sup>26,28</sup> or recommend breaking long-term goals into short-term specific activities and providing timely guidance.<sup>29,30</sup> However, these theories do not explain *how* various factors may work together and change over time.<sup>5</sup> Thus, it is critical to investigate whether existing theories used in JITAIs are effective in promoting healthy behaviors—therefore, this meta-analysis will examine whether the use of theories in JITAIs moderates the effects of JITAIs.

### *Retention*

Furthermore, many theoretical frameworks do not consider fatigue, countereffects, or resistance to persuasion,<sup>31</sup> which may cancel interventions' effects or even backfire. Due to frequent delivery of interventions and assessments as well as strong demand for participants' time and effort, JITAIs may have the potential to cause fatigue or non-adherence, which could result in lower retention rates. Retention rates refer to the percentage of users who continue to use JITAIs over a given period of time. It is a measure of how well a JITAI is able to keep its users engaged and satisfied.<sup>5</sup> As indicated in prior sections, JITAIs which require less time or effort from the participants, such as using passive assessments, may lead to higher retention rates than those that require longer time or active assessments.<sup>8</sup> Besides, theoretically grounded JITAIs may generate better outcomes, such as increasing retention rates and engagement, than JITAIs without a theoretical basis.<sup>26</sup> In order to examine if key design principles and components impact retention rates, besides analyzing the overall effectiveness of JITAIs and key moderators, this meta-analysis will also investigate if theories, assessments of tailoring variables, and intervention time affect retention rates.

### *Other potential moderators*

As the concept of aggregate complexity indicates, differences in initial conditions can largely affect forthcoming events. As initial conditions, such as social demographic factors, first exist when complex systems are established, they may ultimately influence behavioral outcomes at each practice. Besides initial conditions, the nature of the interactions between different agents "contribute to the uniqueness of each complex adaptive system."<sup>2</sup> For example, the interactions between interventions and people who are ill may be different from those between interventions and healthy people as various agents' "attitudes, skills, individual self-efficacy, and the nature of the interactions" are distinctive.<sup>32</sup> Since JITAIs often have various initial conditions and interactions, it is critical to consider these factors when evaluating the effectiveness

of JITAIs. In the current meta-analysis, we will therefore consider *social demographic factors* and *populations* as moderators of JITAIs' effects.

## Present review

Several systematic reviews have reviewed fundamental components of JITAIs and provided useful guidance on designing and testing JITAIs as well as addressing challenges.<sup>5,33</sup> To our knowledge, only one prior meta-analysis assessed the effectiveness of JITAIs, and it only included studies published by February 2018.<sup>11</sup> As an increasing number of JITAIs have been designed and tested in recent years, an updated meta-analysis is needed. Moreover, Wang and Miller's study combined JITAIs with ecological momentary interventions (EMIs) which utilize real-time assessments but do not necessarily provide tailored interventions based on these real-time assessments.<sup>34</sup> Since JITAIs and EMIs are not always the same,<sup>35</sup> the overall effectiveness of a combination of JITAIs and EMIs may not accurately represent the effects of JITAIs. A meta-analysis only focusing on JITAIs is therefore necessary. In the current meta-analysis, the primary goal is to examine the overall effectiveness of JITAIs on objective health outcomes. The secondary goal is to investigate if JITAIs' fundamental components, such as assessments, theories, retention rates, and intervention duration as well as initial conditions, such as demographic variables and population, moderate JITAIs' effects on health outcomes. We also aim to explore if theories, assessments, and intervention time affect retention rates.

## Method

### Literature search

Previous studies have used various terms to describe interventions adaptive to individuals' changing needs although not all of the interventions are adaptive in real time. Such terms include "dynamic tailoring"<sup>36</sup> and "intelligent real-time therapy."<sup>37</sup> In July 2021, keywords including just-in-time adaptive intervention, JITAI, dynamic tailoring, and intelligent real-time therapy were searched in the Academic Search Premier, CMMC, PsycINFO, PUBMED, and Medline via EBSCO databases. A completed checklist of PRISMA guidelines can be found in Appendix 1.

### Inclusion criteria and selection

Since the primary goal of this study is to examine the effectiveness of JITAIs on objective health outcomes and JITAIs need to be delivered in the natural environment to evaluate their effects,<sup>5</sup> we only included English-language studies that (a) measured health outcomes objectively (e.g. step

counts, sedentary time, etc.), (b) had a control condition or pre-post-test design, and (c) were conducted in the real-world setting. Studies were excluded if the intervention did not target individuals' changing internal and contextual states in real time. For instance, a study was excluded as it delivered dietary-related messages that were scheduled based on students' usual time to contemplate food choices (e.g. between lectures) instead of delivered just-in-time based on their changing states.<sup>38</sup> We contacted the authors if their studies met all the above criteria, but the authors did not provide data that can be used to calculate effect sizes in their publications.

The titles and abstracts of the articles were screened by an author and a research assistant separately, and any articles that were deemed potentially relevant were thoroughly assessed in full text. The two researchers then independently reviewed the full texts. In cases where there were disagreements, the two researchers discussed the reasons behind their inclusions and exclusions until a consensus was reached. The identification, screening, and inclusion process followed the PRISMA 2020 guideline.<sup>39</sup> Figure 1 shows the flow diagram.

