Date: 24 June 2021
Version number: 1.6

D7.3 Statistical Analysis Plan (including Economic Analysis)

ECoWeB Project (754657)

Version: 1.6
24 June 2021

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 754657

DOCUMENT CONTROL

Project:	ECoWeB (754657)
Project Title:	Assessing and Enhancing Emotional Competence for Well-Being (ECoWeB) in the Young: A principled, evidence-based, mobile-health approach to prevent mental disorders and promote mental well-being Plan.
Date:	24 June 2021
Author:	Dr Fiona Warren (Clinical Trials Unit, University of Exeter)
Reviewer	Prof Rod Taylor (University of Glasgow)
Activity	WP7, Task 7.9 - Statistical and economic analysis
Dissemination level	PU
Filename:	ECoWeB_D7.3_Statistical Analysis Plan_UNEXE_v1.6

DISSEMINATION LEVEL

PU	Public, fully open access
RE	Restricted to a group specified by the ECoWeB Consortium (including the Commission Services)
CO	Confidential: only for members of the ECoWeB Consortium (including the Commission Services)

Document History

Version	Date	Author	Comments/Change Details
1.0	6th June 2019	Dr Fiona Warren (Clinical Trials Unit, University of Exeter)	Finalised version of Statistical Analysis Plan (including Economic Analysis)
1.1	7th June 2019	Dr Fiona Warren (Clinical Trials Unit, University of Exeter)	Updated to include detail under Section 8 - Health economic analysis plan
1.2	28th June 2019	Dr Fiona Warren (Clinical Trials Unit, University of Exeter)	Insertion of front cover, table of contents. Minor formatting changes
1.3	17 December 2020	Dr Fiona Warren (Clinical Trials Unit, University of Exeter)	Amendments related to effects of COVID-19 on mental health during the trial period

1.4	18/03/2021	Dr Fiona Warren (Clinical Trials Unit, University of Exeter)	Definition of window of opportunity to provide outcome data using link
1.5	17/06/2021	Dr Fiona Warren (Clinical Trials Unit, University of Exeter)	Strategies to address app outage
1.6	24/06/2021	Dr Fiona Warren (Clinical Trials Unit, University of Exeter)	Strategies to address app outage

TABLE OF CONTENTS

CONTENTS

ECoWeB Statistical Analysis plan

This Statistical Analysis Plan (SAP) was finalised on 06 June 2019; version number 1.0.

1. Administrative information

1.1 Trial title and trial registration

The full title of the trial is "Assessing and Enhancing Emotional Competence for Well-Being (ECoWeB) in Young Adults: A principled, evidence-based, mobile-health approach to prevent mental disorders and promote mental well-being". The trial registration number is: **ISRCTN** NCT04148508. This information is also set out in Table 1.

1.2 Trial roles and responsibilities

Roles and responsibilities for the ECoWeB trial are set out in Table 1

Table 1 Roles and responsibilities

Trial full title	Assessing and Enhancing Emotional Competence for Well-Being (ECoWeB) in the Young: A principled, evidence-based, mobile-health approach to prevent mental disorders and promote mental well-being
Trial registration number	**ISRCTN** NCT04148508
Trial chief investigator	Prof Ed Watkins
Exeter CTU co-director	Prof Rod Taylor
Trial manager	Dr Lexy Newbold
Trial statistician	Dr Fiona Warren
SAP author	Dr Fiona Warren
Exeter CTU involvement	Statistics, randomisation, data management; trial management
Trial funder	European Commission
Trial sponsor	University of Exeter

1.3 Preparation of the statistical analysis plan

The SAP has been prepared in accordance with ICH-9 statistical guidelines for clinical trials [1] and JAMA Guidelines for the Content of Statistical Analysis Plans in Clinical Trials [2], and with guidance from the Exeter Clinical Trials Unit (ExeCTU) Standard Operating Procedure (SOP) for Statistical Principles. Results are to be reported in accordance with the CONSORT checklist for trials [3], the CONSORT extension for non-pharmacologic trials [4], and the CONSORT extension for multi-arm trials. [5]

1.4 Statistical analysis plan and trial protocol

This SAP is not intended to be a standalone document and should be used in conjunction with the trial protocol (version 1.3 dated 05/07/2021).

1.5 Statistical analysis plan roles and responsibilities

The roles and responsibilities of those who produced this SAP are set out in Table 2.

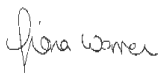
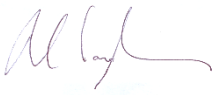
Table 2 Statistical analysis plan roles and responsibilities

Name	Affiliation(s)	Role
Dr Fiona Warren	Statistician, ExeCTU	Author
Prof Rod Taylor	Co-director, ExeCTU	Reviewer

1.6 Signatures

The SAP has been signed off by the TMG. The signatures of the SAP author, senior statistician and chief investigator are set out in Table 3.

Table 3 Signatures

Role	Author of SAP	Co-director, ExeCTU	Chief Investigator
Name	Fiona Warren	Rod Taylor	Ed Watkins
Signature			
Date (DD-MMM-YYYY)	6/6/2019	5/6/2019	3/6/2019

1.7 Revisions to the statistical analysis plan

Revisions to this SAP will be set out in Table 4. Revisions can be made at any time during the trial and will be approved by the TMG. Revisions following scrutiny of the data and unblinded interpretation of the analyses will be considered to be post hoc and will require strong justification.

Table 4 Statistical Analysis Plan revisions

SAP revision number	Date of revision *(DD-MM-YYYY)*	Timing of SAP revision in relation to trial progress	Details of the revision	Justification for the revision
1.0	06/06/2019	Set-up		
1.1	07/06/2019	Set-up	Updated to include detail under Section 8 - Health economic analysis plan	Inclusion of HEAP
1.2	28/06/2019	Set-up	Insertion of front cover, table of	Minor formatting changes.

			contents. Minor formatting changes	
1.3	17/12/2020	Participant recruitment	Section 6.7: Addressing COVID-19 related issues Section 6.8: Addressing protocol violations	Concerns related to effects of COVID-19 on mental wellbeing were not addressed in the original SAP (written prior to the COVID-19 pandemic).
1.4	18/03/2021	Participant recruitment	Section 6.1 Trial outcomes Section 6.3 Analysis methods	Duration of time that the link for submission of outcome data collection at each follow-up timepoint has been confirmed to be just under 3 weeks. This is now reflected in the SAP.
1.5	17/06/2021	Participant recruitment	Section 6.9 Addressing protocol violations due to app outage	Response to app outage for ~4 weeks in March 2021.
1.6	24/06/2021	Participant recruitment	Section 6.9 Addressing protocol violations due to app outage	Response to app outage for ~4 weeks in March 2021.

2. Trial description

2.1 Background and rationale

The ECoWeB project is a multi-country project that aims to develop a personalised self-help app to support mental health and well-being in young adults, and to evaluate the performance of the app against usual practice and an active but non-personalised app intervention; all interventions will include use of the app for self-monitoring. Mental and emotional wellbeing in young adults are known to be associated with mental health disorders, such as anxiety and depression, and to be related to non-health outcomes such as education and employment achievements. The ECoWeB programme incorporates two clinical trials, ECoWeB-PROMOTE and ECoWeB-PREVENT, to be run in parallel. Both trials include the same three interventions and collect the same outcome data. The aim of these trials is to examine the effectiveness of two active forms of psychological self-help delivered by means of an app, both of which are compared with a usual care control group.

2.2 Objectives

Both trials seek to evaluate mental health and emotional wellbeing within their respective participant samples comparing two active interventions (one personalised, one uniform across participants) against usual care (self-monitoring of emotional well-being via an app) and against each other. In addition, a health economic evaluation will be performed (Section 8).

3. Study methods

3.1 Trial design

Both PROMOTE and PREVENT are 3-arm parallel group trials, with individual randomisation in a 1:1:1 ratio. For both trials, an external pilot will be carried out to identify any issues in the administrative aspects of the trial, such as participant recruitment, consent procedures for under 18s (in countries where parental consent is required for trial participation of under-18s. The data from this pilot will not be combined with the results of the main trials, which are fully powered definitive trials. Furthermore, the external pilot has no stopping criteria that would, if fulfilled, lead to the main trial being cancelled prior to commencement.

