# StayWell at Home: A Text Messaging Intervention to Counteract Depression and Anxiety during COVID-19 Social Distancing

Adrian Aguilera, Rosa Hernandez-Ramos, Alein Yaritza Haro, Claire Elizabeth Boone, Tiffany Luo, Jing Xu, Bibhas Chakraborty, Christopher Karr, Sabrina Darrow, Caroline Astrid Figueroa

Submitted to: JMIR Mental Health
on: May 07, 2021

# *Table of Contents*

**Original Manuscript** ..... **5**
**Supplementary Files** ..... **23**
Multimedia Appendixes ..... **24**
Multimedia Appendix 0 ..... **24**
CONSORT (or other) checklists ..... **25**
CONSORT (or other) checklist 0 ..... **25**

# StayWell at Home: A Text Messaging Intervention to Counteract Depression and Anxiety during COVID-19 Social Distancing

Adrian Aguilera[1] PhD; Rosa Hernandez-Ramos[1] BA; Alein Yaritza Haro[2] MPH; Claire Elizabeth Boone[2] MPH; Tiffany Luo[1] BA; Jing Xu[3, 4] PhD; Bibhas Chakraborty[3, 5, 6] PhD; Christopher Karr[7] BSc, MSc; Sabrina Darrow[8] PhD; Caroline Astrid Figueroa[1] MD, PhD

[1]School of Social Welfare, University of California, Berkeley Berkeley US
[2]School of Public Health, University of California, Berkeley Berkeley US
[3]Center for Quantitative Medicine, Duke National University of Singapore Singapore SG
[4]Data Science Programme, Division of Science and Technology Beijing Normal University-Hong Kong Baptist University United International College Zhuhai CN
[5]Department of Biostatistics and Bioinformatics, Duke University Durham US
[6]Department of Statistics and Applied Probability National University of Singapore Singapore SG
[7]Audacious Software Chicago US
[8]Department of Psychiatry and Behavioral Sciences, University of California San Francisco San Francisco US

**Corresponding Author:**
Adrian Aguilera PhD
School of Social Welfare, University of California, Berkeley
120 Haviland Hall, MC7400
School of Social Welfare, University of California, Berkeley
Berkeley
US



## *Abstract*

**Background:** Social distancing and stay-at-home orders are critical interventions to slow down person-to-person transmission of COVID-19. While these societal changes help to contain the pandemic, they also have unintended negative consequences, including anxiety and depression. We developed StayWell, a daily skills-based SMS text messaging program, to mitigate COVID-19 related depression and anxiety symptoms among people who speak English and Spanish in the United States.

**Objective:** This paper describes the changes in the anxiety and depression levels of participants in the StayWell program after 60 days of exposure to skills-based SMS text messages.

**Methods:** We used self-administered, empirically supported web-based questionnaires to assess the demographic and clinical characteristics of StayWell participants. Anxiety and depression were measured using the 2-item Generalized Anxiety Disorder (GAD-2) scale and the 8-item Patient Health Quesstionanire-8 (PHQ-8) scale at baseline and 60-day timepoints. We used paired t-tests to detect the change in PHQ-8 and GAD-2 scores from baseline to follow-up measured 60 days later.

**Results:** The analytic sample includes 193 participants who completed both the baseline and 60-day exit questionnaires. At the 60-day time point, there were statistically significant reductions in both PHQ-8 and GAD-2 scores from baseline. We found an average reduction of -1.72 (95% CI: -2.35, -1.09) in PHQ-8 scores and -0.48 (95% CI: -0.71, -0.25) in GAD-2 scores. This translated to an 18.5% and 17.2% reduction in mean PHQ-8 scores and GAD-2, respectively.

**Conclusions:** StayWell is a low-intensity, cost-effective, and accessible population-level mental health intervention. Participation in StayWell focused on COVID-19 mental health coping skills and was related to improved depression and anxiety symptoms. In addition to improvements in outcomes, we found high levels of engagement during the 60-day intervention period. Text messaging interventions could serve as an important public health tool for disseminating strategies to manage mental health. Clinical Trial: ClinicalTrials.gov Identifier: NCT04473599

(JMIR Preprints 07/05/2021:25298)
DOI: https://doi.org/10.2196/preprints.25298

### Preprint Settings

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

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

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

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

# Original Manuscript

## Title Page

**StayWell at Home: A Text Messaging Intervention to Counteract Depression and Anxiety during COVID-19 Social Distancing**

Adrian Aguilera[1,5], Rosa Hernandez-Ramos[1], Alein Y. Haro[2], Jing Xu[2,6,9], Claire E. Boone[2], Tiffany Luo[1], Bibhas Chakraborty[3,4,7,9], Christopher Karr[8], Sabrina M. Darrow[5], Caroline Figueroa[1]

[1] School of Social Welfare, University of California, Berkeley, United States of America
[2] School of Public Health, University of California, Berkeley, United States of America
[3] Center for Quantitative Medicine, Duke National University of Singapore, Singapore
[4] Department of Biostatistics and Bioinformatics, Duke University, United States of America
[5] UCSF Department of Psychiatry and Behavioral Sciences, San Francisco General Hospital, United States of America
[6] Data Science Programme, Division of Science and Technology, Beijing Normal University-Hong Kong Baptist University United International College, Zhuhai, Guangdong, China
[7] Department of Statistics and Data Science, National University of Singapore, Singapore
[8]Audacious Software
[9]Health Services and Systems Research, Duke National University of Singapore, Singapore

Title character count: 103
Abstract word count: 288
Article word count: 4783
Number of Tables: 6
Number of Figures: 1
Multimedia Appendix: 1

# Corresponding Author:
Adrian Aguilera, PhD,
120 Haviland Hall, MC7400
School of Social Welfare, University of California, Berkeley,
Berkeley, CA 94720, USA
Email: aguila@berkeley.edu



## Abstract

**Background:** Social distancing and stay-at-home orders are critical interventions to slow down

person-to-person transmission of COVID-19. While these societal changes help contain the pandemic, they also have unintended negative consequences, including anxiety and depression. We developed StayWell, a daily skills-based SMS text messaging program, to mitigate COVID-19 related depression and anxiety symptoms among people who speak English and Spanish in the United States.

**Objective:** This paper describes the changes in StayWell participants' anxiety and depression levels after 60 days of exposure to skills-based SMS text messages.

**Methods:** We used self-administered, empirically supported web-based questionnaires to assess the demographic and clinical characteristics of StayWell participants. Anxiety and depression were measured using the 2-item Generalized Anxiety Disorder (GAD-2) scale and the 8-item Patient Health Questionnaire-8 (PHQ-8) scale at baseline and 60-day timepoints. We used two-tailed paired t-tests to detect the change in PHQ-8 and GAD-2 scores from baseline to follow-up measured 60 days later.

