Effectiveness of a self-care smartphone intervention for reducing symptoms of depression and anxiety in university students

Statistical Analysis Plan

Written by Olivia McGowan (OM), Ernesta Motuzaite (EM), Eirini Lamprogiannaki (EL), Gemma Lewis (GL), James A. Bisby (JB)

Introduction

This analysis plan sets out the methods of analyzing the predetermined primary and secondary outcomes for the study of assessing the effectiveness of a smartphone intervention in reducing symptoms of depression and anxiety in university students, which will be reported in the main peer reviewed paper(s) resulting from this randomized controlled trial.

The analysis and reporting of this trial will conform to the CONSORT statement. The protocol is in the process of being published.

Trial summary

Aim

This study aims to assess the effectiveness of a publicly available smartphone application intervention, My Online Therapy, in reducing symptoms of depression and anxiety in university students.

Objectives

The aim will be evaluated using an open label parallel group randomized controlled trial.

Study population

Inclusion criteria

Participants will be included if they:

- are aged 18 years or older
- a current UK university student (including undergraduate and postgraduate students)
- have access to a smartphone with iOS or Android operating system
- self-report symptoms of depression and/or anxiety (confirmed by scoring one or more on the PHQ-9 or GAD-7).

Exclusion criteria

Participants will be excluded if they report any of the following:

- a current diagnosis made by a clinician of any psychiatric disorder, including Major Depressive Disorder and Generalized Anxiety Disorder
- current use of any psychiatric medication for the treatment of a psychiatric disorder
- any past or current major medical condition
- any previous negative experiences with psychological therapy

Trial design

The study is an 8-week, open label, parallel group, randomized controlled trial.

Randomized treatments

Intervention:
Participants randomized to the intervention arm will be provided with immediate, free access to My Online Therapy's self-care smartphone application. The application holds a library of short evidence-based audio recordings that provide the listener with therapy-based skills to use.

Control:
Participants randomized to the control arm will be placed on a waiting list to receive the intervention at the end of the 8-week study period.

Sample size
Using the PHQ-9 and GAD-7 as the primary outcome measures and a medium effect size for between group mean differences (standardized mean difference) of 0.50 and estimating a standard deviation of 0.45 (e.g., Baker et al., 2018), a power calculation using GPower software determined that we would require at least 86 participants in each arm to detect a difference between the groups at the end of the trial with 85% power and alpha of 0.05. Calculating attrition estimates of 25% at 8-week follow-up (Linardon & Fuller-Tyszkiewicz, 2020) results in a minimum recruitment of 230 participants.

Randomization
Participants will be randomized with a 1:1 ratio to the intervention or control arm. To improve the balance of recruitment, we will use a block randomization method. An allocation sequence will be generated by the trial statistician (GL) using pre-specified code in Stata (StataCorp, Version 16) to allocate participants to either study arm in blocks, with block sizes ranging from 30 to 34.

Blinding
Participants, the lead researcher and members of the research team will not be blinded to group allocation due to participants being aware that they have received the psychological therapy intervention or being placed on a waitlist.

Outcomes

Primary outcomes

The joint primary outcomes for the study are depression and anxiety symptoms measured at 8-week follow-up.

Depression will be measured using the Patient Health Questionnaire-9 (PHQ-9), which is a 9-item self-administered questionnaire. Each item includes four responses ranging from not at all (0) to nearly every day (3) and scores from each item are summed to provide a total score between 0 and 27 (Kroenke, Spitzer, & Williams, 2001). If any completed questionnaires have one or two missing items, items will be replaced with the mean score from the completed items.

Anxiety will be measured using Generalized Anxiety Disorder-7 (GAD-7; Spitzer et al., 2006), a 7-item self-administered questionnaire. Each item includes four responses ranging from not at all (0) to nearly every day (3). Scores from each item are summed to provide a total score from 0 to 21. If any completed questionnaires have one or two missing items, items will be replaced with the mean score from the completed items.

Secondary outcomes

All secondary outcomes will be measured at baseline and then either weekly for the duration of the 8-week trial, or once again at week 8 only. General mood, self-compassion, mindfulness, social connectedness and adverse events will be measured each week from the beginning of the study until week 8. General wellbeing, life satisfaction and quality of life will be measured at baseline and at week 8.

General mood will be assessed using the Positive and Negative Affect Schedule (PANAS; Watson et al., 1988), a 20-item self-report questionnaire to measure general mood. The questionnaire consists of 10-items to assess positive affect (PA) and 10-items for negative affect (NA). Participants rate each item on a five-point Likert-type scale and scores for each subscale are summed to give a total score from 10-50.