The usual care intervention comprises only self-monitoring using the app. The personalised active intervention consists of usual care plus two out of four modules for emotional competence self-help training, selected for each participant. The standard active intervention consists of usual care plus a self-help intervention based on CBT principles, also delivered via the app.

The ECoWeB trials will be performed within the framework of a cohort multiple RCT (cmRCT). A top level cohort of potential participants, whose eligibility for either PROMOTE or PREVENT has been established, will be recruited. These participants will consent to participating in the cohort and to being offered potential treatments if applicable via the screening website. Based on their baseline scores on eight measures (collected at baseline), to be converted to z-scores (using means and standard deviations derived from a validation study across all four sites), the participant will be allocated to either PROMOTE (lower risk of mental health problems) or PREVENT (higher risk of mental health problems). Full details of the process for allocation of participants to either PROMOTE

or PREVENT are reported elsewhere (Trial selection, randomisation, selection of emotional competence intervention components document).

Following selection of eligible trial, participants will be randomised to one of the three interventions (Section 3.2). Those participants who are allocated to receive the personalised intervention will receive two of the four available modules (Achievement appraisal; Social appraisal; Rumination; Emotion knowledge and Perception), with six module combinations possible. Module selection will be based on the participant's responses to 11 measures across the 4 components - full details are provided elsewhere (Trial selection, randomisation, selection of emotional competence intervention components document).

3.2 Trial population

The two trials differ with regard to their target populations. Both trials include young adults aged 16–22 years. The PROMOTE trial will recruit participants who do not indicate potential increased risk of poor mental health at the baseline assessment based on emotional competence scores. The PREVENT trial will recruit participants who indicate potential increased risk for mental illness at the baseline assessment based on emotional competence scores. Both trials will exclude people with a past history of mental health problems, including depression and self-harm, and those with current mental health disorders, including those receiving psychological therapy or psychiatric medication. Full inclusion/exclusion criteria are set out in the trial protocol v.1.3. The aim of these trials is to examine the effectiveness of two active forms of psychological self-help delivered by means of an app, both of which are compared with a usual practice control group. Both trials will be delivered in four countries (UK, Germany, Spain, Belgium), i.e. all participants will be resident in one of these four countries at baseline; within each country, participants will have the option to communicate with the trial (via the trial website that will be used to recruit participants and conduct follow-up assessments and via the app, which will be used for intervention delivery) in one of four languages (English, German, Spanish, Dutch).

3.3 Randomisation

Within both PROMOTE and PREVENT separately, participants will be allocated to one of the three interventions in a 1:1:1 ratio using a minimisation algorithm to promote balance across the arms overall and within minimisation factors. The minimisation factors to be used are gender (categorised as male:female:both:neither anticipated to be recruited in a 20:70:5:5 ratio), age (categorised as 16–17:18–22, anticipated to be recruited in a 20:80 ratio), and country of residence (UK:Germany:Spain:Belgium, anticipated to be recruited in a 28:28:28:18 ratio). The first 50 randomisations (for both trials) will be performed using simple randomisation; subsequent randomisations will use the minimisation algorithm, taking account of the previous allocations. Further details of the randomisation procedure are provided elsewhere (Trial selection, randomisation, selection of emotional competence intervention components document).

Randomisation across all countries will be carried out by ExeCTU. The randomisation algorithm will be automatically invoked when the participant has submitted the required baseline data and provided consent. The participant will be immediately informed of the allocated treatment and the downloaded version of the app will provide the appropriate intervention. There is no possibility for a

participant to receive a version of the app other than the version to which the participant was randomly allocated.

3.4 Sample size

PROMOTE and PREVENT have different primary outcomes, therefore the sample size for each trial was calculated using the appropriate outcome. The primary outcome for PROMOTE is WEMWBS, which has a minimum clinically important difference (MCID) of 3.0 units and a standard deviation (SD) of 11.3 units. Using these values, with 90% power and a statistical significance threshold of 0.05, the sample size required for a 2-arm comparison would be 300 participants per arm. Accounting for 40% attrition at 3-month follow-up (the primary follow-up timepoint), a total of 500 participants per arm are required. Therefore, across the three arms of the trial, 1500 participants are required.

The primary outcome for PREVENT is PHQ-9, which has an established MCID of 5.9 and SD of 5.4. Using the same power requirements as for PROMOTE, the sample size required per arm is 93 participants. Accounting for 40% attrition at 3-month follow-up, 155 participants per arm are required, producing a total of 465 participants across the three arms.

A top level cohort, comprising young adults potentially eligible for either PROMOTE or PREVENT is to be recruited. It is anticipated that within such a top level cohort, 70% will be eligible for PROMOTE and 30% for PREVENT. Given that 1500 participants are required for PROMOTE, a top level cohort of 2142 potential participants will be required. The remaining 30% of the cohort, 642 participants, exceeds the required sample size for PREVENT.

For all participants recruited within the first 2 months from start of recruitment (anticipated to be approximately 450), we will calculate loss to 3-month follow-up and should it exceed 40% (as allowed for in the sample size calculation), we will re-adjust the sample size accordingly. Note that in the event of loss to follow-up being less than 40%, the sample size will not be reduced.

Full details of the sample size calculations, including references for the parameters used, are set out in Protocol v.1.3.

3.5 Framework

Both PROMOTE and PREVENT are fully powered superiority trials; both trials seek to evaluate superiority of the personalised intervention compared with both usual care and a standard active intervention. The SAP from this point onwards refers to both PROMOTE and PREVENT as they are identical in their analysis methods; the only difference between the two is in the choice of which of an identical set of outcome measures constitutes the primary outcome measure.

3.6 Statistical interim analysis and stopping guidelines

No interim analyses will be performed for efficacy or harms. As the intervention is considered to be low risk to participants, there are no formal guidelines for early termination of the trial due to potential for harm to participants. All adverse and serious adverse events will be reported to the TSC and DMEC for their consideration; if the TSC and DMEC consider that there is sufficient cumulative evidence of harm to participants due to the intervention(s), the trial will be discontinued. Also, there

are no guidelines for early termination due to futility (inability to achieve statistical significance for a treatment effect) or achieving significant results prior to full data analysis.

3.7 Timing of analyses

We anticipate performing all analysis following final database lock, when all follow-up data (up to and including 12-month follow-up) has been entered and cleaned. However, in the event that follow-up takes longer than expected (for example, if the sample size is increased due to higher loss to follow-up than was accounted for in the power calculation), it may be necessary to perform analysis of data up to 3-month follow-up and then perform analyses that include 12-month follow-up data at a later time.

Participants' timing throughout the trial will be relative to their trial allocation and randomisation, which will occur immediately following collection of baseline data.

3.8 Timing of outcome assessments

Follow-up timepoints are at 1-, 3- and 12-month follow-up (3-month follow-up is the primary timepoint). Each participant will have a scheduled follow-up date for each of the three follow-up timepoints, to be allocated on date of randomisation. Data collection will be sought within a 14-day window (within 7 days before or after the scheduled date) for 1-month follow-up. Data collection will be sought within a 28-day window (within 14 days before or after the scheduled date) for 3- and 12-month follow-up. Data collected outside this window will not be included in the primary analysis, although sensitivity analyses including all collected data will be performed.

3.9 Allocation masking

Participants will not be specifically informed of their treatment allocation. Participants will be informed prior to consenting to participate that they may be randomly selected to receive additional support and education in mental health and wellbeing via the app. Therefore, it is reasonable to assume that participants allocated to the usual care intervention who do not receive any support or information vie the app will be able to deduce that they are in a control arm, although this should be mitigated to some extent by the cohort trial design. Similarly, participants allocated to either of the two active interventions will be aware that they have been selected to receive additional support and education, and may be able to determine whether they are in the personalised intervention arm or the standard intervention arm. The issue of differential participant awareness of intervention arm will be considered when interpreting the results.

All outcome data is participant reported and will be collected via the website with no direct contact between participants and researchers. Contact between a participant and researcher would only occur if the researcher's support is required for concerns about risk of harm to the participant or for technical support with use of the app or as an additional attempt to obtain primary outcome data. Thus, there is the possibility for inadvertent unblinding of the researcher if such contact occurs. Any cases of researcher unblinding will be logged. There is provision for data collection to be conducted by telephone if a participant is unwilling or unable to use the website; such instances are anticipated to be rare, and in such cases, data collection will be performed by a blinded researcher. Data collected by telephone (as opposed to via the website, which is standard) will be recorded as such.