**Results:** The analytic sample includes 193 participants who completed both the baseline and 60-day exit questionnaires. At the 60-day time point, there were statistically significant reductions in both PHQ-8 and GAD-2 scores from baseline. We found an average reduction of -1.72 (95% CI: -2.35, -1.09) in PHQ-8 scores and -0.48 (95% CI: -0.71, -0.25) in GAD-2 scores. These improvements translated to an 18.5% and 17.2% reduction in mean PHQ-8 scores and GAD-2, respectively.

**Conclusions:** StayWell is an accessible, low-intensity population-level mental health intervention. Participation in StayWell focused on COVID-19 mental health coping skills and was related to improved depression and anxiety symptoms. In addition to improvements in outcomes, we found high levels of engagement during the 60-day intervention period. Text messaging interventions could serve as an important public health tool for disseminating strategies to manage mental health.

**Trial Registration:** ClinicalTrials.gov Identifier: NCT04473599

**Keywords:** mobile health; COVID-19; text messaging; CBT; anxiety; depression; micro-randomized trials

## Introduction

The COVID-19 pandemic is a significant public health crisis that has caused devastating physical illness and concurrent mental health challenges [1]. Public health measures, including stay-at-home orders and the closure of non-essential businesses, have been necessary to reduce transmission but have also disrupted social life by limiting social activities and physical interactions with one's networks [2,3].

The societal changes to contain the COVID-19 pandemic have caused significant psychological distress worldwide in people of various backgrounds [1,4–6]. Studies show lowered psychological wellbeing and increased depressive and anxiety symptoms among the general public compared to rates before the pandemic [1]. In the United States (US), risk of depression among adults increased threefold during COVID-19 compared to before the pandemic [7]. Stressors associated with social distancing and loss of usual routines, including infection fear, financial insecurity, frustration, and a sense of isolation, had negative psychological impacts, including increased depression and anxiety symptoms [4]. These stressors also increased insomnia [8], decreased physical activity [9], and increased alcohol and substance use [10] in diverse global samples.

While the prevalence of anxiety and depression has increased in the general population, certain groups are at a higher risk of mental health disorders. Individuals with a greater risk for depression during the pandemic include those from lower socioeconomic backgrounds with insufficient economic resources, inadequate social support, and greater exposure to social stressors, such as pandemic-related job loss [7]. Furthermore, the pandemic has disproportionately affected the health of already at-risk individuals such as those from low-income backgrounds, communities of color, and non-English speaking groups [11].

Text messaging is a promising tool to deliver interventions that address the detrimental mental health effects of the COVID-19 pandemic [12–15], especially for underserved populations [16,17]. Texting, often viewed as the "workhorse" of digital and mobile health, is a widely used communication strategy that has been leveraged to deliver mental health interventions by relaying health information, skills-based messages, and self-monitoring messages to participants [16]. Text messaging interventions can help fill the gap between the need for and availability of mental health services, including behavioral health appointments, a gap that worsened during the pandemic [18]. Because 85% of all US adults and 76% of lower income adults own a smartphone, texting interventions have the capacity to reach a large and diverse group of people [19]. Expanding the reach of mental health programs is especially crucial since depression and anxiety symptoms have increased in the general population and intensified among those with existing mental health disorders and vulnerable communities.

Incorporating cognitive behavioral therapy (CBT) and skills-based text messaging interventions have proved feasible and acceptable among a diverse group of patients with affective disorders [20]. A significant body of research establishes the effectiveness of CBT as an evidence-based, first-line treatment for mental health conditions such as depression and anxiety [21]. CBT has been implemented in diverse populations, including communities of color and individuals of lower socioeconomic status. Additionally, CBT is a focused, directive, and structured form of psychotherapy, making it well suited for delivery via digital platforms such as text messaging. Electronically delivered CBT is at least as effective as face-to-face CBT at reducing mental health symptoms [22].

In April 2020, the authors of this study initiated the StayWell at Home intervention

[ClinicalTrials.gov Identifier: NCT04473599], a 60-day skills-based daily text messaging program using principles of CBT [23]. This paper assesses the effects of StayWell on symptoms of depression and anxiety in a broad adult population living in the US during the COVID-19 pandemic. Texts were based on two core components of CBT: behavioral activation (BA) and psychoeducation. BA aims to help people engage in enjoyable activities, reduce reliance on unhealthy coping mechanisms, and decrease avoidance of anxiety-provoking situations. By directing individuals to pleasurable and meaningful activities, BA can improve mood and decrease loneliness and isolation related to the pandemic. Providing psychoeducation around thoughts, feelings, and behaviors is also an important part of CBT. Messages focused on promoting adaptive cognitive approaches to pandemic-related stress and encouraging behavioral activation within the limits of social distancing. Maladaptive thoughts and behaviors can be identified and replaced to reduce the frequency and intensity of negative emotions. Further, information and reminders related to self-care, sleep, physical activity, and mindfulness may promote positive health behavior change and have beneficial effects on individuals' psychological wellbeing. We hypothesize that participants in the intervention will report fewer depression and anxiety symptoms at the end of the intervention.

# Methods

### StayWell Trial Design

StayWell is a fully remote trial and had various designs: 1) a pre/post comparison, in which we assessed depression and anxiety symptoms for all patients before and after the intervention, 2) a randomized controlled trial with two groups: Uniform Random messaging (UR) and a Reinforcement Learning (RL) messaging. Due to a coding error in the algorithm, all participants received messages randomly (UR condition). Therefore, we altered the study's main aims by focusing on the pre-post effects of participating in StayWell on depression and anxiety symptoms. University of California (UC) Berkeley Institutional Review Board reviewed and approved all study procedures.

Participants were enrolled in the StayWell trial using the HealthySMS platform— an automated text messaging platform developed by the authors (https://staywell.healthysms.org) [20]. HealthySMS has been successfully used with various low-income English and Spanish-speaking populations to send automated text messages and manage participant responses using a secure, HIPPA-compliant platform [20]. Participants received two messages daily for 60 days: one skills-based message and one message inquiring about their mood. The skills-based messages included tips on how to deal with worry and stress brought on by the COVID-19 pandemic. Half of these messages were based on behavioral activation (BA) strategies, and half were based on skill-based strategies. We developed the messages to highlight evidence-based practices used in depression and anxiety interventions that promote behavior change. Templates for these messages were developed from previous work conducted by authors SMD and AA[16]. The study team edited the original messages to fit the COVID-19 context and improve readability. The team also translated and culturally adapted messages in Spanish to expand the reach of the intervention. Messages were sent daily at a random time between 9 am and 6 pm. Three hours following the delivery of the skills-based messages, participants were sent a message inquiring about their mood on a scale of 1-9, with 9 being the best mood.