### Statistical analysis

All analyses were performed using R (version 4.1.2). First, Hedges'  $g$ , which refers to the unbiased estimate of standardized mean differences,<sup>40</sup> was used as the effect size indicator. Both randomized trials and non-randomized studies were included. All randomized trials also reported pre-post-test results in the experimental conditions. Since it is not recommended to compute a summary value across randomized and non-randomized studies,<sup>41</sup> we utilized pre- and post-test means and standard deviations (SDs) or standard errors (SEs) to calculate Hedges'  $g$ . A positive mean difference between pre- and post-test behaviors indicates a positive behavior change, such as an increase in the number of steps from pre- to post-test conditions. For JITAIs that use a negative mean difference to indicate a positive behavior change (i.e. a decrease in the sedentary time), we reversed the computed sign of the effect sizes.<sup>41</sup>

Second, since several studies included different types of behaviors, effect sizes nested within a paper were not independent. To account for dependency among effect sizes, a three-level random meta-analytic model was used to calculate the overall effect size of JITAIs on health outcomes.<sup>42</sup> To test for heterogeneity, log-likelihood-ratio tests were performed to determine whether significant within-study variance and between-study variance exist. To assess the relationship between study sample sizes and effect sizes as well as publication bias, a funnel plot was created.<sup>43,44</sup> The rank correlation test was used to statistically determine the potential asymmetry of the funnel plots. The trim and fill method was applied to estimate an unbiased estimate

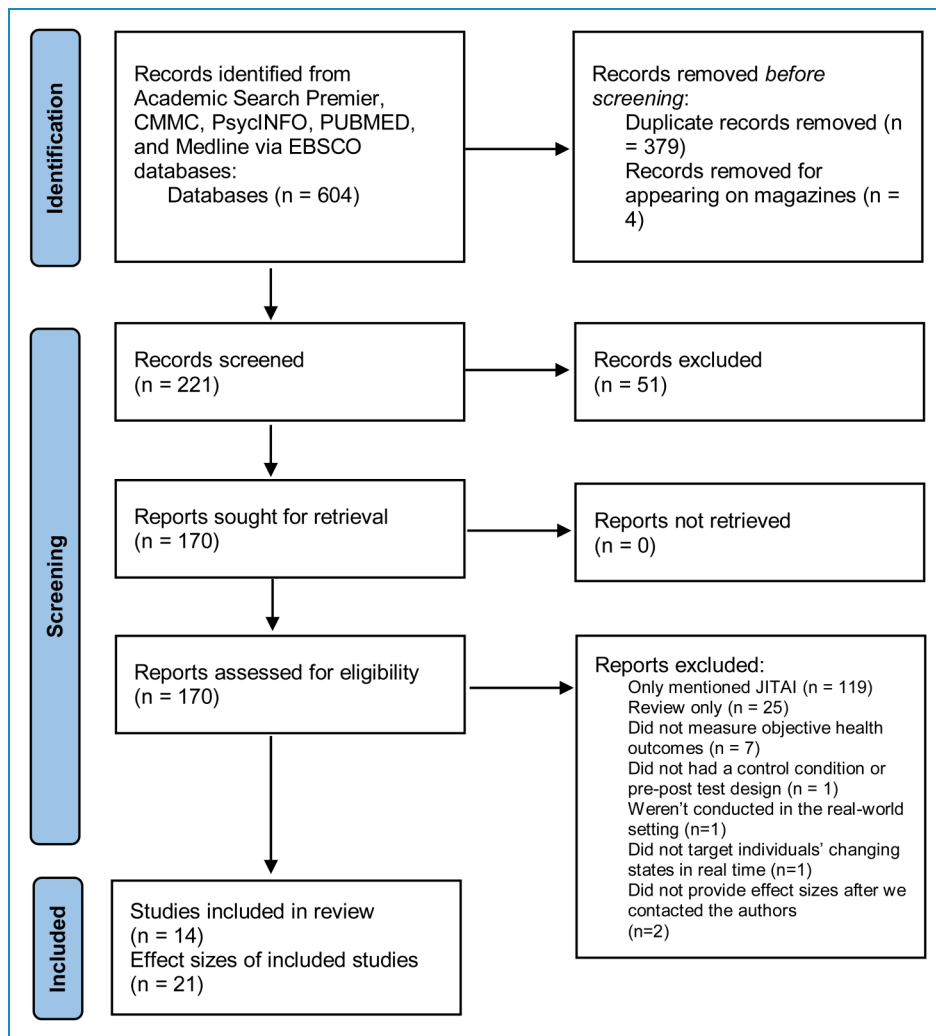


Figure 1. PRISMA 2020 flow diagram.

of the overall effect size, assuming publication bias resulted in the asymmetry of the funnel plot.<sup>45</sup> One of the causes of asymmetry is publication bias, but asymmetry can also be caused by other factors, such as variance in studies’ true effect size.<sup>44</sup> Thus, to assess the robustness of the overall effect size and further assess publication bias, fail-safe N was calculated to examine publication bias using Rosenthal’s<sup>46</sup> method. It calculates “the number of studies averaging null results that would have to be added to the given set of observed outcomes to reduce the combined significance level” (*p*-value) to a particular alpha level (e.g. 0.05 in the present study).<sup>46</sup> A forest plot was used to display the effect size, variance, weight of each study, and summary statistics (Figure 2).

Third, besides general health outcomes, to obtain more precise estimates of the effectiveness of the interventions, subgroup analyses based on different types of outcomes were conducted. Specifically, a three-level random meta-analytic model was used to calculate the effect size

of JITAIs that specifically focused on physical activities and the effect size of JITAIs that specifically targeted dietary behaviors.

Fourth, multivariate moderator analysis was used to examine whether assessments, theories, retention rates, intervention duration, population, behaviors, gender, age, and race significantly moderate the overall effects of JITAIs on health outcomes.

Finally, Pearson correlation analysis was used to assess if retention rates significantly correlated with intervention duration. The Wilcoxon rank sum test was performed to assess whether theories and assessments affected retention rates.