There may be extremely rare instances where data collection by telephone is performed by an unblinded researcher, and such cases will be recorded as such.

Following database lock on completion of 3-month data collection, the database will be made available to the trial statistician with masked treatment allocation. The final database with completed 12-month follow-up data will be made available to the trial statistician following database lock; at this point the trial statistician will be aware of allocations.

3.10 Trial flow of participant events

The flow of participant events within the trial is set out in discrete steps below.

1. Potential participant visits trial website, expresses interest in the trial and provides baseline screening data to allow eligibility (including age and country of residence) to be assessed.
2. The potential participant will be invited to provide data on the emotional competence measurements that will enable the cohort member to be allocated to either PROMOTE or PREVENT.
3. If the potential participant is eligible, the potential participant will be invited to consent to join the top level recruitment cohort (2142 participants) who will provide self-monitoring data via the app and to participate in the appropriate trial within the cohort. (If the potential participant is not eligible due to previous or current mental health conditions, signposting to sources of support will be provided.)
4. The cohort member will be allocated to PROMOTE or PREVENT.
5. If cohort member consents to participate in the allocated trial, the participant will be randomised to treatment allocation and if allocated to the personalised intervention, the appropriate components will be selected based on the data provided previously.
6. Access to the app for intervention delivery is provided as soon as the participant has been randomly allocated to intervention.
7. Follow-up at 1 month.
8. Follow-up at 3 months.
9. Follow-up at 12 months.

4. Statistical principles

4.1 Confidence intervals and p-values

All inferential analyses will be reported using 95% confidence intervals and p-values, with the threshold for statistical significance set at 0.05. No formal testing for multiple comparisons will be performed (whether for the multiple comparisons across the three trial arms or for multiple comparisons across the primary and secondary outcomes); the p-values for the primary outcomes will be interpreted first, and the p-values for the secondary outcomes will be interpreted in the light of the overall results. Global p-values will be used where there are two or more contrasts being examined in the same model; 95% confidence intervals will be used to interpret individual contrasts.

4.2 Intervention adherence and protocol deviations

For participants receiving the active interventions, we will descriptively report data on intervention adherence. Protocol deviations will be reported narratively within the trial report, with descriptive analysis only if required. With regard to treatment allocation, there is no possibility of a participant accessing one of the other active interventions, i.e. other than the allocated intervention.

4.3 Analysis populations

The primary analysis will on the intention to treat (ITT) basis (i.e. participants will be included in the analyses according to their randomised allocation) and will include observed data only). There is no possibility for a participant to receive an active intervention if this was not the participant's randomised allocation. The only possibility for a participant failing to receive the randomised intervention is if the participant did not adhere. An `as treated' analysis is not required as there is no possibility for a participant to receive a non-allocated intervention.

5. Trial population

5.1 Screening data

Screening data for age, gender and country will be available for participants who are eligible to enter the trial (either PROMOTE or PREVENT) but do not consent to participate, and will be reported descriptively.

5.2 Eligibility

Participants will be eligible for the overall cohort if they are (i) aged between 16–22; (ii) resident in the UK, Germany, Spain or Belgium; and (iii) have no past history of mental health disorders or current mental health disorders. Further eligibility criteria are set out in Protocol version 1.3. Participants will be eligible for PROMOTE if they are at lower risk of mental health problems as determined by their responses to eight measures of emotional competence; participants will be eligible for PREVENT if they are at higher risk of mental health problems as determined by their responses to these measures (full details of the deterministic allocation method to either PROMOTE or PREVENT are set out elsewhere: Trial selection, randomisation, selection of emotional competence intervention components document).

5.3 Recruitment

It is anticipated that each country will recruit an average of 14 potential participants (eligible for PROMOTE/PREVENT and consenting to participate) per week. Recruitment will take place over a period of 9 to 12 months within the 4 countries, therefore anticipated recruitment is 2240 to 2912 participants ((14 x 4 x 40) to (14 x 4 x52)), exceeding the required number for the top level cohort (2142 potential participants).

Outlines of the proposed CONSORT flow diagrams are set out below (Figures 1 and 2).

5.4 Withdrawal and loss to follow-up

If a participant formally withdraws from the trial (i.e. a participant contacts the trial team to request withdrawal from the trial), the trial team will ask the participant if the participant is willing to continue to provide outcome data, even if the participant no longer wishes to continue to adhere to the randomised allocation. Also, the trial team will ask the participant if the participant is willing to allow the trial to use the data already provided. If the participant is not willing to allow the trial to use the data already provided, the data will be permanently deleted in accordance with GDPR. The time of withdrawal from the trial will be noted; a reason for trial withdrawal will also be requested. Participants will receive reminders to complete outcome measures at 1-, 3- and 12-month follow-up. Participants who do not complete outcome measures for a specified follow-up will be requested to provide follow-up data for subsequent follow-up timepoints, unless they have contacted the trial team to formally withdraw from the trial.

Participants found to be ineligible following randomisation will be withdrawn from the trial; the numbers of ineligible withdrawn participants will be reported by trial arm with reason for ineligibility.

5.5 Baseline participant characteristics

Participant characteristics at baseline will be reported descriptively by treatment arm (Table A.1). Baseline characteristics will be compared across treatment arms to assess balance; for those characteristics that appear unbalanced (using guidelines of a difference across arms ≥ 1 SD for continuous variables or ≥10 percentage points for categorical variables) across arms will be included as a covariate in analyses (if thought predictive of outcome and at the discretion of the statistician and chief investigator).