*Uniform Random message arm (UR)*

Participants received messages uniform randomly (i.e., a micro-randomized trial (MRT) design [16]), where every day during the study treatment allocation was characterized by a full factorial design with two factors: skills-based messages (M) and the time frame (T) when the message was sent. M has two levels (BA and skill-based), and T has three levels (9 am-12 pm, 12 pm-3 pm, 3 pm - 6 pm). Participants received one daily skills-based message and one daily mood message that did not vary.

**Data Collection**

Adult Spanish and English-speakers aged 18 years and older who had a mobile phone were recruited via online media advertisements on Facebook, Craigslist, and university websites (UC Berkeley and UC San Francisco) to participate. Data were collected through an online Qualtrics survey and HealthySMS. Participants were excluded if they used an online-text messaging app or were outside of the US. Online-text messaging apps are more prone to online scams and facilitate the creation of multiple phone numbers for one participant. The study lasted from April 2020 to December 2020.

*Recruitment*

We designed online ads to target vulnerable populations, including low-income groups and people of color, who are disproportionately impacted by COVID-19 in the US. Using User-Centered Design (UCD), we created 16 user personas. A persona is a fictional characterization of a user that includes specific characteristics and demographics found in the target population [24]. The personas informed the title, picture, and reason for participating in the study used on each ad and the characteristics used for detailed targeted advertising, which is available on Facebook ads.

*Enrollment*

To prevent online scams and fraud, interested subjects were sent to a Qualtrics survey to assess eligibility criteria and human identity through a built-in CAPTCHA. Eligible participants were then sent a unique link to a baseline assessment. Using a different Qualtrics survey, participants consented and answered demographic questions and other measures of interest. Upon completion, participants were enrolled in an automated text messaging intervention for 60 days. On day 61, participants were sent an exit assessment where they were asked the measures of interests originally asked at baseline. Participants were paid US$20 at the end of the study for completing study questionnaires.

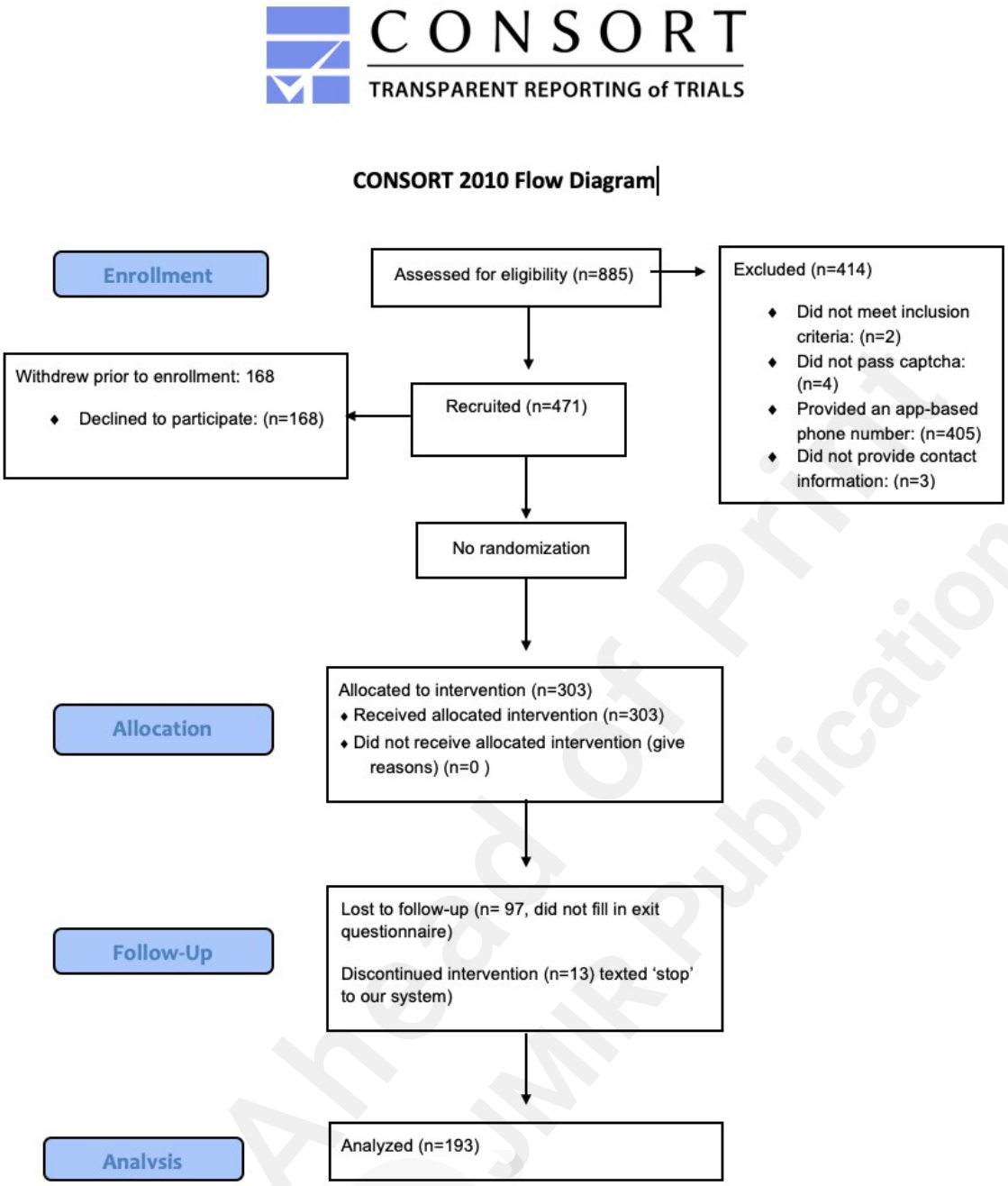


**Figure 1. Flowchart of data collection process**

## Outcome Measures

Our primary outcomes, including the 8-item Patient Health Questionnaire 8 [25] and the 2-item Generalized Anxiety Disorder 2 [26] (GAD-2) scale, were collected through a Qualtrics survey at pre-and post-intervention.

Secondarily, we were interested in assessing engagement in the intervention by measuring response rates to mood rating messages and calculating how many participants stopped the text messaging.