## Results

### Literature search results

We searched five databases and identified 604 articles. After removing duplicate records and non-peer-reviewed

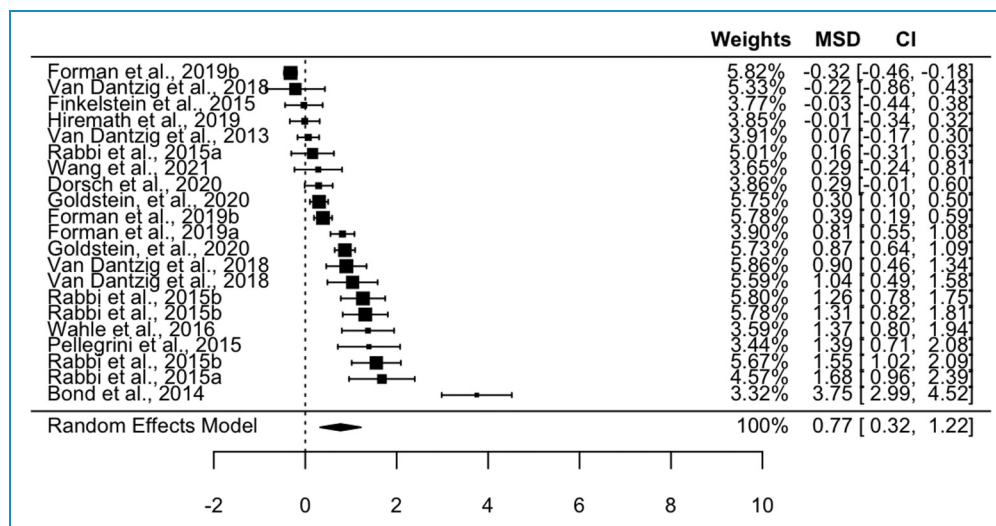


Figure 2. Forest plot.

articles, 170 full-text articles were checked. After removing studies on interventions that did not meet the inclusion criteria and/or did not report on effect sizes, 14 studies were included in the final analysis. As some of the studies examined different types of behaviors (e.g. physical activity and dietary behavior in Rabbi, Pfammatter, Zhang et al.<sup>47</sup>), the effect size for each behavior was coded separately. Thus, 21 effect sizes were included in the final analysis.

### Study characteristics and coding

There was a total of 592 participants across the studies in this review. Coded study characteristics included *total sample size* (with a range from 3 to 119), *mean age* (with a range from 28.3 to 53.1), *percentage of females* (with a range from 20% to 100%), *percentage of White participants* (with a range from 23% to 80%), *intervention duration* (with a range from 1 to 10 weeks), and *retention rates* which measure the completion of follow-up measures (with a range from 22.2% to 100%). Each intervention was also coded for the following characteristics.

**Population.** Eleven interventions were tested among the general population while 10 interventions target at-risk populations such as people with spinal cord injury<sup>25</sup> or with depression.<sup>48</sup>

**Outcomes.** Twelve interventions focused on physical activities such as reducing sedentary time,<sup>21,49</sup> eight interventions focused on dietary behaviors such as limiting caloric intake,<sup>47</sup> and one intervention focused on providing support to people with depression.<sup>48</sup>

**Assessments.** Thirteen interventions only used passive assessments of tailoring variables while eight interventions

used both passive and active assessments of tailoring variables. No interventions used only active assessments.

**Theories.** Nine interventions did not indicate that the messages were designed based on behavioral or intervention theories while 12 interventions' messages were theoretically grounded.

All of the studies and the elements of each intervention that could be coded are presented in Table 1. Other information (i.e. intervention design) regarding each included study is in Appendix 2.

### Meta-analysis results

A three-level random meta-analytic model revealed that the overall effects of JITAIs on health outcomes were significant ( $p < 0.001$ ). The effect size was 0.77 (95% confidence interval (CI) = 0.32, 1.22). Log-likelihood-ratio tests indicated that by comparing the three-level model with the two-level model without within-study variance, the within-study variance was significant ( $p < 0.001$ ). By comparing the three-level model with the two-level model without between-study variance, the between-study variance was significant ( $p < 0.001$ ).

### Publication bias

The rank correlation test confirmed asymmetry of the funnel plot ( $p < 0.001$ ). As shown in Figure 3, the funnel plot was presented with black dots. The trim and fill method indicated that four studies (white dots) would need to be added on the upper right side to symmetrize the funnel plot. These four studies would increase the overall effects of JITAIs from 0.77 to 1. According to the funnel plot, these four studies would have larger

**Table 1.** Study characteristics and coding.

Study	Sample size	Assessments	Theories	Behaviors	Intervention duration (weeks)	Population	Female (%)	Age (M)	White (%)	Retention rates (%)
Bond et al. <sup>21</sup>	30	0	0	1	3	1	83	47.50	67	86
Dorsch et al. <sup>50</sup>	24	1	0	2	8	1	60	-	80	92
Finkelstein et al. <sup>51</sup>	12	0	0	1	4	1	100	52.00	47	90
Forman et al. <sup>52</sup>	62	1	1	2	10	1	84	51.08	70	87
Forman et al. <sup>52</sup>	119	1	1	2	10	1	84	44.02	70	87
Forman et al. <sup>53</sup>	43	1	1	2	8	0	86	50.98	74	76
Goldstein et al. <sup>54</sup>	62	1	1	2	10	1	82	47.27	80	86
Goldstein et al. <sup>54</sup>	59	1	1	2	10	1	86	47.19	68	83
Hiremath et al. <sup>25</sup>	20	1	1	1	4	1	20	39.40	-	80
Pellegrini et al. <sup>49</sup>	8	0	0	1	4	1	88	53.10	23	89
Rabbi et al. <sup>47</sup>	9	0	1	2	3	0	47	28.30	-	94
Rabbi et al. <sup>47</sup>	9	0	1	1	3	0	47	28.30	-	94
Rabbi et al. <sup>55</sup>	16	0	1	1	9	0	56	-	-	100
Rabbi et al. <sup>55</sup>	16	0	1	1	9	0	56	-	-	100
Rabbi et al. <sup>55</sup>	16	0	1	2	9	0	56	-	-	100
Van Dantzig et al. <sup>56</sup>	40	0	0	1	6	0	42	44.50	-	-
Van Dantzig et al. <sup>57</sup>	10	0	0	1	4	0	27	-	-	86

(continued)

Table 1. Continued.