Figure 1 CONSORT flow diagram for ECoWeB-PROMOTE

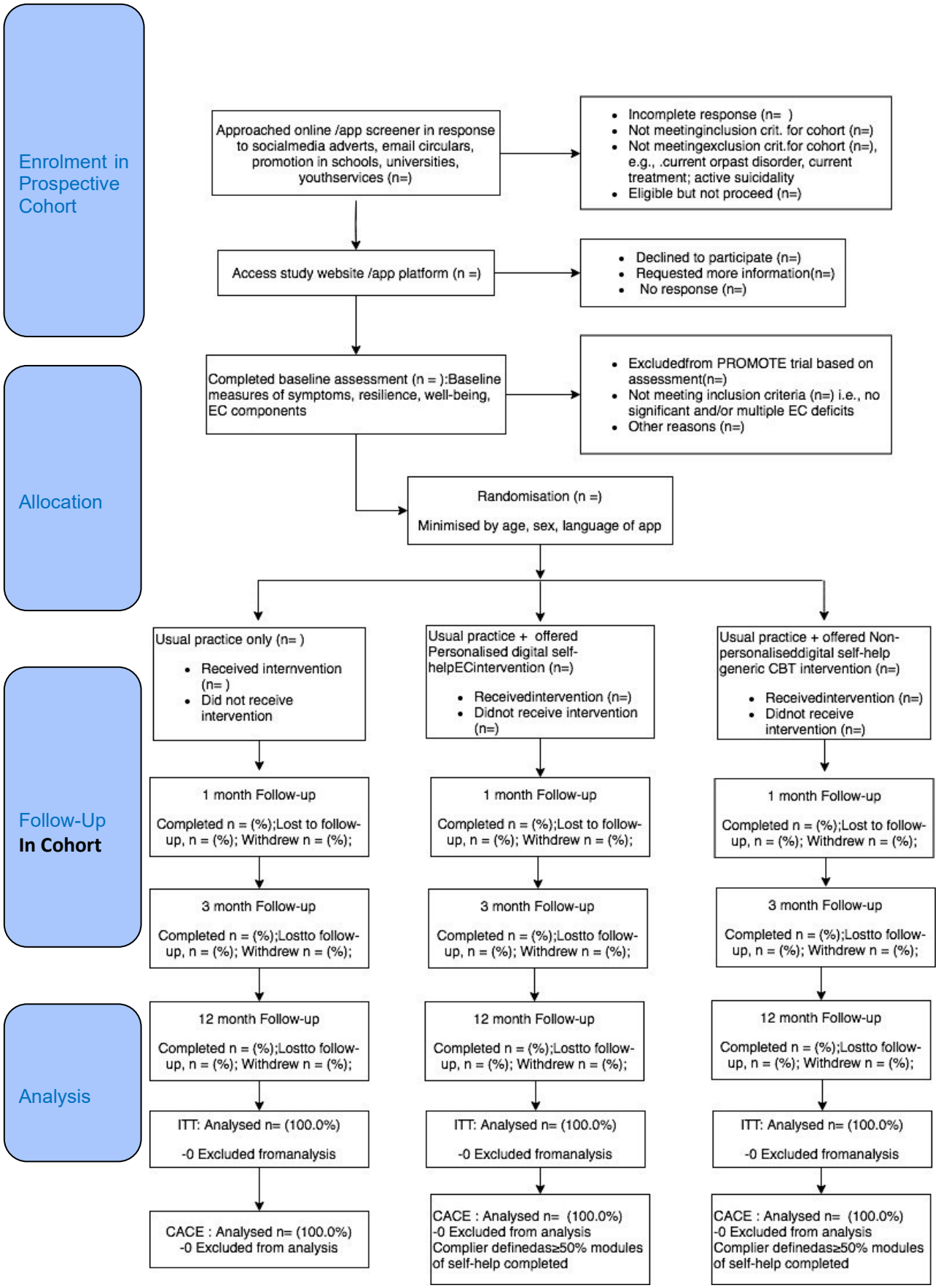
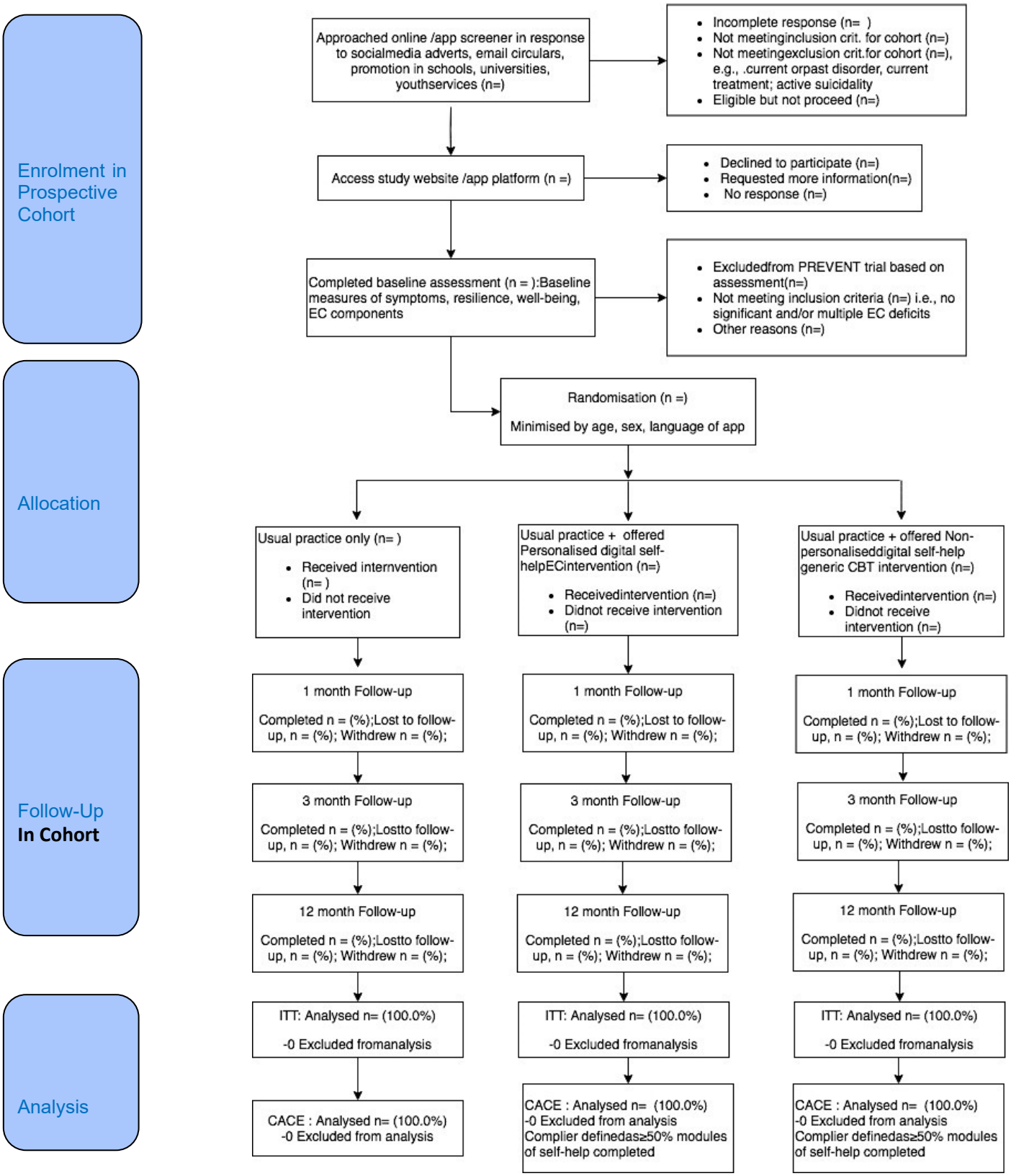


Figure 2 CONSORT flow diagram for ECoWeB-PREVENT



5.6 Adherence to intervention and personalised intervention components

Measures of intervention adherence for the active interventions will include the following.

1. Number of uses of the app
2. Percentage of relevant self-help package completed
3. Number of challenge tasks completed
4. Number of times the tools are practiced (in total across all tools and for each tool)
5. Number of days used (in total and longest streak of consecutive use)
6. Number of days that daily mood rating is completed (maximum of once per day)
7. Number of times that ecological momentary ratings are completed
8. Overall time spent on the app

For the usual care arm, intervention adherence will include level of self-monitoring via the app. Adherence measures will be reported descriptively by trial arm (Table A2). For the personalised active intervention we will report the numbers of participants who receive each of the four components (achievement appraisal, social appraisal, rumination, emotion knowledge and perception), and also the numbers of participants who receive each of the six possible combinations of two of the four components (Table A.3).

6. Statistical analysis

6.1 Trial outcomes

All trial outcomes will be collected for both PROMOTE and PREVENT; the only difference between the trials in terms of outcome variables is in the choice of primary outcome measures. For PROMOTE, the primary outcome measure is WEMWBS, whereas for PREVENT the primary outcome measure is PHQ-9.

Outcomes for PROMOTE and PREVENT include

1. WEMWBS (primary outcome measure for PROMOTE)
2. PHQ-9 (primary outcome measure for PREVENT)
3. WSAS
4. GAD-7
5. EQ5D-3L
6. ADSUS-adapted (health economic evaluation)
7. MINI-IPIP-N

Full details of these outcomes are set out in the Data Management Plan (more specifically Data Dictionary appendix), including the following items.

1. Full name of outcome measure
2. Abbreviation
3. Description of what is being measured
4. Type of measure: screening, outcome, mechanistic
5. Directionality (high scores indicate more positive or negative outcome)
6. Number of items
7. Item ranges

8. Overall score ranges
9. Subscales with number of items, item numbers, range
10. Item numbers for items that require reverse coding (i.e. the meaning of the item is expressed in the opposite direction from the overall scale)
11. Guidelines to address individual item missingness
12. MCID (if available)
13. Metric used for calculation (e.g. mean, total, percentage, algorithm)

All outcomes are questionnaire based continuous measures. The primary follow-up timepoint is 3 months, therefore the primary outcome for PROMOTE is WEMWBS at 3-month follow-up, for PREVENT the primary outcome is PHQ-9 at 3-month follow-up.

For ease of interpretation, all continuous outcomes will be reported using the directionality of higher scores indicating a more positive outcome. Therefore, some scales will be reversed by comparison to their standard usage (see the Data Management Plan for more details on outcome variables and their scoring procedures).

For each follow-up timepoint, there will be an initial email sent to the participant to request outcome data, followed by a reminder after 7 days. The link to submit outcome data will expire after 2 weeks following the reminder, resulting in a time window for data collection of just under 3 weeks. It is possible that some data collection may occur outside the time window, e.g. data collected by telephone.