## Hypotheses

We will conduct a pre-post comparison among all participants. The depression score measured using

the 8-item Patient Health Questionnaire (PHQ-8) and the anxiety score measured using the 2-item Generalized Anxiety Disorder (GAD-2) scale will be improved over the 60-day study.

### Power Analysis

The sample size calculation had been performed in the protocol paper [23], which includes two aims. This paper only considers the primary aim. At a medium standardized effect size (i.e., Cohen d = 0.5), a sample size of 64 is required to detect an improvement of either the depression or anxiety score from baseline to 60-day at 80% power and 5% level of significance. The sample size of this study is 193, which is based on the secondary aim. The secondary outcome focuses on the proximal effect of daily improvement on the mood rating. However, the analysis will be presented in a separate manuscript.

## Statistical Analysis

### Main analyses

To detect the change in depression score (PHQ-8) and anxiety score (GAD-2) from baseline to follow-up measured 60 days later, we used two-tailed paired *t*-tests. The normality assumptions for the change in each score are validated by their corresponding histogram plots, which are relatively symmetric. The goodness of fit to normal distribution is validated by Anderson-Darling test while the skewness for normality is validated by the Shapiro-Wilk skewness test [27].

### Exploratory analyses

We use a simple linear regression to model the change in PHQ-8 and GAD-2 scores as a function of participants' response rates (i.e., the proportion of mood messages answered) to determine whether the improvements in depression and anxiety are predicted by engagement with the intervention. Furthermore, to determine whether any other covariates predict the effects of the StayWell intervention, we used multivariable linear regression to model the change in PHQ-8 and GAD-2 scores (i.e., scores at follow-up minus scores at baseline) as a function of demographic predictors, response rates, self-rated health, and the change in COVID-19 weekly rolling average case rates in each participants' county of residence. Demographic predictors include education (at least high school, some college, college, or graduate degree), age, gender (female, male, or other), language (English or Spanish), and employment (full-time, part-time, unemployed, or other). Self-rated health includes poor/fair, good, very good, excellent. The weekly rolling average of daily new confirmed cases per 100,000 is calculated for the day in which participants enroll and exit the program by averaging the values of that day, the three days before, and the next three days. We then used the change in COVID-19 case rates at 60-day follow-up from the baseline date for each participant.

## Results

A total of 303 people entered the study and received text messages. Of these, 12 were recruited via clinicaltrials.gov, 75 via Craigslist, 184 via Facebook, and 32 via texting the StayWell phone number. We show the distribution of baseline characteristics in Table 1. Most baseline respondents were female (76.0%) and spoke English (88.4%). While almost half of the respondents identified as White or Caucasian (47.9%), the sample is relatively diverse with 20.5% Latino(a)s, 13.2% Asian or Pacific Islander, 11.5% multiethnic, and 6.6% Black or African American participants. Of the baseline participants, 193 also completed an exit questionnaire and were included in the main analysis.

| Table 1. Baseline demographic and clinical characteristics (N = 303). | **Value** |
|---|---|
| **Age in years, Mean (SD)** | 33.3 (11.0) |
| **Gender, female, n (%)** | 230 (76.0%) |

| | |
|---|---|
| **Language, n (%)** | |
| English | 268 (88.4%) |
| Spanish | 35 (11.3%) |
| **Employment, n (%)** | |
| Full time (greater than or equal to 35 hours/week) | 137 (45.2%) |
| Part-time (less than 35 hours) | 61 (20.1%) |
| Homemaker | 28 (9.2%) |
| Unemployed | 50 (16.5%) |
| Disabled/on disability | 12 (4.0%) |
| Retired | 1 (0.3%) |
| Other | 14 (4.6%) |
| **Race/Ethnicity, n (%)** | |
| Asian or Pacific Islander | 40 (13.2%) |
| Black or African American | 20 (6.6%) |
| White or Caucasian | 145 (47.9%) |
| Latino(a) or Hispanic | 62 (20.5%) |
| Multi-Ethnic | 35 (11.5%) |
| Unknown | 1 (0.3%) |
| **Education, n (%)** | |
| Between 6th and 8th grade | 1 (0.3%) |
| Some high school | 10 (3.3%) |
| High school graduate | 29 (9.6%) |
| Some college or technical school | 94 (31.0%) |
| College graduate | 103 (34.0%) |
| Graduate degree | 66 (21.8%) |
| **Paying for basics (e.g., food, housing, medical care, and heating) is, n (%):** | |
| Very hard | 119 (39.3%) |
| Sometimes hard | 137 (45.2%) |
| Not hard at all | 47 (15.5%) |
| **Self-reported-health, n (%)** | |
| Excellent | 8 (2.6%) |
| Very good | 41 (13.5%) |
| Good | 89 (29.4%) |
| Fair | 103 (34.0%) |
| Poor | 61 (20.1%) |
| **Psychological outcomes, Mean (SD)** | |
| Depression (PHQ-8[a]) | 9.41 (5.79) |

| | |
|---|---|
| Anxiety (GAD-2[b]) | 2.71 (1.87) |
| **Impact of COVID-19 (1=completely disagree, 5=completely agree), mean (SD)** | |
| I feel lonelier | 3.55 (1.17) |
| I am running into financial issues | 3.20 (1.36) |
| I feel more stressed | 4.07 (0.971) |
| I feel more anxious | 3.89 (1.09) |

[a] PHQ-8: Patient Health Questionnaire-8, [b] GAD-2: Generalized Anxiety Disorder-2

Table 2 displays the raw scores and the distributions of change in depression (PHQ-8) and anxiety (GAD-2) scores at 60-day from baseline for respondents who completed the baseline and exit surveys, respectively. The data in Table 2 indicate that average PHQ-8 and GAD-2 scores decreased significantly from baseline to the end of the study, suggesting improvements in depression and anxiety symptoms. There was a reduction in the mean PHQ-8 and GAD-2 scores of 18.5% and 17.2%, respectively, at 60 days compared to the baseline scores. Note that the normality assumption of each score change is valid.