Study	Sample size	Assessments	Theories	Behaviors	Intervention duration (weeks)	Population	Female (%)	Age (M)	White (%)	Retention rates (%)
Van Dantzig et al. <sup>57</sup>	15	0	0	1	4	0	27	-	-	86
Van Dantzig et al. <sup>57</sup>	3	0	0	1	4	0	27	-	-	86
Wahle et al. <sup>48</sup>	12	1	1	3	8	1	64	-	-	22
Wang et al. <sup>58</sup>	7	0	0	1	1	0	71	34.40	-	64

Note: outcomes: physical activities = 1; dietary behavior = 2; providing support to people with depression = 3. Assessments: passive assessments = 0; both passive and active assessments = 1; population: general population = 0; risk populations = 1.

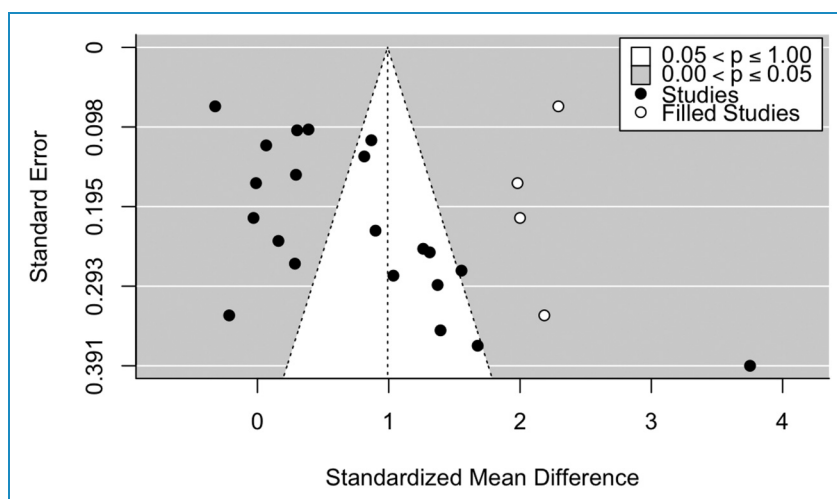


Figure 3. Funnel plot.

standardized mean differences and lower SEs, which suggested that these studies should have larger sample sizes, smaller variances, and larger effect sizes than the existing studies included in this meta-analysis. Realistically, it is easier for larger studies with more significant results to be published. Thus, it is unlikely that publication bias existed. Moreover, Fisher's fail-safe N test showed that 1587 additional studies with non-significant results would be necessary to drag the summary effect below 0.05. Thus, the observed summary effect in our study was demonstrated to be robust to bias.

### Subgroup analysis

Regarding JITAIs that specifically focused on physical activities, a three-level random meta-analytic model

demonstrated that their overall effects on physical activity outcomes were significant ( $p < 0.001$ ). The effect size was 0.98 (95% CI = 0.23, 1.74). Regarding JITAIs that specifically focused on dietary behaviors, a three-level random meta-analytic model demonstrated that their overall effects on dietary behavior outcomes were significant ( $p < 0.001$ ). The effect size was 0.47 (95% CI = 0.12, 0.82).

A series of multivariate moderator analyses revealed that assessments ( $F(1, 19) = 1.22, p = 0.28$ ), theories ( $F(1, 19) = 0.07, p = 0.79$ ), behaviors ( $F(1, 19) = 0.61, p = 0.44$ ), intervention duration ( $F(1, 19) = 0.27, p = 0.61$ ), retention rates ( $F(1, 18) = 0.02, p = 0.89$ ), percentage of female ( $F(1, 19) = 0.47, p = 0.5$ ), mean age ( $F(1, 11) = 0.38, p = 0.55$ ), percentage of White ( $F(1, 7) = 0.03, p = 0.86$ ), population ( $F(1, 19) = 0.08, p = 0.78$ ), and sample size ( $F(1, 19) = 2.02, p = 0.06$ ) did not moderate JITAIs' effects on health outcomes.



The Wilcoxon rank sum test suggested that assessments significantly affected retention rates ( $p=0.03$ ). Specifically, passive assessments of tailoring variables (median retention rates=90%) were more likely than a combination of passive and active assessments of tailoring variables (median retention rates=85%) to relate to higher retention rates. Theories ( $p=0.56$ ) did not affect retention rates. Pearson correlation analysis showed that intervention duration did not significantly correlate with retention rates either ( $r=0.22$ ,  $p=0.83$ ).

### Risk of bias

The RoB 2 tool categorized ten randomized trials into two groups: three as intention-to-treat studies which examined all enrolled participants and four as per-protocol studies which only examined participants who completed all protocol requirements. Among the intention-to-treat studies, 67% were regarded as some concerns and 33% were regarded as high risk of overall bias. A full overview of risk values by domain can be found in Figure 4(a). Among the per-protocol studies, 25% were regarded as low risk of overall bias while 75% were regarded as high risk of overall bias. A full overview of risk values by domain can be found in Figure 4(b). According to the ROBINS-I tool, among the nine non-randomized studies, 29% were regarded as moderate risk of overall bias while 71% were regarded as serious risk of overall bias. Figure 4(c) shows the full overview of risk values by domain.