6.1.1 Derived variables

For the primary outcome in each trial we will create binary outcomes to indicate whether the participant (i) had an improvement ≥MCID from baseline and (ii) had a deterioration ≥MCID from baseline. These outcomes will be included in the inferential analyses for the ITT population using observed data only.

6.2 Purpose of the analyses

The purposes of the statistical analyses are as follows:

1. to provide descriptive baseline data by trial arm;
2. to report attrition at all follow-up timepoints;
3. to report individual outcome missingness at all follow-up timepoints;
4. to report adherence to the intervention in the active treatment arms;
5. to evaluate effectiveness of the active interventions relative to control, and the personalised active intervention relative to the standard active intervention, using inferential analyses;
6. to perform additional analyses to investigate potential moderators of the intervention effect; and
7. to provide descriptive data on serious adverse events.

Details of the quantitative mediation analyses will be described in a separate mediation analysis plan.

6.3 Analysis methods

All continuous outcomes will be reported descriptively (mean, SD) at all follow-up timepoints. Inferential analyses The primary analyses for the primary and continuous secondary outcomes will use linear regression models with adjustment for baseline score, age (as a continuous variable rather than the dichotomised minimisation variable), country, and any other participant characteristics observed at baseline to be unbalanced across treatment arms (meeting the criteria set out in Section 5.5). The primary follow-up time will be at 3 months; analysis of 12-month follow-up data will also be performed using the same approach. The primary analyses will be based on an intention to treat (ITT) approach (i.e. all participants will be included in the analyses according to their randomised allocation) and will use observed data only.

As a secondary analysis, we will perform repeated measures analyses (using a mixed effects linear regression model with a random effect on participant) for primary and secondary continuous outcomes, including data from participants with observed data for at least one of the three follow-up timepoints. Analyses will be performed to compare both active interventions with usual care, and to compare the personalised active intervention with the standard active intervention. A fixed effect interaction between timepoint and trial arm will be used to evaluate differential treatment effects across timepoints. Adjustments for baseline covariates will be made as for the primary analysis regression models; baseline score will be included as a covariate. Binary outcome measures will be analysed using logistic regression models with adjustments for covariates as above.

The primary analyses will use only the data collected during the time window. The amount of such data will be determined, with the possibility of performing a sensitivity analysis should the amount of data collected outside the time window be substantive. These analyses will be performed using observed data for the ITT population (i.e. according to initial random allocation). No formal adjustments of p-values for multiple testing will be performed; the results for the primary outcome will be interpreted first and the results of the secondary outcomes will then be interpreted in the light of multiple testing. The results of all descriptive and inferential analyses for the primary and secondary continuous outcomes are set out in Tables A.4–A.10.

6.4 Interactions between treatment and participant characteristics

To investigate differential treatment effects across subgroups of participants, we will perform a series of models, for the primary outcome only at 3-and 12-month follow-up, including an interaction term between treatment arm and the participant characteristic (as well as the other covariates to be adjusted for; each model will include one interaction term only). The participant characteristics to be investigated for differential treatment effects are age (continuous and dichotomised as for the minimisation algorithm), country and gender. The interaction terms will be reported as coefficients and 95% confidence intervals, with a global p-value (Table A.11). It is acknowledged that these analyses will have limited statistical power, and hence are exploratory in nature, and the results should be viewed with caution.

6.5 Missing data

Demographic characteristics of participants who do not provide data for the primary outcome measure of their trial at 3- and 12-month follow-up will be set out (including participants who have

formally withdrawn (but allowed their previously collected data to be used), and participants not formally withdrawn but did not complete the primary outcome at follow-up) (Tables A.12 and A.13).

Multiple imputation using chained equations (MICE) will be used to impute primary and secondary continuous outcomes. Predictive mean matching will be the method for imputing individual scores; the number of imputed datasets will be determined by the percentage of participants in each trial that have missing 3-month outcome data (i.e. WEMWBS for PREVENT, PHQ-9 for PROMOTE). Imputation models will be informed by treatment arm, baseline scores, other covariates to be included in the model (i.e. minimisation covariates), and other baseline characteristics found to predict outcome or propensity for missingness (logistic regression models will be used to investigate the associations between baseline characteristics and missingness). Regression models using the imputed datasets will be performed using the same methods as described in Section 6.3 for the ITT population (Tables A.4–A.10). Analyses using imputed data will be considered as a secondary analysis.

6.6 Differential adherence analysis

Differential adherence within the active treatment arms is anticipated, which could impact the outcomes. Ideally, we wish to compare participants who adhere to their intervention with those in the comparator arm who would also have adhered to the treatment had they been allocated to it. To address this issue we will undertake a complier average causal effect (CACE) model, using a 2-stage least squares instrumental variable regression model. Such models will be performed for the continuous primary and secondary outcomes at 3- and 12-month follow-up using observed data only, and will include the covariates adjusted for in the linear regression models (Tables A.4–A.10). The CACE analyses will be considered as secondary analyses. The definition of a 'complier' will be: (i) complete at least one Challenge (some psychoeducation) and at least four Challenges completed or ten Tools used (to capture practice) OR (ii) (to capture combined usage) at least ten Challenges or Tools completed (any combination), including at least one Challenge and one Tool. (The definition of a 'complier' may be subject to further revision.)

6.7 Addressing influence of COVID-19

This trial was designed prior to the onset of the COVID-19 pandemic; hence, no account had been taken in the data collection of data analysis to address the potential effects of COVID-19, and the various restrictions and lockdowns that have been applied in all four countries that are participating in the trial, at various times during the period of participant recruitment and follow-up. To address potential effects of COVID-19 and lockdown, data will be collected to indicate whether the participant was being affected by lockdown/restrictions at the time. This data will be used in sensitivity analyses as an additional predictor (in addition to baseline scores and minimisation characteristics) for all primary and secondary outcomes.

6.8 Addressing protocol violations during the recruitment procedure

As participants were being recruited via the website, it was noted that some people who were interested in joining the trial were being found ineligible, and therefore unable to participate; in some cases, the potential participant was repeating the application process using the same email address but entering different data in order to ensure eligibility. This process is designated as a

protocol violation. Such participants are allowed to join the trial, but in the light of the fact that these participants may in fact be ineligible, we will monitor the numbers of such participants, and will consider a sensitivity analysis excluding these participants, depending on the numbers.

6.9 Addressing protocol violations due to app outage

Due to a fire at the Monsenso site on 10/03/2021, the app was unavailable between 10/03/2021 and 25/03/2021. Therefore, participants who registered for the trial and were randomised to receiving a version of the app did not have the app available for a period of time. It is the intention that the app intervention is available to participants for the full 12-month period of follow-up. However, it is anticipated that the majority of therapeutic effect will take place during the first 4 weeks of the intervention period.

Based on a total of 2141 participants who had registered prior to the app outage:

1. Those participants due both 1m and 3m FU (i.e., either completed or past the timepoint to complete) before 10 March (i.e., only 12m FU potentially impacted by lack of app availability), N = *446*
2. Post 1m but prior to 3m FU before 10 March (i.e., potential impact of app outage on 3m and 12m FU). *Those who had not reached the date their 3-month follow-up was due by 9th March but had passed the date on which their 1 month follow up was due (1 month follow up was due on or before 9th February 2021) N= 465*
3. Those signed up to study but not completed 1m FU before 10 March (this will roughly correspond with those who signed up between 12 Feb and 10 March - the participants most severely affected by the app outage – potentially impact of outage on 1m, 3m and 12m FU), N = 992. *This group can be further divided:*

    - *Never logged on: N= 319 (all have been re-invited twice) – on average before the app outage, 20% of participants never signed into the app*
    - *Has old MSS account, N= 83 (old account not lost, badge data will be re-added)*
    - *Had old MSS account and account lost-AND have not set up themselves up on a new account, N= 389*
    - *Data lost, new account set up and ppt starting again (what badge data we have will be added back) N=201*

4. Signed up after 25 March - unaffected by app outage as there was a new server up and running when they signed up. N=238. *These ppts are affected by the intermittent app stability issues, but not the fire or badge data not showing correctly*

Several analytical approaches to address this issue are available.