Table 2. Changes in PHQ-8 and GAD-2 scores for individuals who completed both the baseline and 60-day assessment (N =193).

| Measure | Scores | | | | |
|---|---|---|---|---|---|
| | Baseline score, mean (SD) | 60-day score, mean (SD) | Change from baseline, % | Mean difference (95% CI) | *P* value |
| PHQ-8[a] | 9.30 (5.70) | 7.58 (5.27) | 18.50% | -1.72 (-2.35, -1.09) | <0.001 |
| GAD-2[b] | 2.80 (1.89) | 2.32 (1.83) | 17.20% | -0.48 (-0.71, -0.25) | <0.001 |

[a] PHQ-8: Patient Health Questionnaire-8, [b] GAD-2: Generalized Anxiety Disorder-2

To evaluate the generalizability of our data in terms of anxiety and depression prevalence and symptoms in our baseline samples, we compared the clinical parameters between participants who only responded to the baseline survey versus those who responded to both the baseline and 60-day surveys (Table 3 and Table 4). Likely major depressive disorder (MDD) and likely generalized anxiety disorder (GAD) were assessed using cutoff scores of ≥10 on the PHQ-8 and ≥3 on the GAD-2, respectively. There was no statistically significant difference (all $P > 0.05$) in clinical parameters between people who only responded to the baseline survey and those who responded to the baseline and 60-day assessment. This suggests that the mental health burden was similar between our study sample and individuals who did not complete the 60-day survey.

Table 3. Comparison of the prevalence rates of risk for generalized anxiety disorder and likely major depressive disorder between subscribers who only completed the baseline survey and subscribers who completed both the baseline and 60-day surveys.

| Condition | Prevalence at baseline, n/total responses (%) | | | |
|---|---|---|---|---|
| | Subscribers who completed the baseline assessment but not the 60-day assessment, N = 303 | Subscribers who completed both the baseline and 60-day assessments, N = | Chi(2) | *P* value |

| | | | | |
|---|---|---|---|---|
| | | 193 | | |
| Likely major depressive disorder (PHQ-8[a] ≥ 10) | 137/303 (45.21) | 89/193 (46.11) | 0.039 (1) | 0.84 |
| At risk for generalized anxiety disorder (GAD-2[b] ≥ 3) | 135/303 (44.55) | 88/193 (45.60) | 0.052 (1) | 0.82 |

[a] PHQ-8: Patient Health Questionnaire-8, [b] GAD-2: Generalized Anxiety Disorder-2

Table 4. Comparison of the mean scores on the GAD-2 and PHQ-8 scales between participants who only completed the baseline survey and subscribers who completed both the baseline and 60-day surveys.

| Scale | Score at baseline, mean (SD) | | | |
|---|---|---|---|---|
| | Subscribers who completed the baseline assessment but not the 60-day assessment, N = 303 | Subscribers who completed both the baseline and 60-day assessments, N = 193 | Independent t-test | *P* value |
| PHQ-8[a] | 9.41 (5.79) | 9.30 (5.70) | 0.208 | 0.84 |
| GAD-2[b] | 2.71 (1.87) | 2.80 (1.89) | 0.520 | 0.60 |

[a] PHQ-8: Patient Health Questionnaire-8, [b] GAD-2: Generalized Anxiety Disorder-2

To determine whether the changes in PHQ-8 and GAD-2 scores remain after accounting for participants' engagement in the intervention, we use a simple linear regression model (Table 5) with participants' fraction of mood messages answered (i.e., response rates) as the main predictor. The outcome is the change in GAD-2 and PHQ-8 scores at 60-day from the baseline; thus, a negative coefficient indicates a greater improvement (larger decrease in scores) in anxiety and depressive symptoms than the reference group. We find that even when accounting for engagement in StayWell, the average improvements in both PHQ-8 and GAD-2 scores remain significant. The average improvements in PHQ-8 and GAD-2 scores controlling for engagement are -2.7 points ($p$ = 0.001) and -0.78 points ($p$ = 0.014), respectively.

Table 5. Simple linear regression model: responding predictor of the changes in PHQ-8 and GAD-2 scores at the 60-day exit from baseline for participants who completed both surveys

| **Characteristic** | **Change in PHQ-8[a]** | | | **Change in GAD-2[b]** | | |
|---|---|---|---|---|---|---|
| | Coefficient | 95% CI[1] | *P* value | Coefficient | 95% CI[1] | *P* value |
| Intercept | -2.7 | -4.3, -1.1 | 0.001 | -0.78 | -1.4, -0.16 | 0.014 |
| Response Rate (%) | 0.01 | -0.01, 0.04 | 0.2 | 0.00 | 0.00, 0.01 | 0.3 |
| Observations | 193 | | | 193 | | |
| Adjusted R2 | 0.003 | | | 0.0002 | | |
| Std. Error (df = 191) | 4.42 | | | 1.67 | | |

| | | |
|---|---|---|
| F Statistic (df = 1; 191) | 1.66 | 1.03 |

[1] CI: Confidence Interval, [a] PHQ-8: Patient Health Questionnaire-8, [b] GAD-2: Generalized Anxiety Disorder

To assess the influence of other factors (COVID-19 infection rates, self-rated health, and other demographic variables) on GAD-2 and PHQ-8 score improvements, we conducted a post hoc exploratory analysis (Table 6). This analysis also adjusts for response rates. Eight individuals lacked a valid zip code and were excluded from the analysis. Compared to females, the change in PHQ-8 score at 60-days from baseline was 2.4-points larger ($p$ = 0.01) among males, adjusting for all other covariates; this suggests that males experienced relatively less improvement in depression symptoms. Having very good self-rated health was associated with less improvement in anxiety symptoms (0.83-point higher ($p$ = 0.04) GAD-2 score) at 60-days from baseline compared to those with poor health and adjusting for covariates.

As a sensitivity analysis, we conducted two separate one-way analysis of variance (ANOVA) to assess the differences in the average change in PHQ-8 scores between genders and explore the association between self-rated health and the average change in GAD-2 scores. Results showed that gender had a significant effect on the change in PHQ-8 scores at 60-days from baseline ($F_{2,190}$ = 4.106, $p$ = 0.02). This suggests that there are true differences in the average improvement in PHQ-8 scores among male- and female-identifying participants. However, the mean change in GAD-2 scores did not differ between self-rated health categories ($F_{3,189}$ = 2.954, $p$ = 0.368). Thus, we cannot conclude that there are differences in the average improvement in GAD-2 scores by self-rated health.

