Behaviour Research and Therapy 193 (2025) 104815

Contents lists available at ScienceDirect

Behaviour Research and Therapy

journal homepage: www.elsevier.com/locate/brat

Early changes in passively sensed homestay predict depression symptom improvement during digital behavioral activation

Carter J. Funkhouser [a,b,*], Lauren S. Weiner [c], Ryann N. Crowley [c,d], Jon F. Davis [e], Frank H. Koegler [f], Nicholas B. Allen [c,d], Randy P. Auerbach [a,b]

[a] Department of Psychiatry, Columbia University, New York, NY, USA
[b] Division of Child and Adolescent Psychiatry, New York State Psychiatric Institute, New York, NY, USA
[c] Ksana Health, Inc., Eugene, OR, USA
[d] Center for Digital Mental Health, University of Oregon, Eugene, OR, USA
[e] Novo Nordisk Research Center, Lexington, MA, USA
[f] Novo Nordisk Research Center, Seattle, WA, USA

ARTICLE INFO

Keywords:
Digital phenotyping
Early response
mHealth
Mobile sensing
Behavior change

ABSTRACT

Digital behavioral activation (BA) is scalable, accessible, and efficacious for depression. However, some individuals do not improve during digital BA, and identifying non-responders early is critical for facilitating adaptive intervention approaches (e.g., stepped care). To explore whether passive sensing data might serve as early predictors of symptom change, we tested whether early changes in passively sensed behavioral targets of BA predicted depression symptom changes during app-based BA. Young adults (N = 47) with elevated depressive symptoms completed a 12-week trial of an app-based BA intervention, Vira. The Vira app provided BA psychoeducation, assessed self-reported daily mood, and used smartphone sensors to passively assess time spent at home (i.e., homestay), walking, stationary time, time in bed, bedtime, and waketime each day. We quantified early behavioral changes by fitting a multilevel growth model for each behavior over the first 2 weeks of the intervention. Models included a random slope reflecting each participant's average day-to-day change in that behavior. We extracted these slope estimates and tested whether they predicted depressive symptom (PHQ-8) change from pre-to post-intervention. We hypothesized that individuals with greater early changes in intervention-targeted behaviors would experience greater reductions in depressive symptoms. As hypothesized, individuals with greater early decreases in passively sensed homestay (i.e., reduced behavioral withdrawal) experienced a greater reduction in depressive symptoms by the end of treatment ($b = 0.94$, $p = .025$). Early changes in other behaviors did not significantly predict depressive symptom change ($ps > .158$). Passively monitoring early changes in homestay during app-based BA may support the early identification of individuals at risk of symptom persistence, thus providing earlier opportunities to adjust treatment.

1. Introduction

Depression is a leading cause of global disability (Friedrich, 2017) and is most prevalent in young people (Avenevoli et al., 2015). Alarmingly, the prevalence of depression among young people has more than doubled in the past 15 years (Daly, 2022). Novel intervention approaches are needed to meet the growing demand for services. Digital interventions are a particularly promising approach due to their acceptability, scalability, accessibility, efficacy, and low cost (Lehtimaki et al., 2021). However, some individuals do not improve while using a digital intervention, and identifying these individuals early is critical for facilitating adaptive intervention strategies (e.g., stepped care to higher intensity interventions), which reduces the duration of distress and improves resource allocation (Meeuwissen et al., 2019; Rivero-Santana et al., 2021).

Information collected early in treatment may predict treatment response and can inform treatment decisions, thereby enhancing clinical outcomes for those at risk of not improving without treatment modifications (Beard & Delgadillo, 2019; Lutz et al., 2022). Specifically, early changes in symptom severity (i.e., early response) and skill utilization have been identified as early predictors of symptom reduction during digital depression interventions (Domhardt et al., 2021; Jackson et al.,

* Corresponding author. Department of Psychiatry, Columbia University, New York, NY, USA.
E-mail address: carter.funkhouser@nyspi.columbia.edu (C.J. Funkhouser).

https://doi.org/10.1016/j.brat.2025.104815
Received 29 August 2024; Received in revised form 21 February 2025; Accepted 30 June 2025
Available online 1 July 2025

2023; Lutz et al., 2017). For example, many digital depression interventions are based on behavioral activation (BA), an empirically supported treatment that reduces depressive symptoms by promoting behavioral activation and engagement with rewarding and meaningful activities (e.g., physical activity, spending time with loved ones) (Kanter et al., 2012). Meta-analyses indicate that both face-to-face and digital BA are effective for depression, and qualitative studies suggest that young people prefer to use behavioral strategies to manage their depression (Alber et al., 2023; Stein et al., 2021). Increases in rewarding and meaningful behaviors precede improvements in depressive symptoms during BA (Webb et al., 2019) and mediate the effects of digital BA on depressive symptoms (Bisby et al., 2022; Fu et al., 2021). Monitoring early changes in these behaviors may therefore improve treatment decisions by identifying individuals who are not utilizing BA skills and are less likely to improve.

Despite the clinical benefits of monitoring early skill utilization during treatment, this early predictor is rarely monitored in real-world settings due to concerns about the burdensomeness and validity of the retrospective self-report measures used to assess them (Hermens et al., 2014; Lewis et al., 2019). Mobile sensing directly addresses this barrier by leveraging built-in smartphone sensors to assess clinically relevant behaviors continuously, objectively, and with minimal burden (Mohr et al., 2017). Smartphones' GPS and accelerometer sensors can passively assess homestay (i.e., the percentage of time spent at home versus out of home) and other behavioral indicators of physical activity (e.g., time spent walking, time spent stationary). These behaviors relate to depressive symptom severity both cross-sectionally and longitudinally (de Angel et al., 2022; Rohani et al., 2018), and are measured and explicitly targeted by some app-based BA interventions. For example, the Vira app provides BA psychoeducation, passively monitors behaviors associated with depression, identifies passively sensed behaviors associated with users' daily mood, and teaches behavioral skills to help users change behaviors associated with worse daily mood (Tse et al., 2025). In a recent report on the current trial comparing engagement and efficacy across self-guided and coach-guided versions of the Vira intervention, the coach-guided version significantly reduced the primary outcome measures of depressive and anxious symptoms whereas the self-guided version did not (although both versions improved some secondary outcomes such as symptoms of stress; Weiner et al., 2024). However, a proportion of the participants in both conditions did not experience any reduction in depressive symptoms, underscoring the importance of identifying early markers that could be used to guide treatment adjustments.