All approaches are based on replenishment of the sample: we will recruit an additional number of participants to replace those affected by the outage. A minimum of 992 (Group 3 above) and a maximum of 1457 (Groups 2 and 3) additional participants should be recruited. Groups 1 and 4 do not need to be replaced. The overall rate of non-logins is 20%; hence it is assumed that 20% of participants recruited for replenishment would also be non-logins.

Therefore, the overall revised target recruitment is 2142 (initial sample size) plus an additional 1457 participants to replenish those disrupted by the app outage: 3599 participants in total.

Assuming adequate replenishment of the sample, a series of post hoc sensitivity analyses will be performed on the primary outcomes (WEMWBS for PROMOTE; PHQ-9 for PREVENT).

3-month analyses (observed outcome data only)

The primary analysis using the intention to treat principle (all recruited participants will be analysed according to their randomised group) will be carried out using all observed data, combining data from the original sample with data from the additional participants recruited to replenish the sample following the disruption due to app outage.

For comparison with the primary analysis, two further post hoc sensitivity analyses are proposed.

1. Exclusion of participants whose intervention was disrupted prior to 3-month follow-up (Groups 2 and 3).
2. Primary analysis with addition of a covariate to indicate postulated level of disruption due to app outage (0 for Groups 1 and 4 (little/no disruption); 1 for Group 2 (moderate disruption); 2 for Group 3 (severe disruption)). Although this is a post-randomisation covariate, it is thought to be distributed randomly across treatment groups (i.e. all treatment groups were affected equally by the app outage). As an extension to this model, an interaction term, to model the interaction between treatment group and level of disruption, will be added to this model, with the aim of exploring whether there is any evidence to indicate any variation in the effect of disruption across the three treatment groups.

Comparison of these analyses will allow estimation of the between treatment group differences at specified levels of disruption (i.e. due to disruption occurring at different times during treatment delivery and follow-up). All of the above are post hoc exploratory analyses only, performed in response to the unforeseen app outage. Analysis 2 in particular, should be viewed with caution as it includes a post-randomisation covariate. However, disruption due to app outage is distributed evenly across treatment groups, and it is reasonable to assume that distribution is also effectively random (i.e. timing of recruitment is thought to be random; there is no difference between participants recruited at such a time as to be affected by the app outage and those that were not).

12-month analyses

Following the analysis of the 3-month outcome data, a decision will be taken on whether to include a similar set of sensitivity analyses as were undertaken using the 3-month follow-up data. Whilst it is recognised that there are issues with regard to planning subsequent analyses after viewing the results of analyses of data from a preceding timepoint, in this circumstance, it is helpful to use 3-month data to inform sensitivity analyses at 12 months. If the 3-month analyses indicate little difference comparing analyses including all participants with those including non-disrupted participants only, then it is likely that sensitivity analyses at 12 months would produce similar

results, as it is likely that the disruption due to app outage would have a stronger influence on treatment effect at 3 months than at 12 months.

Options for sensitivity analyses at 12-month follow-up are similar to those for 3-month follow-up. The primary analysis will again be based on the intention to treat principle: all recruited participants will be analysed according to their randomised group. This is the originally intended primary analysis and adds the participants recruited to replenish the sample following the disruption due to app outage to the original intended sample.

For comparison with the primary analysis, three further post hoc sensitivity analyses would be proposed.

1. Exclusion of participants whose intervention was disrupted prior to 12-month follow-up (Groups 1, 2 and 3).
2. Exclusion of participants whose intervention was disrupted prior to 3-month follow-up (Groups 2 and 3; moderate/severe disruption).
3. Primary analysis with addition of a covariate to indicate level of disruption at the specified timepoint (0 for Groups 1 and 4; 1 for Group 2; 2 for Group 3). As an extension to this model, an interaction term, to model the interaction between treatment group and level of disruption, will be added to this model, with the aim of exploring whether there is any evidence to indicate any variation in the effect of disruption across the three treatment groups.

Comparison of these analyses with the primary analysis of 12-month data will allow exploration of the effects of disruption due to the app outage, including the effects of app outage on participants who experienced disruption at different times during treatment delivery and follow-up.

Repeated measures analyses (including observed outcome data only at 1, 3 and 12 months)

Similarly to the analyses of 12-month data, we will await results of the sensitivity analyses at 3-month follow-up prior to making any decisions regarding sensitivity analyses of the repeated measures analyses including data from 1-, 3- and 12-month follow-up.

Again, the primary analysis will use the intention to treat principle : all recruited participants will be analysed according to their randomised group. This is the originally intended primary analysis and adds the participants recruited to replenish the sample following the disruption due to app outage to the original intended sample.

Options for sensitivity analyses for repeated measures analyses include the following.

1. Exclusion of participants whose intervention was disrupted prior to 12-month follow-up (Groups 1, 2 and 3).
2. Exclusion of participants whose intervention was disrupted prior to 3-month follow-up (Groups 2 and 3; moderate/severe disruption).

CACE analyses

CACE analyses will include only data from those participants who did not experience disruption prior to 3-month follow-up, as it will be impossible to distinguish between participants who chose not to adhere to the intervention and those that were unable to adhere due to the app outage; this inability to distinguish between causes of non-adherence could lead to bias in the analysis.

Missing data

The primary analysis (including all randomised participants according to their allocation using observed data only) will be repeated using observed and imputed data for the full sample (i.e. the original sample and replenished sample). See Section 6.5 above for further details on methods to address missing data.

6.10 Adverse events

The main anticipated adverse event in this trial is deliberate self-harm. Serious adverse events will be reported descriptively by treatment arm, also the number of participants in each treatment arm who have experienced at least one serious adverse event will be reported. Inferential analyses will also be performed using the count of events (Poisson or negative binomial model) and the number of participants experiencing the outcome at least once during the 12-month follow-up period (logistic regression), using observed data on an ITT principle with adjustment for covariates as for the outcome data analyses (Table A.14).

6.11 Statistical software

All statistical analyses will be performed using Stata v.14 (or later).

7. General analysis considerations

7.1 Timing of analyses

The first set of statistical analyses will be performed when the 3-month outcome data is available and the database has been locked. At this point, the statistician will perform the 3-month inferential analyses (as well as descriptive analyses for 1- and 3-month data), whilst being unaware of treatment arm allocation. These analyses will be performed on the ITT population using observed data only. Following presentation of the results with anonymised allocations, the allocations will be revealed for interpretation of the results. Following revelation of treatment allocations, the remaining analyses of 3-month outcome data will be performed. Following completion of 12-month data collection and database lock, the remaining analyses of the 12-month outcome data will be performed. At this point the repeated measures analyses and analyses including imputed and observed data will be performed.

7.2 Related documents

The related documents that should be read in conjunction with this SAP are set out below.

1. Trial protocol version 1.3
2. Trial master file (TMF)
3. Trial selection, randomisation, selection of emotional competence intervention components document
4. Data Management Plan
5. Process evaluation analysis plan
6. SOP 019 Deviations, Misconduct and Serious Breaches of GCP and(or) the Protocol

7.3 Mediation analyses

Quantitative mediational analyses will be described in the process evaluation analysis plan.

8. Health economic analysis plan

8.1 Aim of economic evaluation

Both trials seek to evaluate mental health and emotional wellbeing within their respective participant samples comparing two active interventions (one personalised, one uniform across participants) against usual care (self-monitoring of emotional well-being via an app) and against each other. The aim of the health economic evaluation is to estimate costs and outcomes associated with the interventions.

8.2 Objective of economic evaluation

To undertake within country cost consequence analyses

8.3 Overview of economic analysis

The economic analyses will use a cost-consequence approach. Because all the resources used, costs, and outcomes are transparently listed in the cost-consequence analysis, decision makers can select the information that is of most interest to them. Furthermore, the choice of cost-consequence analysis supports current approaches used by healthcare decision-makers to value the efficiency of healthcare interventions across Europe.[6]

8.4 Jurisdiction

A cost consequence analysis will be present for each country.

8.5 Perspective

The analyses will take a societal perspective.

8.6 Time horizon

12 months.

8.7 Identification of resources

Health and social care utilisation will be identified and collected using an adapted version of the Adult Service Use Schedule (AD-SUS). Developed for mental health trials the AD-SUS questionnaire quantifies the use of healthcare resources, use of medication and employment and time off work over the trial including follow up.[7] Adaptation will include time away from school or college together with changes to the wording to ensure relevance across the four countries.