Levels of self-compassion will be measured using the Self-Compassion Scale Short Form (SComS-SF; Raes et al., 2011), a 12-item self-report questionnaire that assesses understanding of and kindness toward oneself in painful circumstances which participants respond to on a five-point Likert-type scale. Scores on each item are summed to provide a total score from 12-60, with higher scores reflecting greater self-compassion.

Dispositional mindfulness will be measured using the Mindfulness Attention Awareness Scale (MAAS; Brown and Ryan, 2003) at baseline and then weekly for 8 weeks. The MAAS is a 15-item self-report questionnaire, responded to on a six-point Likert-type scale ranging from 1-6, that captures trait consciousness of internal and external states. The mean score is calculated to give an overall MAAS score, with higher scores reflecting greater dispositional mindfulness.

The extent to which participants feel connected to others will be measured by the Social Connectedness Scale (SCS; Lee and Robbins, 1995). Participants rate 20-items in the questionnaire on a six-point Likert-type scale, which are summed to give a social connectedness score ranging from 20-120. Higher scores reflect greater social connectedness.

The Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM; Evans et al., 2000) questionnaire will be used to measure general wellbeing at baseline and at the end of the study period (week 8) only. The CORE-OM is a self-report questionnaire of 34-items rated on a five-point Likert-type scale. Scores for each item are summed to give a total wellbeing score between 0 and 136, with higher scores reflecting lower wellbeing and greater psychological distress.

Global life satisfaction will be measured using the Satisfaction with Life Scale (SWLS; Diener,et al., 1985). Participants rate each of the 5-items on a 7-point Likert-type scale, which are summed to calculate a total SWLS score with scores ranging from 5-35.

The 12-item Short Form Health Survey (SF-12; Ware et al., 1996) will be used to measure participants' quality of life. Items relating to the impact of health on respondents' everyday life are responded to in varying formats including Likert-type and 'yes/no'. Standardized quality of life scores are calculated using the SF-12 algorithm, with scores above and below 50 indicating above and below average quality of life respectively.

Adverse events, should they occur, will be reported by participants on a weekly basis. Participants will respond to a questionnaire which asks whether they have experienced any adverse events as a result of the My Online Therapy app ("yes" / "no") and how severe this adverse event was ("mild", "moderate" or "severe"). Participants will be able to provide further details of any adverse events and will be given the opportunity to discuss these with a member of the research team.

Data collection

Baseline
Depressive symptoms (PHQ-9)
Anxiety symptoms (GAD-7)
General mood (PANAS)
Self-compassion (SComS-SF)
Mindfulness (MAAS)
Social connectedness (SCS)
Wellbeing (CORE-OM)
Life satisfaction (SWLS)
General health (SF-12)

Weeks 1 to 7
Depressive symptoms (PHQ-9)
Anxiety symptoms (GAD-7)
General mood (PANAS)
Self-compassion (SComS-SF)
Mindfulness (MAAS)
Social connectedness (SCF)
Adverse events

Week 8
Depressive symptoms (PHQ-9)
Anxiety symptoms (GAD-7)
General mood (PANAS)
Self-compassion (SComS-SF)
Mindfulness (MAAS)
Social connectedness (SCF)
Wellbeing (CORE-OM)
Life satisfaction (SWLS)
General health (SF-12)
Adverse events

Data entry

The data will be collected using online surveys generated using REDCap software (Harris et al., 2009). Researchers will also check for any duplicate sign-ups by removing data provided by the same email address. The data will be checked by researchers (OM, EM, EL, JB) before the analysis and if any concerns arise, consultation will be sought from trial's statistician (GL) before proceeding to data analysis. Researchers will ensure that all data meets enrolment criteria (that participants have signed a consent form and fully completed at least co-primary measures at baseline and week 8). Unsuitable data such as dropouts will be excluded from the final analysis. Upon data analysis, the data will be de-identified, leaving only participants' IDs, which are needed to link it with their vouchers, and exported to Stata (version 16; 2019). Participants' email addresses will be erased.

Statistical analyses

Primary and secondary outcomes will be analyzed using Stata v1626.

Interim analyses

A subsample of data will be extracted for analysis to be included as a MSc dissertation project for authors OM, EM, and EL. The results of these interim analyses will not be used to adapt or alter the trial protocol in any way. There are no planned analyses. However, this does not preclude the Data Monitoring Committee from requesting interim analyses.

Final analyses

All analyses will be complete case, intention-to-treat (defined as all patients randomized, analyzed according to their randomized group regardless of treatment received).