To identify these individuals early, this exploratory study conducted secondary analyses of data from this trial (Weiner et al., 2024) to investigate whether early changes in passively sensed behavioral intervention targets predicted depressive symptom change during an app-based BA intervention, Vira. We hypothesized that individuals with greater early changes in intervention-targeted behaviors (i.e., decreased time spent at home, increased time spent walking, and decreased time spent stationary) would experience greater reductions in depressive symptoms. Our primary analyses focused on these physical activity behaviors, as the other behaviors passively assessed in this study pertain to sleep, which empirical and theoretical work suggests is less relevant to symptom change during digital BA (Domhardt et al., 2021; Kanter et al., 2012). However, we also examined these sleep features (i.e., sleep duration, bedtime, waketime) as a secondary aim. As an exploratory aim, we also tested whether associations between early behavioral changes and depressive symptom change differed depending on the Vira intervention version (self-guided versus coach-guided).

2. Materials and methods

2.1. Study design

This study conducted secondary analyses of a fully remote, two-arm RCT comparing self-guided and coach-guided versions of the Vira intervention (ClinicalTrials.gov NCT05638516) (Weiner et al., 2024). After completing a baseline assessment, participants were randomized to either a self-guided or coach-guided version of the Vira intervention. The intervention period was 12 weeks. Depressive symptoms were assessed at baseline and a post-intervention follow-up at 12 weeks. Study procedures were approved by the Oregon Research Institute Institutional Review Board and are described in detail elsewhere (Weiner et al., 2024). All participants provided informed consent.

2.2. Participants

Young adults ages 18-25-years-old with elevated depressive symptoms and an obesity risk factor were recruited between November 2022 and January 2023 by a market research firm, KJT Group Inc. (Rochester, NY, USA). The study was advertised to members of KJT's existing databases and consumer research panels who opted in to participate in research. Eligibility criteria included: (1) ages 18–25 years old, (2) resident of the United States, (3) English fluency and literacy, (4) access to an Android or iOS smartphone, (5) elevated depressive symptoms (Patient Health Questionnaire [PHQ-8]$\geq$10), and (6) overweight (BMI$\geq$25 kg/m^2) or a parental history of overweight or obesity. Recruitment efforts aimed to recruit a sample with a demographic composition resembling that of the United States. Given this study's focus on passive sensing data collected early in the intervention, we excluded 10 participants who did not download the Vira app and 11 participants who did not have any passive sensing data during the first two weeks. Participants were included if they had any passive sensing data in the first 2 weeks, and the final sample included 52 young adults. However, 10 participants only had data available for certain passive sensing features during the first two weeks, and each passive sensing feature was available for 5 of these 10 participants during the first two weeks. Thus, each analysis included 47 participants. Sample characteristics are presented in Table 1. Included and excluded participants did not differ on baseline sociodemographic or clinical characteristics (Table S1).

2.3. Intervention conditions

For participants randomized to the self-guided condition, the Vira app passively collected data from smartphone sensors to assess behaviors relevant to depression and BA (Funkhouser et al., 2024; Weiner et al., 2024). The app passively assessed mobility behaviors, including GPS-derived homestay (i.e., the percentage of the day spent at home), time spent walking, and time spent stationary, which were the primary focus in this study. Sleep features (i.e., time spent sleeping, bedtime, and waketime) were also assessed, and were examined in secondary analyses.[1] Each day, the Vira app also prompted users to rate their enjoyment during the previous day on a scale of 0 ("not at all") to 4 ("super enjoyable"). Daily enjoyment ratings over the past week and passive sensing data for the prior day were presented in the app. Beginning on day 10, the app offered up to two personalized insights per week based on associations between daily enjoyment ratings and passively sensed behaviors (e.g., "you tend to enjoy the day more when you wake up earlier"). The app encouraged users to explore more detailed information about the insight and offered in-app and external resources describing strategies for modifying the behavior (e.g., "set a gentle alarm to wake up 15 min earlier"; Fig. 1). Thus, participants in the self-guided condition received feedback on daily behaviors, daily mood

[1] The Vira app also assessed time spent exercising and time spent in transit, but we did not analyze these behaviors because they had minimal variation over the first two weeks, leading to singular fits. Language patterns in key inputs were also assessed, but were not analyzed because numerous participants opted out of key input data collection.

Table 1
Baseline demographic and clinical characteristics.

Characteristic	Overall (N = 52)	Self-Guided (n = 26)	Coach-Guided (n = 26)
Age	22.04 (2.15)	22.92 (1.67)	21.15 (2.24)
Gender (%)			
Cisgender Female	35 (67.3%)	18 (69.2%)	17 (65.4%)
Cisgender Male	13 (25.0%)	6 (23.1%)	7 (26.9%)
Transgender Female	3 (5.8%)	1 (3.8%)	2 (7.7%)
Non-Conforming	1 (1.9%)	1 (3.8%)	0 (0.0%)
Race (%)			
White	30 (57.7%)	15 (57.7%)	15 (57.7%)
African American	14 (26.9%)	7 (26.9%)	7 (26.9%)
Asian American	2 (3.8%)	1 (3.8%)	1 (3.8%)
Biracial	4 (7.7%)	1 (3.8%)	3 (11.5%)
Other	2 (3.8%)	2 (7.7%)	0 (0.0%)
Ethnicity (% Hispanic)	7 (13.5%)	3 (11.5%)	4 (15.4%)
Education (%)			
Some high school	2 (3.8%)	1 (3.8%)	1 (3.8%)
Graduated high school	18 (34.6%)	8 (30.8%)	10 (38.5%)
Some college	19 (36.5%)	10 (38.5%)	9 (34.6%)
Graduated 2-year college	4 (7.7%)	2 (7.7%)	2 (7.7%)
Graduated 4-year college	9 (17.3%)	5 (19.2%)	4 (15.4%)
Employment (%)			
Employed full time	17 (32.7%)	11 (42.3%)	6 (23.1%)
Employed part time	6 (11.5%)	2 (7.7%)	4 (15.4%)
Homemaker	2 (3.8%)	2 (7.7%)	0 (0.0%)
Student	18 (34.6%)	6 (23.1%)	12 (46.2%)
Unemployed	8 (15.4%)	4 (15.4%)	4 (15.4%)
Other	1 (1.9%)	1 (3.8%)	0 (0.0%)
Body Mass Index, kg/m^2	29.42 (8.18)	29.14 (7.58)	29.71 (8.88)
Overweight Status (%)			
Self-overweight	7 (13.5%)	3 (11.5%)	4 (15.4%)
Parental history of overweight	19 (36.5%)	10 (38.5%)	9 (34.6%)
Self & parental history of overweight	26 (50.0%)	13 (50.0%)	13 (50.0%)
Smartphone operating system (%)			
iOS	30 (57.7%)	16 (61.5%)	14 (53.8%)
Android	22 (42.3%)	10 (38.5%)	12 (46.2%)
Depressive symptoms (PHQ-8)	14.90 (4.65)	14.81 (4.45)	15.00 (4.93)

self-assessment, insights into connections between their behavior and mood, and in-app and external psychoeducational resources. Participants in the coach-guided condition used the Vira app as described above, and additionally received text-messaging support and in-app prompts and reminders related to their behavior change goals from a trained health coach (see Weiner et al., 2024). Participants could not be blinded, as they were aware that they were either receiving coaching or not.