8.8 Measurement of resource use data

The AD-SUS will be administered at baseline to assess use of services in previous 6-months, at 3-month follow-up, ask about use of services since the baseline interview, and 12-month follow-up, ask about use of services since the 3-month interview.

8.9 Valuation of resource use data

Unit costs of health and social care will be taken from appropriate national publications to reflect differences in costs between countries (for example, Curtis & Burns (2018)).[8] Productivity losses will be valued using the human capital approach.[9] Wage rates for each country will be derived from European sources such as Eurostat.[10]

8.10 Identification of outcomes

The analysis will use the EQ-5D-3L.[11]

8.11 Measurement of outcomes

The EQ-5D-3L will be administered with the adapted ADSUS at baseline, 3 and 12-months

8.12 Valuation of outcomes

Responses to the EQ-5D-3L will be converted to utility values using the EuroQoL general population tariff values for each country.[12–15]

8.13 Analysis population

The analyses will be on the intention to treat (ITT) basis (i.e. participants will be included in the analyses according to their randomised allocation) and will include observed data only).

8.14 Timing of analyses

Analyses will be undertaken following 12-month data collection and associated database lock.

8.15 Discount rates for costs and benefits

Given the time horizon no discounting is required.

8.16 Analysis of resource use, costs and outcomes

Within country analyses of the differences between the trial arms will be undertaken. Although the distribution of costs is commonly skewed in populations of this kind, analyses will compare mean costs between groups using standard parametric regression models adjusted for minimisation variables and baseline costs. The robustness of the parametric tests will be confirmed using bias-corrected, non-parametric bootstrapping.[16,17] As with primary and secondary outcomes, between group differences in costs and EQ-5D-3L will be presented as means and 95% CIs using STATA v14.2.

Given the differences in cost structures (health and social care, and wage rates) and differences in the population tariffs between countries, no formal analyses of differences between costs and utility values between countries will be undertaken.[18] However, the differences between the number and type of resource use and days away from work and employment will be examined.

9. References

1. ICH E9 Expert Working Group. ICH Harmonised Tripartite Guideline: Statistical principles for clinical trials. Stat Med. 1999;18(15):1905-42.

2. Gamble C, Krishan A, Stocken D, et al. Guidelines for the content of statistical analysis plans in clinical trials. JAMA. 2017;318(23):2337-43.

3. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet. 2001. 14;357:1191-4.

4. Boutron I, Altman DG, Moher D, Schulz KF, Ravaud P, for the CONSORT NPT Group. CONSORT statement for randomized trials of nonpharmacologic treatments: A 2017 update and a CONSORT extension for nonpharmacologic trial abstracts. Ann Intern Med. Ann Intern Med. 2017;167(1):40-47.

5. Juszczak E, Altman DG, Hopewell S, Schulz K. Reporting of multi-arm parallel group randomized trials: extension of the CONSORT 2010 statement. JAMA. 2019;321(16):1610-1620.

6. Angelis A, Lange A, Kanavos P. Using health technology assessment to assess the value of new medicines: the results of a systematic review and expert consultation across eight European countries. European Journal of Health Economics 2018; 19(1): 123-152. doi:10.1007/s10198-017-0871-0.

7. Kuyken W, Hayes R, Barrett B, et al. The effectiveness and cost-effectiveness of mindfulness-based cognitive therapy compared with maintenance antidepressant treatment in the prevention of depressive relapse/recurrence: results of a randomised controlled trial (the PREVENT study). Health Technology Assessment, No. 19.73. Southampton (UK): NIHR Journals Library; 2015 Sep.

8. Curtis L, Burns A. Unit Costs of Health & Social Care 2018. Kent: Personal Social Services Research Unit. Available at: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2018/ [07 June 2019].

9. Koopmanschap MA, Rutten FFH. Pharmacoeconomics (1996) 10: 460. https://doi.org/10.2165/00019053-199610050-00003.

10.Wage and labour costs. Eurostat. Available at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Wages_and_labour_costs [07 June 2019].

11. Brooks R. EuroQol: the current state of play. Health Policy. 1996; 37(1)53-72.

12. Dolan P. Modeling valuations for EuroQol health states. Med Care. 1997; 35(11): 1095-108.

13. Cleemput I. A social preference valuations set for EQ-5D health states in Flanders, Belgium. Eur J Health Econ. 2010;11(2):205-13.

14. Badia X, Roset R, Herdman, M, Kind P.A comparison of United Kingdom and Spanish general population time trade-off values for EQ-5D health states. Med Decis Making 2001; 21(1): 7-16.

15. Claes C, Greiner W, Uber A, Graf von der Schulenburg JM. An interview-based comparison of the TTO and VAS values given to EuroQol states of health by the general German population.In: Greiner W, J-M. Graf v.d. Schulenburg, Piercy J, editors. EuroQol Plenary Meeting, 1-2 October 1998. Discussion papers. Centre for Health Economics and Health Systems Research, University of Hannover, Germany. Uni-Verlag Witte, 1999; 13-39.

16. Efron, B. and Tibshirani, R.J. An Introduction to the Bootstrap (1993). Chapman and Hall, New York.

17. Barber JA, Thompson SG. Analysis of cost data in randomized trials: an application of the non-parametric bootstrap. Statistics in Medicine, 2000; 9(23):3219-36.

18. Kiadaliri AA, Eliasson B, Gerdtham UG. 2015. Does the choice of EQ-5D tariff matter? A comparison of the Swedish EQ-5D-3L index score with UK, US, Germany and Denmark among type 2 diabetes patients. Health and Quality of Life Outcomes201513:145. Available at: https://doi.org/10.1186/s12955-015-0344-z [07 June 2019].

10. Abbreviations and definitions

Abbreviation	Meaning	Detail
ADSUS	Adult Service Use Schedule	
CACE	Complier Average Causal Effect	
CI	Confidence interval	
cmRCT	cohort multiple Randomised Controlled Trial	
DMEC	Data monitoring and ethics committee	
EQ5D-3L	European Quality of Life 5 Dimensions-3 Levels	
ExeCTU	Exeter Clinical Trials Unit	
GAD-7	Generalised Anxiety Disorder-7	
ITT	Intention to treat	
MINI-IPIP-N		
PHQ-9	Patient Health Questionnaire-9	Primary outcome for PREVENT; secondary outcome for PROMOTE
SAP	Statistical Analysis Plan	

SD	Standard deviation	
SOP	Standard Operating Procedures	SOPs document a series of steps to be followed to accomplish a process. They include the purpose and scope of a process, describes procedure (i.e. what, where, and when), assigs responsibility for action and decisions (i.e. who), identify records to be kept, and identify associated and reference documents (as applicable)
TMF	Trial Master File	
TMG	Trial Management Group	
TSC	Trial Steering Committee	
WASS	Work and Social Adjustment Scale	
WEMWBS	Warwick-Edinburgh Mental Well Being Scale	Primary outcome for PROMOTE; secondary outcome for PREVENT

11. Appendices

Table A.1 Participant demographic characteristics at baseline

Characteristic	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**
Age (years); mean (SD), median [25th centile; 75th centile]			
Age (years)			
16–17			
18–22			
Gender			
Male			
Female			
Both			
Neither			
Country			
UK			
Germany			
Spain			
Belgium			
Ethnic group (tbc)			
Education (tbc)			
Occupation (tbc)			
Post-randomisation			
App outage disruption			
Disruption prior 1-month follow-up			
Disruption prior to 3-month follow-up			
Disruption prior to 12-month follow-up			
No disruption			

Table A.2 Adherence measures for active intervention arms

	Personalised digital self-help + usual care (N=X)	**Non-personalised digital self-help +usual care (N=X)**
Number of usages of the app		
Percentage of relevant self-help package completed		
Number of challenge tasks completed		
Number of times the specified tool is practiced (total)		
Number of days used (total)		
Longest streak of consecutive use (days)		
Number of days that daily mood rating is completed (maximum once per day)		
Number of times that ecological momentary ratings are completed		
Overall time spent on the app (hours)		

Reported as mean (SD), n; median [min, max].