## Discussion

### Summary effect size

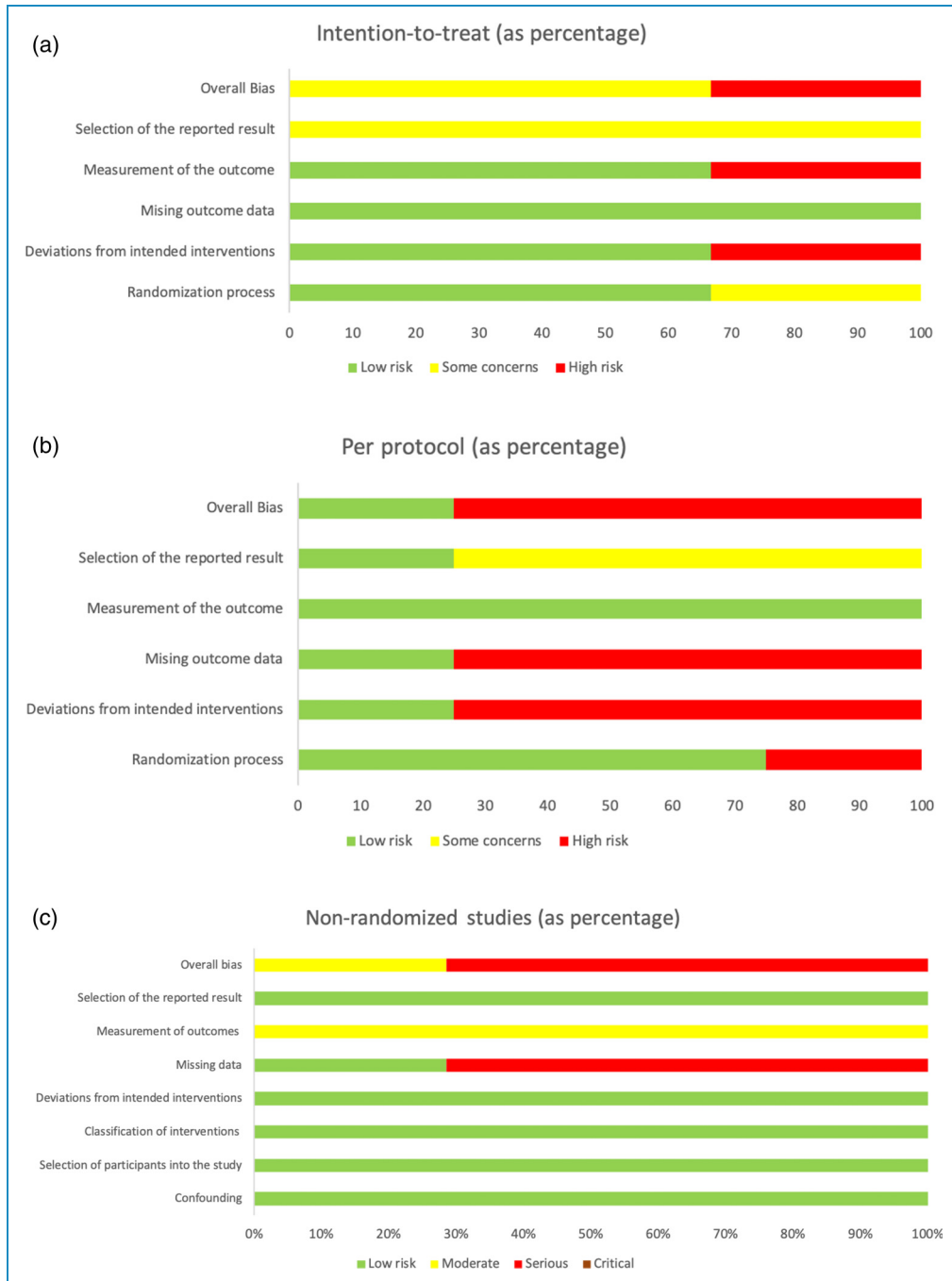
This review demonstrated some support for the efficacy of JITAIs in improving objective health outcomes. According to Cohen,<sup>59</sup> a Hedges'  $g$  between 0.5 and 0.8 indicates a medium effect, and a Hedges'  $g$  that equals or is larger than 0.8 indicates a large effect.<sup>59</sup> Thus, our findings suggest that, on the whole, JITAIs had a moderate impact on health outcomes. JITAIs that were designed to enhance physical activities produced a large impact on physical activity outcomes, whereas JITAIs that aimed at improving dietary habits had a comparatively modest effect on dietary behavior outcomes. The test for heterogeneity indicated that JITAIs were consistently effective across different study settings. Fisher's fail-safe  $N$  test demonstrated that JITAIs' effects were robust to bias. These findings reveal the promise of JITAIs. As explained by the complexity science approach, the reason may be due to their attention to dynamic interactions among internal and contextual factors. It also suggests that these JITAIs may be well able to determine the right timing and frequency of intervention message delivery. However, most of the

current JITAIs have focused on physical activity and dietary behavior. More JITAIs that target other health behaviors, such as cancer prevention behaviors, can be developed and tested in the future.

### Moderating effects

Contrary to our prediction, theoretically grounded JITAIs did not generate better outcomes than JITAIs without theoretical perspectives. Moreover, when theories were used in the JITAIs included, this did not seem to increase retention rates. These findings imply that the theories used in these JITAIs, such as social cognitive theory used in Hiremath et al.,<sup>25</sup> Rabbi et al.,<sup>55</sup> Rabbi et al.,<sup>47</sup> and Wahle et al.,<sup>48</sup> fail to contribute to the persuasiveness of JITAIs. A possibility is that these theories do not address the dynamic nature of underlying behavior change mechanisms and thus provide little guidance on adaptive intervention design for complex behavior change. As mentioned in prior research, the lack of behavioral and intervention theories that are dynamic is a major gap that hinders JITAIs' development. Most existing theories treat underlying behavior change mechanisms as relatively stable and do not allow them to vary over time.<sup>27,60</sup> Therefore, it is necessary to advance theories of health behavior or develop dynamic behavior theories to explain "time- and context-embedded change mechanisms."<sup>61</sup> The complexity science approach and chaos theory can help develop theories that explain and predict "non-linear and quantum influences on human behavior."<sup>62</sup>

Although researchers express concerns regarding intervention adherence and retention,<sup>5</sup> we found that retention rates did not moderate JITAIs' effects. In other words, JITAIs that attract participants over a longer time period may not necessarily lead to better outcomes than JITAIs that only keep participants over a shorter time period. As mentioned in some research, many current JITAI studies primarily focus on usability testing (i.e. engagement and retention) without involving lots of effort in assessing behavior outcomes.<sup>58</sup> Thus, more attention should be devoted to the evaluation of JITAIs' effects on health behaviors. At the same time, it is worthy to note that the majority of the studies that were included show high retention rates. As most of the JITAIs included in this study offered incentives for participation, it is possible that incentives, instead of intervention engagement design, encouraged participants to continue using JITAIs. Currently, in the absence of evidence, it cannot be concluded that the intervention will yield a similar level of effect size when the retention rate is low. Therefore, it remains crucial to make efforts to ensure a high level of retention and adherence during the development of a JITAI. When doing usability testing, future research is expected to rule out the influences of incentives or other factors on intervention adherence.