2.4. Measures

2.4.1. Depressive symptoms

Depressive symptoms were assessed at baseline and 12-week follow-up using the Patient Health Questionnaire-8 (Kroenke et al., 2009). The PHQ-8 has strong reliability and construct validity (de la Torre et al., 2023), and had adequate internal consistency at baseline (α = .80) and 12-week follow-up (α = .84).

2.5. App-derived passive features

The Vira app captured GPS measurements of latitude and longitude every 15 min for Android phones and each time an iPhone traveled >100 m. Homestay was calculated as the percentage of the day spent at home, with home defined as the location at which participants spent the most time between 2 a.m. and 6 a.m. Time spent walking, stationary, and sleeping and bedtime and waketime were assessed by Vira using Android's and iOS's built-in motion capture capabilities, which leverage accelerometer and gyroscope data to recognize activity states. Time spent asleep was operationalized as the longest stationary period in each 24-h period (6 p.m.–6 p.m.). The start and end of the sleep period represented bedtime and waketime, respectively. Data were encrypted and uploaded to a secure cloud server each night, and passive sensing data were automatically processed to generate daily and weekly aggregates for each feature. These passive features can accurately measure behaviors (Straczkiewicz et al., 2021) and have demonstrated validity via associations with daily mood (Bitran et al., 2024) and depression symptoms and risk (Auerbach et al., 2022, 2025; de Angel et al., 2022; Rohani et al., 2018). Availability of passive data features in the first 2 weeks was quite high for most participants, with the average participant providing at least 9 days of data for each feature (Fig. S1).

2.6. Data analyses

2.6.1. Early changes in passively assessed behaviors

To quantify early changes in each behavior, we fit a multilevel growth model for each behavior over the first 2 weeks of the intervention. A 2-week period was selected because users received BA psychoeducation on day 0 and began receiving insights on day 10. Each growth model included a random participant-level intercept and a random slope for day number. Growth models were fit with restricted maximum likelihood to minimize bias in the variance component estimates (McNeish & Stapleton, 2016). Best linear unbiased predictions (Robinson, 1991) of each participant's random slope and intercept were extracted from each model. Extracted slopes reflect each participant's average day-to-day change in the behavior over the first two weeks, and intercepts reflect each participant's estimated 'level' of the behavior on

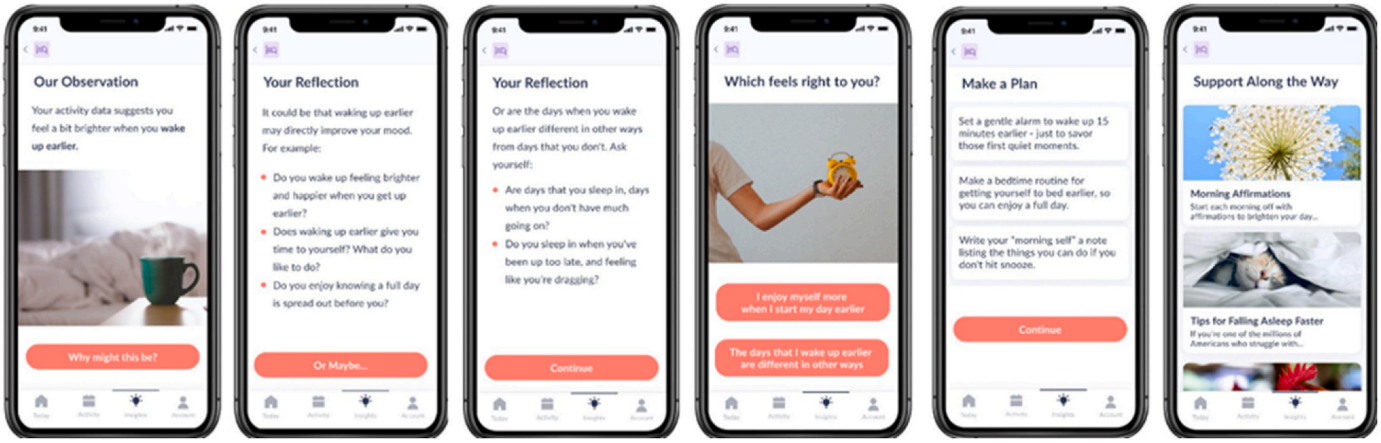

Fig. 1. An example of the content presented in the Vira insights feature. Figure originally published in JMIR Mental Health by Weiner et al. (2024).

the first day of Vira data collection (day 0). Fig. 2 visualizes this approach using homestay as an example.

2.6.2. Associations between early behavioral changes and depressive symptom reduction

Multilevel modeling was used to investigate whether early behavioral changes predicted change in depressive symptoms from baseline to the 12-week follow-up (i.e., post-intervention). Each behavior was examined in a separate model, and each model tested whether early change in the behavior moderated the effect of timepoint (baseline versus 12-week follow-up) on depressive symptom severity. All models included a participant-level random intercept and covaried for intervention version. Models were fit with restricted maximum likelihood assuming missingness at random (Bloom et al., 2024), and *p*-values were estimated using Kenward-Roger's method. To examine whether associations between early behavioral changes and depressive symptom change differed between the self-guided versus coach-guided intervention versions, exploratory analyses tested whether intervention version moderated the interactions between timepoint and early behavioral change in predicting depression symptom severity. Significant interactions were probed using the Johnson-Neyman method (Johnson & Neyman, 1936). This technique calculates regions of significance and confidence intervals to delineate the exact values of a moderator (e.g., early behavioral change) at which the effect of a predictor (e.g., time point, representing change in depressive symptoms over the 12-week intervention) is and is not statistically significant, thus revealing the specific pattern of early behavioral change significantly associated with symptom improvement. Analyses were conducted in R using the lme4 (Bates et al., 2015), lmerTest (Kuznetsova et al., 2017), interactions (Long, 2024), and simr (Green & Macleod, 2016) packages.

2.6.3. Sensitivity analyses

A priori power analyses were not performed, but post-hoc sensitivity analyses were conducted using Monte Carlo simulations. We substituted the effect of the focal interaction (e.g., Timepoint x Early behavioral change) on depressive symptoms with a range of effect sizes separated by increments of 0.2. For each effect size, we simulated the model 1000 times and extracted the proportion of iterations for which the focal effect was statistically significant (i.e., power). Additionally, a series of sensitivity analyses were conducted to test the robustness and specificity of any significant associations between early behavioral change and depression symptom change. We tested whether the effect was independent of: (1) participants' initial levels of the behavior (i.e., estimated intercepts), as intercepts and slopes are often confounded; (2) early changes in other passively sensed behaviors, as early behavioral changes may be correlated; (3) month of baseline assessment, as the examined behaviors vary seasonally (Chan et al., 2006); (4) changes in the behavior over the entire 12-week intervention period, as the effect may not be specific to *early* change; and (5) early changes in self-reported enjoyment, as passive data and self-report data may capture some of the same information.