Table 6. Multivariable linear regression models: demographic, clinical, and engagement predictors of the changes in PHQ-8 and GAD-2 scores at the 60-day exit from baseline for participants who completed both surveys

| **Characteristic** | **Change in PHQ-8[a]** | | | **Change in GAD-2[b]** | | |
|---|---|---|---|---|---|---|
| | Coefficient | 95% CI[1] | p-value | Coefficient | 95% CI[1] | p-value |
| ***Intercept*** | -4.84 | -8.4, -1.3 | **0.01** | -0.46 | -1.8, 0.87 | **0.50** |
| **Weekly COVID-19 Case Rates by 100,000** | -0.01 | -0.04, 0.02 | 0.60 | 0.00 | -0.01, 0.01 | 0.80 |
| **Education** | | | | | | |
| At least HS | — | — | | — | — | |
| Some College | 0.68 | -1.6, 2.9 | 0.60 | -0.62 | -1.5, 0.23 | 0.20 |
| College | 1.60 | -0.65, 3.8 | 0.20 | -0.63 | -1.5, 0.22 | 0.15 |
| Grad School | 1.20 | -1.3, 3.7 | 0.30 | -0.93 | -1.9, 0.02 | 0.06 |
| **Self-rated health** | | | | | | |
| Poor/Fair | — | — | | — | — | |
| Good | 1.30 | -0.92, 3.5 | 0.30 | 0.69 | -0.15, 1.5 | 0.11 |
| Very Good | 2.00 | -0.11, 4.1 | 0.06 | 0.83 | 0.04, 1.6 | 0.04 |
| Excellent | 1.40 | -1.1, 3.9 | 0.30 | 0.42 | -0.52, 1.4 | 0.40 |
| **Age** | -0.01 | -0.07, 0.05 | 0.80 | -0.02 | -0.04, 0.01 | 0.20 |
| **Gender** | | | | | | |
| Female | — | — | | — | — | |
| Male | 2.40 | 0.52, 4.2 | 0.01 | 0.45 | -0.24, 1.1 | 0.20 |
| Other | -2.80 | -8.0, 2.4 | 0.30 | 0.08 | -1.9, 2.0 | >0.9 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Employment** | | | | | | |
| Full Time | — | — | | — | — | |
| Part Time | 0.44 | -1.4, 2.3 | 0.60 | 0.45 | -0.25, 1.2 | 0.20 |
| Unemployed | -0.44 | -2.4, 1.5 | 0.70 | 0.13 | -0.60, 0.86 | 0.70 |
| Other | 0.00 | -1.9, 1.9 | >0.9 | -0.18 | -0.89, 0.54 | 0.60 |
| **Language** | | | | | | |
| English | — | — | | — | — | |
| Spanish | 1.00 | -1.3, 3.3 | 0.40 | 0.63 | -0.23, 1.5 | 0.15 |
| **Response Rate (%)** | 0.01 | -0.01, 0.03 | 0.40 | 0.00 | 0.00, 0.01 | 0.30 |
| Observations | 185 | | | 185 | | |
| Adjusted R2 | 0.02 | | | 0.03 | | |
| Std. Error (df = 169) | 4.41 | | | 1.66 | | |
| F Statistic (df = 15; 169) | 1.24 | | | 1.36 | | |

[1] CI: Confidence Interval, [a] PHQ-8: Patient Health Questionnaire-8, [b] GAD-2: Generalized Anxiety Disorder

The results in Table 2 show that both the depression and anxiety scores decreased over the 60-day period. In Table 5, we observe that the changes of both scores remain significant, adjusting for the response rate using a simple linear regression model, (i.e., see the p-value of intercept terms). We also find that participants' response rates are not significant predictors of improved PHQ-8 and GAD-2 scores (see Table 5 and Table 6). In Table 6, we adjust for the response rate and other demographic and clinical variables using a multivariable linear regression model, and we observe that the change in depression score remains significant, but the change in anxiety score is not significant. However, we had previously modeled the change in PHQ-8 and GAD-2 scores only adjusting for demographic and clinical variables, and the change in GAD-2 scores was also not significant [see Multimedia Appendix]. Therefore, the statistically insignificant results for the change in GAD-2 scores are not necessarily attributed to response rates.

**Engagement with the text messages.**
Participants answered the mood text messages on average 60.0% of the time (ranging from 0% to 100%). Furthermore, 21 people did not respond to any mood messages, and 13 participants opted out of text messaging by texting "STOP" to our system. The 303 baseline participants were in the study ranging from 2 to 72 days, with a mean of 59 days. Seventy participants went beyond the 60-day timeframe due to a system glitch; these participants were in the study for an average number of 63 days.

# Discussion

## Principal Findings

Participants who received the StayWell text messaging program showed improved depression and anxiety symptoms at completion of the program (60 days) on average. These results are similar to previous studies utilizing text messaging as a public mental health intervention to counteract the deleterious emotional and mental health effects of the COVID-19 pandemic in Canada [12–15]. Additionally, engagement with our texting study (two messages per day) was relatively high. Response rates averaged to 60% in the daily mood check-in, and only 4% of participants opted out of the text messages during the study. This study supports the use of text messaging as a broad-based tool for improving mental health, especially in the backdrop of a global pandemic when in-person behavioral health visits are inaccessible.

It was important to assess whether other factors (local COVID-19 infection rates, self-rated health,

and other demographic variables) influenced the positive outcomes; however, we found that improvements in GAD-2 scores were not related to other measured variables. We did find a greater improvement in PHQ-8 for female-identifying participants, but improvements held despite local infection rates or other demographic factors. These findings suggest that women may experience greater benefits in their depression symptoms from participating in the StayWell program. Nonetheless, it is particularly notable that weekly local infection rates were not related to change in outcomes. For example, it is possible that decreases in symptoms could be influenced by reductions in local infection rates and accompanying lowered concerns of infection or reduction in policies like stay-at-home orders. However, improvements remained for PHQ-8 after accounting for changes in weekly infection rates and other covariates, suggesting that the intervention improved symptoms beyond the influence of any change in the severity of the COVID-19 pandemic in participants' area of residence.

Over a third (36.3%) of participants did not complete the final assessment, and we took steps to assess how noncompletion impacted our results. First, we assessed whether reductions in depression and anxiety symptoms remained after controlling for participants' response rates, and our findings suggest that the intervention may be beneficial overall even if participants do not respond. It is possible that reading messages can be beneficial even if participants do not respond to subsequent mood ratings. On average, participants who completed the study had a significantly higher response rate to mood rating messages (70%, 42/60 responses) compared to those that dropped out (35%, 21/60 responses) ($p<.001$). Despite no significant differences in baseline mood ratings, PHQ-8, or GAD-2 scores between completers and non-completers, those who completed the final assessment showed increased daily mood ratings (0.0037, $p<0.001$), whereas those who did not complete the final assessment reported decreases in mood (0.0037-0.0078 = -0.0041, $p<0.001$). Therefore, it is possible that people with worsening symptoms respond less often or that less engagement is related with worse outcomes. It is possible that the outcomes of the study are biased based on the mood rating post-hoc analyses (attrition bias), however this contrasts with significant differences in outcomes remaining after controlling for response rate. Ideally, we would be able to assess any differences in primary outcomes (PHQ-8 and GAD-2) but we do not have those data for participants who dropped out of the study. This is a mixed picture but suggests that it is more likely that participants experiencing worsening symptoms drop out of the study more often.