The CONSORT flow diagram (Eysenbach, 2011) will be constructed by the Trial Manager who will have logs of patients who do and do not agree to take part in the study. It will include number of patients randomized to each arm of the trial, and the numbers who have follow up data available.

Descriptive statistics

Descriptive statistics will be summarized according to randomized group, with continuous variables summarized using means and standard deviations, and categorical variables summarized using frequencies or proportions.

Analysis of primary outcomes

To evaluate the primary outcomes (PHQ-9 and GAD-7 scores), we will carry out linear regression analyses to compare mean scores of the outcome in intervention and control arms at week 8, adjusting for baseline scores to improve precision of estimates. We will use a separate model for each co-primary outcome. Two sensitivity analyses of the primary outcome will be conducted. First, we will include any variables that were imbalanced by randomized group at baseline as covariates. Second, we will include as covariates all baseline variables that are associated with missing primary outcome data. These variables will be identified using univariable logistic regression models.

Additionally, to compare total mean scores of co-primary outcomes between study arms at week 8 we will conduct separate multilevel linear regression models that will include GAD-7 and PHQ-9 scores at all follow-up time points as repeated measures outcomes, with a random intercept for individuals. Fixed effects will include randomized group, the baseline version of the relevant continuous co-primary outcome and time (as a categorical variable). We will investigate whether the treatment effect for each co-primary outcome varies by time-point by running an additional model which includes an interaction term between randomized group and time. Adjusting for covariates.

Analysis of secondary outcomes

Continuous secondary outcomes which are measured at multiple follow-up time-points will be analyzed using similar multilevel linear regression models as per primary analysis, including interaction analysis between study arm and time. Secondary outcomes measured at a single follow-up will be analyzed using linear regression models. We will conduct two sensitivity analyses for each secondary outcome. First, we will include any variables that were imbalanced by randomized group at baseline as covariates. Second, we will include as covariates all baseline variables that are associated with missing data on each secondary outcome. Associations between baseline variables and missing outcome data will be investigated using univariable logistic regression models.

Subgroup analyses

Subgroup analyses will be run to inspect if the association between intervention (exposure variable) and depression and anxiety symptoms (modelled as separate co-primary outcomes)

differs according to demographic variable, including level of study, ethnicity, and socioeconomic status, and the duration/frequency of application being used (i.e., once per week vs twice per week). Analyses will use the same models used for the primary analysis, including the addition of an interaction term between randomized group and the subgroup variable. A separate model will be run for each subgroup variable and each outcome.

Per protocol analysis

We will inspect the treatment effect on co-primary outcomes when restricting the sample to participants in the intervention group who have used the app at least once per week.

References:

Baker, D., Katantzis, N., Rickwood, D., Rickard, N. (2018). A randomized controlled trial of three smartphone apps for enhancing public mental health. *Behaviour research and therapy*, 109, 75-83.

Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: mindfulness and its role in psychological well-being. *Journal of personality and social psychology*, *84*, 822.

Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of personality assessment*, *49*(1), 71-75.

Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., McGrath, G. (2000). CORE: clinical outcomes in routine evaluation. *Journal of Mental Health*, *9*(3), 247-255.

Eysenbach, G. (2011). CONSORT-EHEALTH: Improving and Standardizing Evaluation Reports of Web-based and Mobile Health Interventions. *Journal Of Medical Internet Research*, *13*(4), e126. doi: 10.2196/jmir.1923

Lee, R. M., & Robbins, S. B. (1995). Measuring belongingness: The social connectedness and the social assurance scales. *Journal of counseling psychology*, *42*(2), 232.

Linardon, J., & Fuller-Tyszkiewicz, M. (2020). Attrition and Adherence in Smartphone-Delivered Interventions for Mental Health Problems: A Systematic and Meta-Analytic Review. *Journal of Consulting and Clinical Psychology, 88*(1), 1-13

Neff, K. D. (2003). The development and validation of a scale to measure self-compassion. *Self and identity*, *2*(3), 223-250.

Raes, F., Pommier, E., Neff, K. D., & Van Gucht, D. (2011). Construction and factorial validation of a short form of the Self-Compassion Scale. Clinical Psychology & Psychotherapy. 18, 250-255.

Spitzer, R., Kroenke, K., Williams, J., & Löwe, B. (2006). A Brief Measure for Assessing Generalized Anxiety Disorder. *Archives Of Internal Medicine*, *166*(10), 1092. doi: 10.1001/archinte.166.10.1092

Ware Jr, J. E., Kosinski, M., & Keller, S. D. (1996). A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Medical care*, 220-233.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology*, *54*(6), 1063.

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9. Journal of General Internal Medicine, 16(9), 606-613. doi:10.1046/j.1525-1497.2001.016009606.x