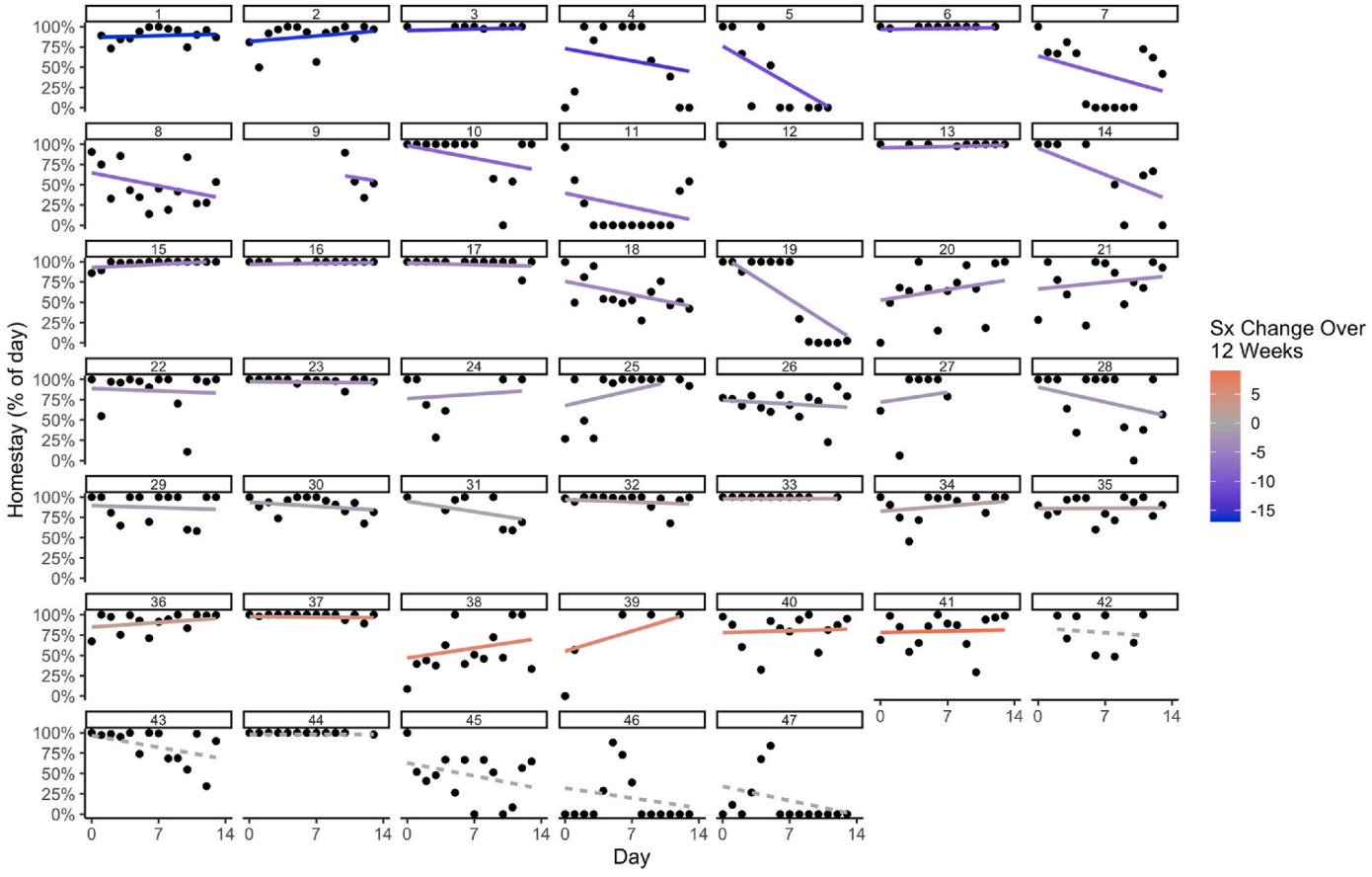

Fig. 2. Individual participants' slopes of homestay over the first two weeks of the intervention. Each facet represents a participant. Black dots represent the participant's observed daily homestay. Lines represent the participant's estimated slope of homestay. Lines are colored by participants' change in depressive symptoms from baseline to 12-week follow-up. Blue lines indicate symptom improvement and red lines indicate symptom worsening. Solid gray lines indicate no change in symptom severity. Participants with dashed gray lines did not complete the depressive symptom assessment at follow-up. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

3. Results

Descriptive statistics characterizing day-to-day behavioral changes over the first two weeks of the intervention are presented in Table S2. On average, behavioral features did not significantly change over the first two weeks (*ps* > .100). However, there was considerable between-person variation in the amount of early behavioral change (Fig. S2).

3.1. Associations between early behavioral changes and depressive symptom changes

Early changes in GPS-assessed homestay over the first 2 weeks of the intervention predicted changes in depressive symptoms (b = 0.94, p = .025), as individuals with greater early decreases in homestay experienced a greater reduction in depressive symptoms from pre-to post-intervention. Fig. 3A visualizes early change in homestay in relation to depression symptom change. Johnson-Neyman analyses probing this effect indicated that depressive symptoms significantly decreased only for individuals whose homestay decreased, stayed the same, or increased by <1% day-over-day over the first two weeks (Fig. 3B). The predictive performance of an exploratory model using early change in homestay to predict reliable change in depressive symptoms (operationalized using the reliable change index; Jacobson & Truax, 1991) was as follows: area under the curve = 0.70, sensitivity = 0.54, specificity = 0.86. Early changes in time spent walking (b = −2.37, p = .520) and time spent stationary (b = 5.03, p = .545) did not significantly predict change in depressive symptoms during the intervention. Early changes in sleep characteristics were examined as a secondary aim and did not predict depressive symptom change (*ps* > .158; Fig. S3). Exploratory analyses revealed that associations between early behavioral changes and depression symptom improvement did not significantly differ between the self-guided versus coach-guided versions of Vira (*ps* > .244). However, Monte Carlo simulations indicated that these analyses were underpowered and only likely to detect large differences between intervention versions (Fig. S5).