**Figure 4.** (a) Intention-to-treat studies' publication bias. (b) Per-protocol studies' publication bias. (c) Non-randomized studies' publication bias.

Our results also revealed the relationship between several JITAI's fundamental components and retention rates. Specifically, JITAI's using passive assessments tended to keep participants for a longer time than JITAI's using a combination of passive and active assessments. Besides, intervention duration was not related to retention rates. These findings suggest that repeatedly asking participants to self-report their thoughts and behaviors may lead to

low retention rates. However, interventions that last long may not lead to low retention rates. According to prior work, intervention engagement and retention is also a dynamic system that contains the interplay among affective, cognitive, and behavioral states.<sup>63</sup> Besides the demands of time and effort, individuals' attitudes toward themselves and the intervention; their daily life tasks, moods, and symptoms; and many other factors can all impact their

adherence to the JITAI.<sup>64</sup> Future research can benefit from including other individual-level data (i.e. psychological data) when examining engagement and retention rates.

### Risk of bias

Overall, randomized studies had less risk of overall bias than non-randomized studies, and intention-to-treat studies had less risk of overall bias than per-protocol studies. Among three intention-to-treat studies, one study demonstrated high risk while the risk of bias across six domains among other studies was reasonable. Most per-protocol studies demonstrated high risk due to missing outcome data and deviations from intended interventions mainly because they did not evaluate whether non-adherence could have affected the outcome. For instance, if a JITAI is designed to prevent substance misuse, participants who do not complete the follow-up assessments are highly likely to continue using substances; however, as they do not respond to the assessments, their behavior outcomes are not recorded, and thus, those who are classified as continuous substance users at the post-test are not representative of missing participants.<sup>65</sup> Most non-randomized studies had serious bias due to missing data. Similarly, it was because more than 10% of participants dropped out of the study, and their reasons for non-adherence were not reported. Future JITAI studies are recommended to use high-quality randomized trials and collect data on non-adherence.

### Limitations

It is important to acknowledge several limitations. First, our conclusions, especially related to the moderator analysis, are limited by a few small numbers of subgroups. Several subgroup analyses did not generate statistically significant results, maybe because of insufficient power.

Second, all the JITAI studies included in this analysis used a pre- and post-test design, which may result in an overestimation of the effects. To obtain a more reliable estimate of the effectiveness of JITAI, future meta-analyses could concentrate on randomized control trials that compare the outcomes between arms at follow-up rather than solely examining pre–post changes in the intervention arm. By comparing the effect size from these studies with the one obtained in this study, a more robust estimation of the effectiveness of JITAI could be achieved.

Third, the inclusion criteria for this study included a requirement for objective measurement of health outcomes. It might create a certain bias in the results, as some health outcomes are more easily measured objectively than others. Additionally, there was no evidence to suggest that self-reported behaviors are less valid or reliable. For example, a study revealed that “biomarkers have been

found not to add sufficiently to the accuracy of self-reported measures to warrant being used.”<sup>8</sup> It is worth noting that the lack of active-only assessment-based JITAIs may be attributed to the fact that active assessments are typically employed in cases where outcomes cannot be measured objectively.

Fourth, we applied a step-wise model building approach. Since no factor was significant in the univariate analyses, we did not perform the multivariate analysis. With more JITAI studies published in the future, future meta-analyses can conduct multivariate analysis and examine how design principles may be related to each other.

Finally, the funnel plot indicated that more JITAI studies with large sample sizes were needed. As many current JITAI studies primarily focus on usability testing, it is reasonable to test the designs and features among a relatively small sample.<sup>51</sup> Before generalizing research findings, more larger JITAI studies needed to be tested in real-world environment.

### Conclusion

Despite the limitations, this meta-analysis evaluated and explained the effectiveness of JITAIs through the lens of the complexity science approach and provided practical guidance on key design components that could increase retention rates of JITAIs.

**Acknowledgements:** We would like to thank Hao Guo for her help with data analysis. We deeply appreciate the insightful comments from the editors and the reviewers.



**Contributorship:** Author 1: writing—original draft, conceptualization, methodology, and funding acquisition. Author 2: writing—review and editing.

**Declaration of conflicting interests:** The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Ethical approval:** Not applicable. This study did not involve human subjects.