As stated in the trial protocol, an additional aim of this study was to test whether a reinforcement learning algorithm could improve personalization and outcomes beyond randomly selecting messages within different categories (behavioral activation and coping skills). Given technical difficulties, we could not randomize participants into the distinct conditions, and all participants received the same intervention (random message condition). To prevent technical errors, future studies using RL should incorporate pilot work before commencing the trial to check for errors. Because we needed a quick roll-out during the pandemic, we were unable to do so. Continued research is needed to compare the effectiveness of personalizing messages using a reinforcement learning algorithm compared to a random message condition. Other studies in progress may be better suited to answer questions related to the impact of reinforcement learning models' utility for improving personalization [23].

**Strengths**

This study had a racially/ethnically diverse sample in the US that included 11.3% Spanish speakers. On average, participants entered the study with mild/moderate symptoms of depression, which improved to mild symptoms at the end of the intervention period. While reductions in PHQ-8 and GAD-2 were not large, the low-intensity intervention approach is highly scalable and can reach a large number of people. This study also shows that participants are open to receiving two messages

per day, including one mood response, while maintaining a high response rate (60%). Lastly, the longitudinal data collected provides opportunities for further inquiry into daily mood ratings and response rates.

**Limitations**

The main limitation of this study is the lack of a control group with no or an inactive intervention. Given the impacts of the COVID-19 pandemic on mental health, the study team felt it was unethical to withhold mental health support at this time. In addition, we were interested in assessing whether the intervention might be improved by applying a reinforcement learning algorithm to personalize messages, but technical errors prevented us from examining this. Lastly, a significant portion of the sample did not complete the final assessment, which may result in attrition bias and further exploration of the engagement/outcome relationship is merited.

**Implications for future studies**

Our study provides support for low-intensity text messaging interventions to improve mental health at the population level. Our data show the feasibility of sending two messages a day and asking for daily mood responses. Texting and other mobile interventions could serve as ways to identify individuals in need of more intensive interventions based on reported daily mood ratings. Future studies should continue to assess the impacts of text messaging and related mobile health interventions for mental health and continue to assess methods including machine learning to improve personalization.

**Conclusion**

Participation in a CBT-based text messaging program focused on COVID-19 cognitive flexibility, behavioral activation and acceptance was related to improved depression and anxiety symptoms. In addition to improvements in outcomes, this study reported high levels of engagement during a 60-day intervention that sent two messages per day. Text messaging interventions could serve as an important public health tool to disseminate strategies for managing mental health.

## Acknowledgments

This project was funded by the Health Resources & Services Administration (HRSA) Coronavirus Aid, Relief, and Economic Security (CARES) Grant #T1NHP39159.

## Authors' Contributions

AA provided the platform, developed intervention content, design, and writing of the manuscript
RHR contributed to intervention development, content and implementation, writing, and editing of manuscript;
AYH contributed to intervention conceptualization, data analysis, writing, and editing of manuscript;
CEB contributed to intervention conceptualization, data analysis, and editing of manuscript;
TL helped develop user-centered personas;
JX Advised on the experimental design and the statistical analysis plan and conducted analyses;
BC helped conceptualize the trial design and advised on the analysis plan;
SMD contributed to the development of the intervention content;
CF contributed to intervention development, design, writing, and analysis

## Conflicts of Interest

AA is the creator and owner of HealthySMS.

## Abbreviations

BA: behavioral activation
CBT: cognitive behavioral therapy
GAD-2: 2-item Generalized Anxiety Disorder scale
MDD: major depressive disorder
PHQ-8: 8-item Patient Health Questionnaire scale
UCD: User-Centered Design
UR: Uniform Random
RL: Reinforcement Learning



## References

1. Vindegaard N, Benros ME. COVID-19 pandemic and mental health consequences: Systematic review of the current evidence. Brain Behav Immun 2020;89:531–542. PMID:32485289

2. Bavel JJV, Baicker K, Boggio PS, Capraro V, Cichocka A, Cikara M, Crockett MJ, Crum AJ, Douglas KM, Druckman JN, Drury J, Dube O, Ellemers N, Finkel EJ, Fowler JH, Gelfand M, Han S, Haslam SA, Jetten J, Kitayama S, Mobbs D, Napper LE, Packer DJ, Pennycook G, Peters E, Petty RE, Rand DG, Reicher SD, Schnall S, Shariff A, Skitka LJ, Smith SS, Sunstein CR, Tabri N, Tucker JA, Linden S van der, Lange P van, Weeden KA, Wohl MJA, Zaki J, Zion SR, Willer R. Using social and behavioural science to support COVID-19 pandemic response. Nat Hum Behav 2020;4(5):460–471. PMID:32355299

3. Pedrosa AL, Bitencourt L, Fróes ACF, Cazumbá MLB, Campos RGB, de Brito SBCS, Simões E Silva AC. Emotional, Behavioral, and Psychological Impact of the COVID-19 Pandemic. Front Psychol 2020;11:566212. PMID:33117234

4. Brooks SK, Webster RK, Smith LE, Woodland L, Wessely S, Greenberg N, Rubin GJ. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. Lancet 2020 14;395(10227):912–920. PMID:32112714

5. Leigh-Hunt N, Bagguley D, Bash K, Turner V, Turnbull S, Valtorta N, Caan W. An overview of systematic reviews on the public health consequences of social isolation and loneliness. Public Health [Internet] 2017 Nov 1 [cited 2020 Nov 11];152:157–171. [doi: 10.1016/j.puhe.2017.07.035]

6. Dsouza DD, Quadros S, Hyderabadwala ZJ, Mamun MA. Aggregated COVID-19 suicide

incidences in India: Fear of COVID-19 infection is the prominent causative factor. Psychiatry Res 2020;290:113145. PMID:32544650

7. Ettman CK, Abdalla SM, Cohen GH, Sampson L, Vivier PM, Galea S. Prevalence of Depression Symptoms in US Adults Before and During the COVID-19 Pandemic. JAMA Netw Open [Internet] 2020 Sep 1 [cited 2020 Sep 20];3(9):e2019686–e2019686. [doi: 10.1001/jamanetworkopen.2020.19686]

8. Cellini N, Canale N, Mioni G, Costa S. Changes in sleep pattern, sense of time and digital media use during COVID-19 lockdown in Italy. J Sleep Res [Internet] 2020 May 15 [cited 2020 Nov 21]; PMID:32410272