3.2. Sensitivity analyses

Sensitivity analyses examining the specificity of the association between early change in homestay and depressive symptom change (Table S3) found that it remained significant when additionally controlling for initial levels (i.e., intercept) of early homestay (b = 1.00, p = .023), early changes in all other passively sensed behaviors (b = 0.89, p = .039; Table S4), month of baseline assessment (b = 0.86, p = .047), or

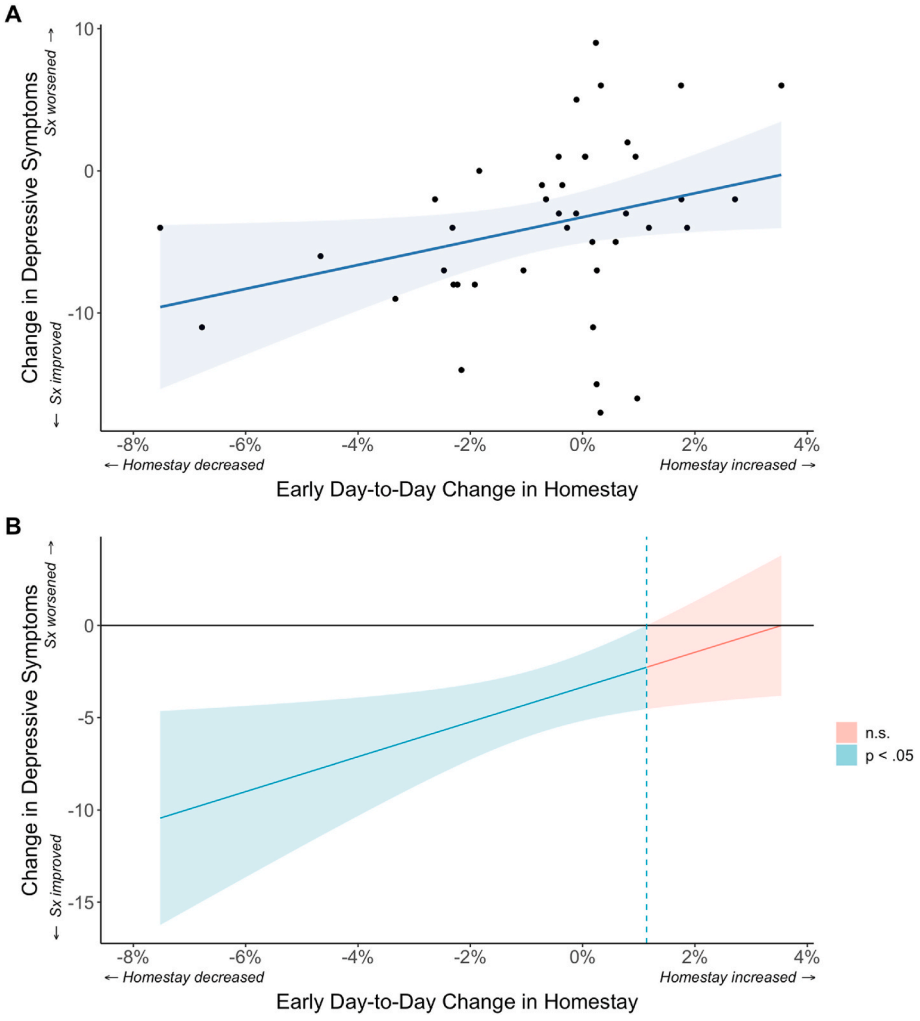

Fig. 3. (A) The relationship between early change in homestay and change in depressive symptoms from pre-to post-intervention. (B) Johnson-Neyman plot of the effect of early slopes of homestay on change in depressive symptoms. On the y-axis, negative values reflect an improvement in depressive symptoms and positive values reflect a worsening of symptoms. On the x-axis, negative values reflect decreases in homestay and positive values reflect increases in homestay. The region in which there was a statistically significant reduction in depressive symptoms is shaded in light blue and bounded by a dashed vertical line. The region in which depressive symptoms did not significantly decrease is shaded in light red. In both panels, shaded ribbons represent 95% confidence intervals. Sx = Symptoms. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

changes in homestay over the entire 12-week intervention period (b = 0.91, $p = .032$). When also covarying for early changes in self-reported enjoyment, the effect size was similar in magnitude, but no longer statistically significant at the conventional level (b = 0.84, $p = .061$). We also explored whether early changes in homestay over longer time periods (i.e., 3 weeks, 4 weeks, or the entire 12-week intervention period) predicted depressive symptom improvement. The effect size was larger—but not statistically significant—when using 3 weeks (b = 1.40, $p = .059$), 4 weeks (b = 1.58, $p = .070$), or 12 weeks (b = 1.88, $p = .278$) of data instead of 2 weeks (Fig. S4). Further inspection revealed that using more weeks of data to estimate early homestay change resulted in less between-person variability in homestay change (e.g., $SD_{2\text{weeks}} = 2.52$ versus $SD_{12\text{weeks}} = 0.61$). Finally, we considered that some participants' early changes in homestay may have been driven by potentially unusual homestay patterns (e.g., a vacation) during the first two weeks. To test this possibility, we re-estimated the analysis excluding participants with 2+ consecutive days with 0% homestay during the first two weeks ($n = 7$). When excluding these individuals, the effect was numerically similar, but no longer statistically significant (b = 0.98, $p = .130$).

4. Discussion

Innovative intervention approaches that are scalable, accessible, and cost-effective are urgently needed to address the rising demand for depression treatment in young people. Delivering interventions digitally is a promising way to address these needs. However, many individuals do not improve with digital interventions, and identifying early predictors of symptom improvement is critical for facilitating timely treatment modifications. Early changes in self-reported behavioral targets of app-based BA are early predictors of symptom improvement, but are rarely assessed in real-world settings due to concerns about the burdensomeness and validity of retrospective self-report methods (Hermens et al., 2014; Lewis et al., 2019). An alternative way to assess these behavioral targets of BA is through passive sensing, which–unlike retrospective self-report methods–assesses these behaviors objectively, continuously, and with minimal burden. Importantly, young adults with depression are interested in app-based interventions that passively monitor behaviors related to their mental health (Beltzer et al., 2023), and some app-based interventions are already using passive sensing to assess these behaviors throughout the intervention.

This study tested whether passive sensing data collected during one of these app-based interventions can be leveraged for early identification of individuals at risk of not improving. As hypothesized, individuals with greater early decreases in passively assessed homestay experienced greater reductions in depressive symptoms during the intervention, providing preliminary evidence that passively monitoring early changes in homestay may help identify individuals at risk of not improving. The association between early change in homestay and later symptom reduction was independent of early changes in all other passively sensed behaviors (none of which predicted symptom change) and various other potential confounds such as initial levels of homestay and longer-term changes in homestay. This pattern of results suggests that decreases in depressive symptoms may be specifically predicted by early reductions in homestay. Consistent with this finding, homestay is one of the strongest and most consistent passively sensed predictors of depression (Shin & Bae, 2023), and increases in homestay precede increases in depressive symptoms (Stamatis et al., 2024; Zhang et al., 2022). Decreased homestay is also particularly relevant to symptom change during digital BA as it is a behaviorally anchored marker of reduced behavioral withdrawal, a key target and putative mechanism of BA (Kanter et al., 2012). If these findings are replicated, early changes in passively assessed homestay may be a candidate for inclusion in an automated early identification tool. Such a tool could use early changes in homestay and other early predictors as inputs to predict app-based BA users' expected symptom reduction or likelihood of response. Treatment modifications (e.g., 'stepping up' to a higher intensity intervention) could then be implemented for individuals unlikely to respond.