**Funding:** The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Scripps League Newspaper Fund.

**Guarantor:** Scripps League Newspaper Fund.

**ORCID iDs:** Zhan Xu  <https://orcid.org/0000-0002-6140-6421>  
Eline Smit  <https://orcid.org/0000-0001-8588-6194>

**Supplemental material:** Supplemental material for this article is available online.

## References

1. Bonell C, Fletcher A, Morton M, et al. Realist randomised controlled trials: a new approach to evaluating complex public health interventions. *Soc Sci Med* 2012; 75: 2299–2306.
2. Litaker D, Tomolo A, Liberatore V, et al. Using complexity theory to build interventions that improve health care delivery in primary care. *J Gen Intern Med* 2006; 21: S30–S34.
3. Sheeran P, Goldstein AO, Abraham C, et al. Reducing exposure to ultraviolet radiation from the sun and indoor tanning: a meta-analysis. *Heal Psychol* 2020; 39: 600–616.
4. Young B, Lewis S, Katikireddi SV, et al. Effectiveness of mass media campaigns to reduce alcohol consumption and harm: a systematic review. *Alcohol Alcohol* 2018; 53: 302–316.
5. Nahum-Shani I, Smith SN, Spring BJ, et al. Just-in-time adaptive interventions (JITAI) in mobile health: key components and design principles for ongoing health behavior support. *Ann Behav Med* 2018; 52: 446–462.
6. Smit ES, Linn AJ and Van Weert JCM. Taking online computer-tailoring forward: the potential of tailoring the message frame and delivery mode of online health behavior change interventions. *Eur Heal Psychol* 2015; 17: 25–31.
7. Spruijt-Metz D, Wen CKF, O'Reilly G, et al. Innovations in the use of interactive technology to support weight management. *Curr Obes Rep* 2015; 4: 510–519.
8. Gustafson DH, McTavish FM, Chih MY, et al. A smartphone application to support recovery from alcoholism a randomized clinical trial. *JAMA Psychiatry* 2014; 71: 566–572.
9. Consolvo S, McDonald DW, Toscos T, et al. Activity sensing in the wild: a field trial of UbiFit Garden. In: Conference on human factors in computing systems – proceedings, 2008, pp.1797–1806.
10. 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: e1.
11. Wang L and Miller LC. Just-in-the-moment adaptive interventions (JITAI): a meta-analytical review. *Health Commun* 2020; 35: 1531–1544.
12. Miles A. Complexity in medicine and healthcare: people and systems, theory and practice. *J Eval Clin Pract* 2009; 15: 409–410.
13. Manson SM. Simplifying complexity: a review of complexity theory. *Geoforum* 2001; 32: 405–414.
14. Naughton F, Hopewell S, Lathia N, et al. A context-sensing mobile phone app (Q sense) for smoking cessation: a mixed-methods study. *JMIR mHealth uHealth* 2016; 4: e106.
15. Saleheen N, Ali AA, Hossain SM, et al. Puffmarker: a multi-sensor approach for pinpointing the timing of first lapse in smoking cessation. In: UbiComp 2015 - proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing, 2015, pp.999–1010.
16. Collins LM, Murphy SA and Bierman KL. A conceptual framework for adaptive preventive interventions. *Prev Sci* 2004; 5: 185–196.
17. Patrick K, Intille SS and Zabinski MF. An ecological framework for cancer communication: implications for research. *J Med Internet Res* 2005; 7: e23.
18. Finch L, Janda M, Loescher LJ, et al. Can skin cancer prevention be improved through mobile technology interventions? A systematic review. *Prev Med* 2016; 90: 121–132.
19. Plarre K, Raij A, Hossain SM, et al. Continuous inference of psychological stress from sensory measurements collected in the natural environment. In: Proceedings of the 10th ACM/IEEE international conference on information processing in sensor networks, IPSN'11, Chicago, IL, USA, 2011, pp.97–108.
20. Ali AA, Hossain SM, Hovsepian K, et al. Mpuff: automated detection of cigarette smoking puffs from respiration measurements. In: IPSN'12 - proceedings of the 11th international conference on information processing in sensor networks, Beijing, China, 2012, pp.269–280.
21. Bond DS, Thomas JG, Raynor HA, et al. B-MOBILE - a smartphone-based intervention to reduce sedentary time in overweight/obese individuals: a within-subjects experimental trial. *PLoS One* 2014; 9: e100821.
22. Ben-Zeev D, Brenner CJ, Begale M, et al. Feasibility, acceptability, and preliminary efficacy of a smartphone intervention for schizophrenia. *Schizophr Bull* 2014; 40: 1244–1253.
23. Riley W, Obermayer J and Jean-Mary J. Internet and mobile phone text messaging intervention for college smokers. *J Am Coll Heal* 2008; 57: 245–248.
24. Coughlin LN, Nahum-Shani I, Philyaw-Kotov ML, et al. Developing an adaptive mobile intervention to address risky substance use among adolescents and emerging adults: usability study. *JMIR mHealth uHealth* 2021; 9: e24424.
25. Hiremath SV, Amiri AM, Thapa-Chhetry B, et al. Mobile health-based physical activity intervention for individuals with spinal cord injury in the community: a pilot study. *PLoS One* 2019; 14: e0223762.
26. Webb TL, Joseph J, Yardley L, et al. Using the Internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J Med Internet Res* 2010; 12: e4.
27. Riley WT. Theoretical models to inform technology-based health behavior interventions. In: Marsch L, Lord S and Dallery J (eds) *Behavioral health care and technology*. New York, NY, USA: Oxford, 2016, pp.13–26.
28. Abraham C and Michie S. A taxonomy of behavior change techniques used in interventions. *Heal Psychol* 2008; 27: 379–387.
29. Bandura A. Health promotion by social cognitive means. *Health Educ Behav* 2004; 31: 143–164.
30. Strecher VJ, Seijts GH, Kok GJ, et al. Goal setting as a strategy for health behavior change. *Heal Educ Behav* 1995; 22: 190–200.
31. Fransen ML, Smit EG and Verlegh PWJ. Strategies and motives for resistance to persuasion: an integrative framework. *Front Psychol* 2015; 6: 1201.
32. Tallia AF, Stange KC, McDaniel RR, et al. Understanding organizational designs of primary care practices. *J Healthc Manag* 2003; 48: 45–59.
33. Hardeman W, Houghton J, Lane K, et al. A systematic review of just-in-time adaptive interventions (JITAI) to promote physical activity. *Int J Behav Nutr Phys Act* 2019; 16: 31.
34. Heron KE and Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. *Br J Health Psychol* 2010; 15: 1–39.
35. Smith KE and Juarascio A. From Ecological Momentary Assessment (EMA) to Ecological Momentary Intervention