9. Tison GH, Avram R, Kuhar P, Abreau S, Marcus GM, Pletcher MJ, Olgin JE. Worldwide Effect of COVID-19 on Physical Activity: A Descriptive Study. Ann Intern Med [Internet] 2020 Jun 29 [cited 2020 Nov 21]; PMID:32598162

10. Rajkumar RP. COVID-19 and mental health: A review of the existing literature. Asian J Psychiatr [Internet] 2020 Aug [cited 2020 Sep 9];52:102066. PMID:32302935

11. Williams DR, Cooper LA. COVID-19 and Health Equity—A New Kind of "Herd Immunity." JAMA [Internet] 2020 Jun 23 [cited 2020 Nov 2];323(24):2478. [doi: 10.1001/jama.2020.8051]

12. Agyapong VIO. Coronavirus Disease 2019 Pandemic: Health System and Community Response to a Text Message (Text4Hope) Program Supporting Mental Health in Alberta. Disaster Med Public Health Prep 2020 Oct;14(5):e5–e6. PMID:32317038

13. Agyapong VIO, Hrabok M, Vuong W, Shalaby R, Noble JM, Gusnowski A, Mrklas KJ, Li D, Urichuk L, Snaterse M, Surood S, Cao B, Li X-M, Greiner R, Greenshaw AJ. Changes in Stress, Anxiety, and Depression Levels of Subscribers to a Daily Supportive Text Message Program (Text4Hope) During the COVID-19 Pandemic: Cross-Sectional Survey Study. JMIR Ment Health 2020 Dec 18;7(12):e22423. PMID:33296330

14. Agyapong VIO, Hrabok M, Shalaby R, Vuong W, Noble JM, Gusnowski A, Mrklas K, Li D, Urichuck L, Snaterse M, Surood S, Cao B, Li X-M, Greiner R, Greenshaw AJ. Text4Hope: Receiving Daily Supportive Text Messages for 3 Months During the COVID-19 Pandemic Reduces Stress, Anxiety, and Depression. Disaster Med Public Health Prep 2021 Feb 8;1–5. PMID:33551009

15. Agyapong VIO, Shalaby R, Hrabok M, Vuong W, Noble JM, Gusnowski A, Mrklas K, Li D, Snaterse M, Surood S, Cao B, Li X-M, Greiner R, Greenshaw AJ. Mental Health Outreach via Supportive Text Messages during the COVID-19 Pandemic: Improved Mental Health and Reduced Suicidal Ideation after Six Weeks in Subscribers of Text4Hope Compared to a Control Population. Int J Environ Res Public Health 2021 Feb 23;18(4). PMID:33672120

16. Berrouiguet S, Baca-García E, Brandt S, Walter M, Courtet P. Fundamentals for Future Mobile-Health (mHealth): A Systematic Review of Mobile Phone and Web-Based Text Messaging in Mental Health. Journal of Medical Internet Research [Internet] 2016 [cited 2020 Nov 7];18(6):e135. [doi: 10.2196/jmir.5066]

17. Figueroa CA, Aguilera A. The Need for a Mental Health Technology Revolution in the COVID-19 Pandemic. Front Psychiatry [Internet] Frontiers; 2020 [cited 2020 Sep 25];11. [doi: 10.3389/

fpsyt.2020.00523]

18. Pfefferbaum B, North CS. Mental Health and the Covid-19 Pandemic. New England Journal of Medicine [Internet] Massachusetts Medical Society; 2020 Aug 6 [cited 2020 Nov 21];383(6):510–512. PMID:32283003

19. Demographics of Mobile Device Ownership and Adoption in the United States [Internet]. [cited 2021 Apr 8]. Available from: https://www.pewresearch.org/internet/fact-sheet/mobile/

20. Aguilera A, Bruehlman-Senecal E, Demasi O, Avila P. Automated Text Messaging as an Adjunct to Cognitive Behavioral Therapy for Depression: A Clinical Trial. Journal of Medical Internet Research [Internet] 2017 May 8 [cited 2021 Apr 8];19(5):e6914. [doi: 10.2196/jmir.6914]

21. Klasnja P, Hekler EB, Shiffman S, Boruvka A, Almirall D, Tewari A, Murphy SA. Micro-Randomized Trials: An Experimental Design for Developing Just-in-Time Adaptive Interventions. Health Psychol [Internet] 2015 Dec [cited 2021 Apr 8];34(0):1220–1228. PMID:26651463

22. Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, Hróbjartsson A, Mann H, Dickersin K, Berlin JA, Doré CJ, Parulekar WR, Summerskill WSM, Groves T, Schulz KF, Sox HC, Rockhold FW, Rennie D, Moher D. SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials. Ann Intern Med [Internet] 2013 Feb 5 [cited 2021 Apr 8];158(3):200–207. PMID:23295957

23. Figueroa CA, Hernandez-Ramos R, Boone CE, Gómez-Pathak L, Yip V, Luo T, Sierra V, Xu J, Chakraborty B, Darrow S, Aguilera A. A Text Messaging Intervention for Coping With Social Distancing During COVID-19 (StayWell at Home): Protocol for a Randomized Controlled Trial. JMIR Res Protoc [Internet] 2021 Jan 14 [cited 2021 Apr 8];10(1). PMID:33370721

24. Norman D, Draper SW. User Centered System Design: New Perspectives on Human-computer Interaction. 1986. ISBN:0-89859-872-9

25. Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1–3):163–173. PMID:18752852

26. Hughes AJ, Dunn KM, Chaffee T, Bhattarai J (Jackie), Beier M. Diagnostic and clinical utility of the GAD-2 for screening anxiety symptoms in individuals with multiple sclerosis. Arch Phys Med Rehabil [Internet] 2018 Oct [cited 2021 Apr 7];99(10):2045–2049. PMID:29964000

27. Shapiro SS, Wilk MB, Chen MHJ. A Comparative Study of Various Tests for Normality. Journal of the American Statistical Association [Internet] Taylor & Francis; 1968 Dec 1 [cited 2021 Jul 5];63(324):1343–1372. [doi: 10.1080/01621459.1968.10480932]

# Supplementary Files

# Multimedia Appendixes

Multivariable linear regression modeling the change in PHQ-8 and GAD-2 scores at 60-day exit from baseline for participants who completed both surveys.
URL: http://asset.jmir.pub/assets/4a258b7d22672070015ea9fe08f781d6.docx

# CONSORT (or other) checklists

Untitled.
URL: http://asset.jmir.pub/assets/99b290b3bb917f6e53bbdd414bb116af.pdf