This study is intended to serve as a proof of concept exploring the potential for passive sensing data to serve as early predictors of symptom change, and thus, results should be replicated. These findings also highlight directions for future research examining digital mental health interventions and early identification of non-improvers. For example, the sample was fairly representative of young adults who would use a digital depression intervention (Rickwood et al., 2016), but the generalizability of these results to other age groups and those without an obesity risk factor is unclear. Digital BA has previously been shown to be efficacious for other mental health problems in addition to depression (e.g., anxiety symptoms; Alber et al., 2023), and future research could similarly examine whether findings generalize to other symptoms. More broadly, it is important to build upon research investigating whether and how individual characteristics (e.g., overweight status, low mood) and intervention characteristics (e.g., human support) impact digital intervention effects (Moshe et al., 2021). Other potential future directions include: (1) identifying other early predictors of symptom improvement and treatment response, (2) developing and evaluating a multivariable algorithm that uses multiple early predictors to improve the early identification of individuals unlikely to improve, and (3) implementing this predictive algorithm in a sequential, multiple assignment, randomized trial (SMART; Collins et al., 2007) to empirically test whether 'stepping up' these individuals to a higher intensity intervention (e.g., teletherapy) leads to improved outcomes.

This study has several notable limitations. First, the sample size was relatively small and analyses testing whether associations differed between the self-guided and coach-guided versions were likely underpowered. Second, it is unclear whether early decreases in homestay were predicting depressive symptom reduction or response to app-based BA, as there was not a no-treatment control group. Relatedly, it is unclear if the results are specific to app-based BA or generalizable to other app-based interventions. However, the association between early change in homestay and later depressive symptom reduction did not differ between self-guided and coach-guided versions of the intervention, suggesting that monitoring early changes in homestay may help identify individuals at risk of not improving regardless of whether they are using a self-guided or coach-guided version of app-based BA. Third, depressive symptoms were assessed at baseline and a post-intervention follow-up 12 weeks later, but not during the intervention. Thus, it is unclear when changes in depressive symptoms occurred. Last, we fit growth models to quantify early behavioral change and used slopes from these growth models to predict depressive symptom change. Quantifying early behavioral change using growth models is more accurate and unbiased than alternative approaches (e.g., difference scores; Snijders & Bosker, 2012), but this two-step approach introduced the potential for generated regressor bias.

5. Conclusions

Passive sensing data are already being collected by some app-based BA interventions, and early changes in these passive data may prospectively predict depressive symptom reduction. Specifically, passively monitoring early changes in homestay may facilitate the early identification of non-responders and accordingly, provide novel opportunities to adjust treatment.

CRediT authorship contribution statement

Carter J. Funkhouser: Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization. **Lauren S. Weiner:** Investigation, Data curation. **Ryann N. Crowley:** Writing – review & editing, Data curation. **Jon F. Davis:** Writing – review & editing. **Frank H. Koegler:** Writing – review & editing. **Nicholas B. Allen:** Funding acquisition, Investigation, Writing – review & editing. **Randy P. Auerbach:** Investigation, Writing – review & editing, Funding acquisition.

Data availability statement

Data are available from NBA upon reasonable request.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Carter Funkhouser reports financial support was provided by National Institute of Mental Health and the Klingenstein Third Generation Foundation. Nicholas Allen reports financial support was provided by National Institute of Mental Health. Nicholas Allen has received research funding from Google Health and consulting fees from Snap Inc. Randy Auerbach reports financial support was provided by National Institute of Mental Health. Randy Auerbach reports financial support was provided by Morgan Stanley Foundation. Randy Auerbach reports a relationship with Ksana Health that includes: consulting or advisory. In the past 3 years, Randy Auerbach has received consulting fees and equity from Get Sonar Inc. Randy Auerbach also has received consulting fees from RPA Health Consulting, Inc. and Covington & Burling LLP, which is representing a social media company in litigation. Lauren Weiner reports a relationship with Ksana Health that includes: employment and equity or stocks. Ryann Crowley reports a relationship with Ksana Health that includes: employment and equity or stocks. Nicholas Allen reports a relationship with Ksana Health that includes: employment and equity or stocks. Jon F. Davis and Frank H. Koegler were employed by Novo Nordisk at the time of the trial. Novo Nordisk provided funding for the randomized controlled trial as part of the Prevention Accelerator, but did not support the secondary analyses reported in this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.brat.2025.104815.

Data availability

Data will be made available on request.

References

Alber, C. S., Krämer, L. V., Rosar, S. M., & Mueller-Weinitschke, C. (2023). Internet-based behavioral activation for depression: Systematic review and meta-analysis. *Journal of Medical Internet Research, 25*(1), Article e41643. https://doi.org/10.2196/41643

Auerbach, R. P., Bloom, P. A., Pagliaccio, D., Lan, R., Galfalvy, H., Bitran, A., Durham, K., Crowley, R., Joyce, K., Blanchard, A., Chernick, L. S., Dayan, P. S., Greenblatt, J., Kahn, L. E., Porta, G., Tse, T. C., Cohn, J. F., Morency, L.-P., Brent, D. A., & Allen, N. B. (2025). Using smartphone GPS data to detect the risk of adolescent suicidal thoughts and behaviors. *JAMA Network Open, 8*(1), Article e2456429. https://doi.org/10.1001/jamanetworkopen.2024.56429

Auerbach, R. P., Srinivasan, A., Kirshenbaum, J. S., Mann, J. J., & Shankman, S. A. (2022). Geolocation features differentiate healthy from remitted depressed adults. *Journal of Psychopathology and Clinical Science, 131*(4), 341–349. https://doi.org/10.1037/abn0000742

Avenevoli, S., Swendsen, J., He, J. P., Burstein, M., & Merikangas, K. R. (2015). Major depression in the national comorbidity survey–adolescent supplement: Prevalence, correlates, and treatment. *Journal of the American Academy of Child & Adolescent Psychiatry, 54*(1), 37–44.e2. https://doi.org/10.1016/j.jaac.2014.10.010

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Beard, J. I. L., & Delgadillo, J. (2019). Early response to psychological therapy as a predictor of depression and anxiety treatment outcomes: A systematic review and meta-analysis. *Depression and Anxiety, 36*(9), 866–878. https://doi.org/10.1002/da.22931