- (EMI): past and future directions for ambulatory assessment and interventions in eating disorders. *Curr Psychiatry Rep* 2019; 21: 53.
36. Kennedy CM, Powell J, Payne TH, et al. Active assistance technology for health-related behavior change: an interdisciplinary review. *J Med Internet Res* 2012; 14: e80.
  37. Kelly J, Gooding P, Pratt D, et al. Intelligent real-time therapy: harnessing the power of machine learning to optimise the delivery of momentary cognitive-behavioural interventions. *J Ment Heal* 2012; 21: 404–414.
  38. Brookie KL, Mainvil LA, Carr AC, et al. The development and effectiveness of an ecological momentary intervention to increase daily fruit and vegetable consumption in low-consuming young adults. *Appetite* 2017; 108: 32–41.
  39. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Br Med J* 2021; 372: n71.
  40. Hedges LV. Distribution theory for Glass's estimator of effect size and related estimators. *J Educ Stat* 1981; 6: 107–128.
  41. Borenstein M, Hedges LV, Higgins JPT, et al. *Introduction to meta-analysis*. Chichester, West Sussex, UK: John Wiley & Sons Ltd, 2009, pp.21–32.
  42. Konstantopoulos S. Fixed effects and variance components estimation in three-level meta-analysis. *Res Synth Methods* 2011; 2: 61–76.
  43. Sterne JAC and Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *J Clin Epidemiol* 2001; 54: 1046–1055.
  44. Egger M, Smith GD, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. *Br Med J* 1997; 315: 629–634.
  45. Duval S and Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 2000; 56: 455–463.
  46. Rosenthal R. The file drawer problem and tolerance for null results. *Psychol Bull* 1979; 86: 638–641.
  47. Rabbi M, Pfammatter A, Zhang I, et al. Automated personalized feedback for physical activity and dietary behavior change with mobile phones: a randomized controlled trial on adults. *JMIR mHealth uHealth* 2015; 3: e42.
  48. Wahle F, Kowatsch T, Fleisch E, et al. Mobile sensing and support for people with depression: a pilot trial in the wild. *JMIR mHealth uHealth* 2016; 4: e111.
  49. Pellegrini CA, Hoffman SA, Daly ER, et al. Acceptability of smartphone technology to interrupt sedentary time in adults with diabetes. *Transl Behav Med* 2015; 5: 307–314.
  50. Dorsch Michael P, Cornellier Maria L, Poggi Armella D, et al. Effects of a Novel Contextual Just-In-Time Mobile App Intervention (LowSalt4Life) on Sodium Intake in Adults With Hypertension: Pilot Randomized Controlled Trial. *JMIR mHealth and uHealth* 2020; 8(8): e16696. <http://dx.doi.org/10.2196/16696>
  51. Finkelstein J, Bedra M, Li X, et al. Mobile App to Reduce Inactivity in Sedentary Overweight Women. *Studies in health technology and informatics* 2015; 216: 89–92.
  52. Forman Evan M, Goldstein Stephanie P, Crochiere Rebecca J, et al. Randomized controlled trial of OnTrack, a just-in-time adaptive intervention designed to enhance weight loss. *Translational Behavioral Medicine* 2019; 9(6): 989–1001. <http://dx.doi.org/10.1093/tbm/ibz137>
  53. Forman Evan M, Goldstein Stephanie P, Zhang Fengqing, et al. OnTrack: development and feasibility of a smartphone app designed to predict and prevent dietary lapses. *Translational Behavioral Medicine* 2019; 9(2): 236–245. <http://dx.doi.org/10.1093/tbm/iby016>
  54. Goldstein Stephanie P, Thomas J Graham, Foster Gary D, et al. Refining an algorithm-powered just-in-time adaptive weight control intervention: A randomized controlled trial evaluating model performance and behavioral outcomes. *Health Informatics Journal* 2020; 26(4): 2315–2331. <http://dx.doi.org/10.1177/1460458220902330>
  55. Rabbi M, Aung MH, Zhang M, et al. Mybehavior: automatic personalized health feedback from user behaviors and preferences using smartphones. In: UbiComp 2015 - proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing, New York, NY, USA, 2015, pp.707–718.
  56. van Dantzig Saskia, Geleijnse Gijs and van Halteren Aart Tijmen. Toward a persuasive mobile application to reduce sedentary behavior. *Pers Ubiquitous Comput* 2013; 17(6): 1237–1246. <http://dx.doi.org/10.1007/s00779-012-0588-0>
  57. Van Dantzig Saskia, Bulut Murtaza and Krans Martijn et al. Enhancing physical activity through context-aware coaching. In: PervasiveHealth '18: Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare, New York, NY, USA, 21–24 May 2018, pp. 187–190.
  58. Wang S, Sporrel K, Van Hoof H, et al. Reinforcement learning to send reminders at right moments in smartphone exercise application: a feasibility study. *Public Health* 2021; 18: 6059.
  59. Cohen J. *Statistical power analysis for the behavioral sciences*. New York, NY, USA: Lawrence Erlbaum Associates, 1988, pp.1–17.
  60. Riley WT, Rivera DE, Atienza AA, et al. Health behavior models in the age of mobile interventions: are our theories up to the task? *Transl Behav Med* 2011; 1: 53–71.
  61. Heino MTJ, Knittle K, Noone C, et al. Studying behaviour change mechanisms under complexity. *Behav Sci (Basel)* 2021; 11: 77.
  62. Resnicow K and Page SE. Embracing chaos and complexity: a quantum change for public health. *Am J Public Health* 2008; 98: 1382–1389.
  63. King G, Currie M and Petersen P. Child and parent engagement in the mental health intervention process: a motivational framework. *Child Adolesc Ment Health* 2014; 19: 2–8.
  64. Heckman BW, Mathew AR and Carpenter MJ. Treatment burden and treatment fatigue as barriers to health. *Curr Opin Psychol* 2015; 5: 31–36.
  65. Businelle MS, Ma P, Kendzor DE, et al. An ecological momentary intervention for smoking cessation: evaluation of feasibility and effectiveness. *J Med Internet Res* 2016; 18: e321.
-