Beltzer, M. L., Meyerhoff, J., Popowski, S. A., Mohr, D. C., & Kornfield, R. (2023). Mental health self-tracking preferences of non-treatment engaged young adults with depression and anxiety: Qualitative analysis. *JMIR Formative Research, 7*, Article e48152. https://doi.org/10.2196/48152

Bisby, M. A., Titov, N., Dear, B. F., Karin, E., Wilhelms, A., Nugent, M., & Hadjistavropoulos, H. D. (2022). Examining change in the frequency of adaptive actions as a mediator of treatment outcomes in internet-delivered therapy for depression and anxiety. *Journal of Clinical Medicine, 11*(20), Article 20. https://doi.org/10.3390/jcm11206001

Bitran, A. M., Sritharan, A., Trivedi, E., Helgren, F., Buchanan, S. N., Durham, K., Li, L. Y., Funkhouser, C. J., Allen, N. B., Shankman, S. A., Auerbach, R. P., & Pagliaccio, D. (2024). The effects of family support and smartphone-derived homestay on daily mood and depression among sexual and gender minority adolescents. *Journal of Psychopathology and Clinical Science, 133*(5), 358–367. https://doi.org/10.1037/abn0000917

Bloom, P. A., Lan, R., Galfalvy, H., Liu, Y., Bitran, A., Joyce, K., Durham, K., Porta, G., Kirshenbaum, J. S., Kamath, R., Tse, T. C., Chernick, L., Kahn, L. E., Crowley, R., Trivedi, E., Brent, D., Allen, N. B., Pagliaccio, D., & Auerbach, R. P. (2024). Identifying factors impacting missingness within smartphone-based research: Implications for intensive longitudinal studies of adolescent suicidal thoughts and behaviors. *Journal of Psychopathology and Clinical Science, 133*(7), 577–597. https://doi.org/10.1037/abn0000930

Chan, C. B., Ryan, D. A. J., & Tudor-Locke, C. (2006). Relationship between objective measures of physical activity and weather: A longitudinal study. *International Journal of Behavioral Nutrition and Physical Activity, 3*, 21. https://doi.org/10.1186/1479-5868-3-21

Collins, L. M., Murphy, S. A., & Strecher, V. (2007). The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): New methods for more potent eHealth interventions. *American Journal of Preventive Medicine, 32*(5 Suppl), S112–S118. https://doi.org/10.1016/j.amepre.2007.01.022

Daly, M. (2022). Prevalence of depression among adolescents in the U.S. from 2009 to 2019: Analysis of trends by sex, race/ethnicity, and income. *Journal of Adolescent Health, 70*(3), 496–499. https://doi.org/10.1016/j.jadohealth.2021.08.026

de Angel, V., Lewis, S., White, K., Oetzmann, C., Leightley, D., Oprea, E., Lavelle, G., Matcham, F., Pace, A., Mohr, D. C., Dobson, R., & Hotopf, M. (2022). Digital health tools for the passive monitoring of depression: A systematic review of methods. *npj Digital Medicine, 5*(1), Article 1. https://doi.org/10.1038/s41746-021-00548-8

de la Torre, J. A., Vilagut, G., Ronaldson, A., Valderas, J. M., Bakolis, I., Dregan, A., … Alonso, J. (2023). Reliability and cross-country equivalence of the 8-item version of the patient health questionnaire (PHQ-8) for the assessment of depression: Results from 27 countries in Europe. *The Lancet Regional Health – Europe, 31*. https://doi.org/10.1016/j.lanepe.2023.100659

Domhardt, M., Steubl, L., Boettcher, J., Buntrock, C., Karyotaki, E., Ebert, D. D., Cuijpers, P., & Baumeister, H. (2021). Mediators and mechanisms of change in internet- and mobile-based interventions for depression: A systematic review. *Clinical Psychology Review, 83*, Article 101953. https://doi.org/10.1016/j.cpr.2020.101953

Friedrich, M. J. (2017). Depression is the leading cause of disability around the world. *JAMA, 317*(15), 1517. https://doi.org/10.1001/jama.2017.3826

Fu, Z., Burger, H., Arjadi, R., Nauta, M. H., & Bockting, C. L. H. (2021). Explaining the efficacy of an internet-based behavioral activation intervention for major depression: A mechanistic study of a randomized-controlled trial. *Clinical Psychology in Europe, 3*(3), Article e5467. https://doi.org/10.32872/cpe.5467

Funkhouser, C. J., Tse, T. C., Weiner, L. S., deLuise, D., Pagliaccio, D., Durham, K., Cullen, C. C., Blumkin, Z. K., O'Brien, C. T., Allen, N. B., & Auerbach, R. P. (2024). Study preregistration: Testing a digital suicide risk reduction platform for adolescents: A pragmatic randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry, 63*(8), 845–847. https://doi.org/10.1016/j.jaac.2024.03.012

Green, P., & Macleod, C. J. (2016). Simr: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution, 7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504

Hermens, M. L., Muntingh, A., Franx, G., van Splunteren, P. T., & Nuyen, J. (2014). Stepped care for depression is easy to recommend, but harder to implement: Results of an explorative study within primary care in the Netherlands. *BMC Family Practice, 15*(1), 5. https://doi.org/10.1186/1471-2296-15-5

Jackson, H. M., Calear, A. L., Batterham, P. J., Ohan, J. L., Farmer, G. M., & Farrer, L. M. (2023). Skill enactment and knowledge acquisition in digital cognitive behavioral therapy for depression and anxiety: Systematic review of randomized controlled trials. *Journal of Medical Internet Research, 25*, Article e44673. https://doi.org/10.2196/44673

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12–19. https://doi.org/10.1037/0022-006X.59.1.12

Johnson, P. O., & Neyman, J. (1936). Tests of certain linear hypotheses and their application to some educational problems. *Statistical Research Memoirs, 1*, 57–93.

Kanter, J. W., Puspitasari, A. J., Santos, M. M., & Nagy, G. A. (2012). Behavioural activation: History, evidence and promise. *The British Journal of Psychiatry: Journal of Mental Science, 200*(5), 361–363. https://doi.org/10.1192/bjp.bp.111.103390

Kroenke, K., Strine, T. W., Spitzer, R. L., Williams, J. B. W., Berry, J. T., & Mokdad, A. H. (2009). The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders, 114*(1), 163–173. https://doi.org/10.1016/j.jad.2008.06.026

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13). https://doi.org/10.18637/jss.v082.i13

Lehtimaki, S., Martic, J., Wahl, B., Foster, K. T., & Schwalbe, N. (2021). Evidence on digital mental health interventions for adolescents and young people: Systematic overview. *JMIR Mental Health, 8*(4), Article e25847. https://doi.org/10.2196/25847

Lewis, C. C., Boyd, M., Puspitasari, A., Navarro, E., Howard, J., Kassab, H., Hoffman, M., Scott, K., Lyon, A., Douglas, S., Simon, G., & Kroenke, K. (2019). Implementing measurement-based care in behavioral health: A review. *JAMA Psychiatry, 76*(3), 324–335. https://doi.org/10.1001/jamapsychiatry.2018.3329

Long, J. A. (2024). *Interactions: Comprehensive, user-friendly toolkit for probing interactions* [Computer software] Version 1.2.0. https://cran.r-project.org/web/packages/interactions/index.html.

Lutz, W., Arndt, A., Rubel, J., Berger, T., Schröder, J., Späth, C., Meyer, B., Greiner, W., Gräfe, V., Hautzinger, M., Fuhr, K., Rose, M., Nolte, S., Löwe, B., Hohagen, F., Klein, J. P., & Moritz, S. (2017). Defining and predicting patterns of early response in a web-based intervention for depression. *Journal of Medical Internet Research, 19*(6), Article e7367. https://doi.org/10.2196/jmir.7367

Lutz, W., Schwartz, B., & Delgadillo, J. (2022). Measurement-based and data-informed psychological therapy. *Annual Review of Clinical Psychology, 18*(1), 71–98. https://doi.org/10.1146/annurev-clinpsy-071720-014821

McNeish, D. M., & Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: A review and illustration. *Educational Psychology Review, 28*, 295–314. https://doi.org/10.1007/s10648-014-9287-x

Meeuwissen, J. A. C., Feenstra, T. L., Smit, F., Blankers, M., Spijker, J., Bockting, C. L. H., van Balkom, A. J. L. M., & Buskens, E. (2019). The cost-utility of stepped-care algorithms according to depression guideline recommendations – Results of a state-transition model analysis. *Journal of Affective Disorders, 242*, 244–254. https://doi.org/10.1016/j.jad.2018.08.024

Mohr, D. C., Zhang, M., & Schueller, S. M. (2017). Personal sensing: Understanding mental health using ubiquitous sensors and machine learning. *Annual Review of Clinical Psychology, 13*, 23–47. https://doi.org/10.1146/annurev-clinpsy-032816-044949

Moshe, I., Terhorst, Y., Philippi, P., Domhardt, M., Cuijpers, P., Cristea, I., Pulkki-Råback, L., Baumeister, H., & Sander, L. B. (2021). Digital interventions for the treatment of depression: A meta-analytic review. *Psychological Bulletin, 147*(8), 749–786. https://doi.org/10.1037/bul0000334

Rickwood, D., Webb, M., Kennedy, V., & Telford, N. (2016). Who are the young people choosing web-based mental health support? Findings from the implementation of australia's national web-based youth mental health service, eheadspace. *JMIR Mental Health, 3*(3), Article e5988. https://doi.org/10.2196/mental.5988

Rivero-Santana, A., Perestelo-Perez, L., Alvarez-Perez, Y., Ramos-Garcia, V., Duarte-Diaz, A., Linertova, R., Garcia-Perez, L., & Serrano-Aguilar, P. (2021). Stepped care for the treatment of depression: A systematic review and meta-analysis. *Journal of Affective Disorders, 294*, 391–409. https://doi.org/10.1016/j.jad.2021.07.008

Robinson, G. K. (1991). That BLUP is a good thing: The estimation of random effects. *Statistical Science, 6*(1), 15–32. https://doi.org/10.1214/ss/1177011933

Rohani, D. A., Faurholt-Jepsen, M., Kessing, L. V., & Bardram, J. E. (2018). Correlations between objective behavioral features collected from Mobile and wearable devices and depressive mood symptoms in patients with affective disorders: Systematic review. *JMIR mHealth and uHealth, 6*(8), Article e9691. https://doi.org/10.2196/mhealth.9691

Shin, J., & Bae, S. M. (2023). A systematic review of location data for depression prediction. *International Journal of Environmental Research and Public Health, 20*(11), Article 11. https://doi.org/10.3390/ijerph20115984

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. SAGE Publications.

Stamatis, C. A., Meyerhoff, J., Meng, Y., Lin, Z. C. C., Cho, Y. M., Liu, T., Karr, C. J., Liu, T., Curtis, B. L., Ungar, L. H., & Mohr, D. C. (2024). Differential temporal utility of passively sensed smartphone features for depression and anxiety symptom prediction: A longitudinal cohort study. *Npj Mental Health Research, 3*(1), Article 1. https://doi.org/10.1038/s44184-023-00041-y

Stein, A. T., Carl, E., Cuijpers, P., Karyotaki, E., & Smits, J. A. J. (2021). Looking beyond depression: A meta-analysis of the effect of behavioral activation on depression, anxiety, and activation. *Psychological Medicine, 51*(9), 1491–1504. https://doi.org/10.1017/S0033291720000239

Straczkiewicz, M., James, P., & Onnela, J.-P. (2021). A systematic review of smartphone-based human activity recognition methods for health research. *npj Digital Medicine, 4*, 148. https://doi.org/10.1038/s41746-021-00514-4

Tse, T. C., Weiner, L. S., Funkhouser, C. J., deLuise, D., Cullen, C. C., Blumkin, Z. K., … Allen, N. B. (2025). Acceptability and usability of a digital behavioral health platform for youth at risk of suicide: User-centered design study with patients, practitioners, and business gatekeepers. *JMIR Formative Research, 9*, Article e65418. https://doi.org/10.2196/65418

Webb, C. A., Stanton, C. H., Bondy, E., Singleton, P., Pizzagalli, D. A., & Auerbach, R. P. (2019). Cognitive versus behavioral skills in CBT for depressed adolescents: Disaggregating within-patient versus between-patient effects on symptom change. *Journal of Consulting and Clinical Psychology, 87*(5), 484–490. https://doi.org/10.1037/ccp0000393

Weiner, L. S., Crowley, R. N., Scheeber, L. B., Koegler, F. H., Davis, J. F., Wells, M., … Allen, N. B. (2024). Engagement, acceptability, and effectiveness of the vira digital behavior change platform: A randomized pilot trial comparing self-care and coach-supported versions in young adults at risk for depression and obesity. *JMIR Mental Health, 11*, Article e51366. https://doi.org/10.2196/51366

Zhang, Y., Folarin, A. A., Sun, S., Cummins, N., Vairavan, S., Bendayan, R., Ranjan, Y., Rashid, Z., Conde, P., Stewart, C., Laiou, P., Sankesara, H., Matcham, F., White, K. M., Oetzmann, C., Ivan, A., Lamers, F., Siddi, S., Vilella, E., … Dobson, R. J. (2022). Longitudinal relationships between depressive symptom severity and phone-measured mobility: Dynamic structural equation modeling study. *JMIR Mental Health, 9*(3), Article e34898. https://doi.org/10.2196/34898