Table A.3 Component allocation (Personalised digital self-help plus usual care arm only)

Individual components **n (%)**	**Personalised digital self-help + usual care (N=X)**
Achievement appraisal	
Social appraisal	
Rumination	
Emotion knowledge and perception	
Component combinations	
Achievement appraisal/Social appraisal	
Achievement appraisal/Rumination	
Achievement appraisal/Emotion knowledge and perception	
Social appraisal/Rumination	
Social appraisal/Emotion knowledge and perception	
Rumination/Emotion knowledge and perception	

Table A.4 WEMWBS at baseline, 1-, 3- and 12-month follow-up

Descriptive statistics: mean (SD), n												
	Baseline			1-month follow-up			3-month follow-up			12-month follow-up		
	Personalised digital self-help + usual care (N=X)	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**
ITT, observed data only												
Inferential analyses: between group mean difference (95% confidence interval)												
	3-month follow-up				**12-month follow-up**							
	Personalised digital self-help + usual care vs. Usual care	**Non-personalised digital self-help vs. Usual care**	**Personalised digital self-help + usual care vs Non-personalised digital self-help**	**Global p-value**	**Personalised digital self-help + usual care vs. Usual care**	**Non-personalised digital self-help vs. Usual care**	**Personalised digital self-help + usual care vs Non-personalised digital self-help**	**Global p-value**				
ITT, observed data only												
CACE, observed data only												

ITT, observed and imputed data								
Repeated measures analyses (interaction between intervention arm (usual care as reference) and timepoint (1-month follow-up as reference)) Interaction coefficient (95% confidence interval)								
	3-month follow-up			**12-month follow-up**				
	Personalised digital self-help + usual care	**Non-personalised digital self-help +usual care**		**Personalised digital self-help + usual care**	**Non-personalised digital self-help +usual care**			
ITT observed data only								
	Global p-value							

ITT: Intention to treat; CACE: Complier Average Causal Effect.

Table A.5 PHQ-9 at baseline, 1-, 3- and 12-month follow-up

Descriptive statistics: mean (SD), n												
	Baseline			1-month follow-up			3-month follow-up			12-month follow-up		
	Personalised digital self-help + usual care (N=X)	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**
ITT, observed data only												
Inferential analyses: between group mean difference (95% confidence interval)												
	3-month follow-up				**12-month follow-up**							
	Personalised digital self-help + usual care vs. Usual care	**Non-personalised digital self-help vs. Usual care**	**Personalised digital self-help + usual care vs Non-personalised digital self-help**	**Global p-value**	**Personalised digital self-help + usual care vs. Usual care**	**Non-personalised digital self-help vs. Usual care**	**Personalised digital self-help + usual care vs Non-personalised digital self-help**	**Global p-value**				
ITT, observed data only												
CACE, observed data only												

ITT, observed and imputed data								
Repeated measures analyses (interaction between intervention arm (usual care as reference) and timepoint (1-month follow-up as reference)) Interaction coefficient (95% confidence interval)								
	3-month follow-up			**12-month follow-up**				
	Personalised digital self-help + usual care	**Non-personalised digital self-help +usual care**		**Personalised digital self-help + usual care**	**Non-personalised digital self-help +usual care**			
ITT observed data only								
	Global p-value							

ITT: Intention to treat; CACE: Complier Average Causal Effect.

Table A.6 WSAS at baseline, 1-, 3- and 12-month follow-up

Descriptive statistics: mean (SD), n												
	Baseline			1-month follow-up			3-month follow-up			12-month follow-up		
	Personalised digital self-help + usual care (N=X)	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**
ITT, observed data only												
Inferential analyses: between group mean difference (95% confidence interval)												
	3-month follow-up				**12-month follow-up**							
	Personalised digital self-help + usual care vs. Usual care	**Non-personalised digital self-help vs. Usual care**	**Personalised digital self-help + usual care vs Non-personalised digital self-help**	**Global p-value**	**Personalised digital self-help + usual care vs. Usual care**	**Non-personalised digital self-help vs. Usual care**	**Personalised digital self-help + usual care vs Non-personalised digital self-help**	**Global p-value**				
ITT, observed data only												
CACE, observed data only												

ITT, observed and imputed data								
Repeated measures analyses (interaction between intervention arm (usual care as reference) and timepoint (1-month follow-up as reference)) Interaction coefficient (95% confidence interval)								
	3-month follow-up			**12-month follow-up**				
	Personalised digital self-help + usual care	**Non-personalised digital self-help +usual care**		**Personalised digital self-help + usual care**	**Non-personalised digital self-help +usual care**			
ITT observed data only								
	Global p-value							

ITT: Intention to treat; CACE: Complier Average Causal Effect.

Table A.7 GAD-7 at baseline, 1-, 3- and 12-month follow-up

Descriptive statistics: mean (SD), n												
	Baseline			1-month follow-up			3-month follow-up			12-month follow-up		
	Personalised digital self-help + usual care (N=X)	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**
ITT, observed data only												
Inferential analyses: between group mean difference (95% confidence interval)												
	3-month follow-up				**12-month follow-up**							
	Personalised digital self-help + usual care vs. Usual care	**Non-personalised digital self-help vs. Usual care**	**Personalised digital self-help + usual care vs Non-personalised digital self-help**	**Global p-value**	**Personalised digital self-help + usual care vs. Usual care**	**Non-personalised digital self-help vs. Usual care**	**Personalised digital self-help + usual care vs Non-personalised digital self-help**	**Global p-value**				
ITT, observed data only												
CACE, observed data only												

ITT, observed and imputed data								
Repeated measures analyses (interaction between intervention arm (usual care as reference) and timepoint (1-month follow-up as reference)) Interaction coefficient (95% confidence interval)								
	3-month follow-up			**12-month follow-up**				
	Personalised digital self-help + usual care	**Non-personalised digital self-help +usual care**		**Personalised digital self-help + usual care**	**Non-personalised digital self-help +usual care**			
ITT observed data only								
	Global p-value							

ITT: Intention to treat; CACE: Complier Average Causal Effect.

Table A.8 EQ5D-3L at baseline, 1-, 3- and 12-month follow-up

Descriptive statistics: mean (SD), n												
	Baseline			1-month follow-up			3-month follow-up			12-month follow-up		
	Personalised digital self-help + usual care (N=X)	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**
ITT, observed data only												
Inferential analyses: between group mean difference (95% confidence interval)												
	3-month follow-up				**12-month follow-up**							
	Personalised digital self-help + usual care vs. Usual care	**Non-personalised digital self-help vs. Usual care**	**Personalised digital self-help + usual care vs Non-personalised digital self-help**	**Global p-value**	**Personalised digital self-help + usual care vs. Usual care**	**Non-personalised digital self-help vs. Usual care**	**Personalised digital self-help + usual care vs Non-personalised digital self-help**	**Global p-value**				
ITT, observed data only												
CACE, observed data only												

ITT, observed and imputed data								
Repeated measures analyses (interaction between intervention arm (usual care as reference) and timepoint (1-month follow-up as reference)) Interaction coefficient (95% confidence interval)								
	3-month follow-up				**12-month follow-up**			
	Personalised digital self-help + usual care		**Non-personalised digital self-help +usual care**		**Personalised digital self-help + usual care**		**Non-personalised digital self-help +usual care**	
ITT observed data only								
	Global p-value							

ITT: Intention to treat; CACE: Complier Average Causal Effect.

Table A.9 EQ5D Health State at baseline, 1-, 3- and 12-month follow-up

Descriptive statistics: mean (SD), n												
	Baseline			1-month follow-up			3-month follow-up			12-month follow-up		
	Personalised digital self-help + usual care (N=X)	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**
ITT, observed data only												
Inferential analyses: between group mean difference (95% confidence interval)												
	3-month follow-up				**12-month follow-up**							
	Personalised digital self-help + usual care vs. Usual care	**Non-personalised digital self-help vs. Usual care**	**Personalised digital self-help + usual care vs Non-personalised digital self-help**	**Global p-value**	**Personalised digital self-help + usual care vs. Usual care**	**Non-personalised digital self-help vs. Usual care**	**Personalised digital self-help + usual care vs Non-personalised digital self-help**	**Global p-value**				
ITT, observed data only												
CACE, observed data only												

ITT, observed and imputed data								
Repeated measures analyses (interaction between intervention arm (usual care as reference) and timepoint (1-month follow-up as reference)) Interaction coefficient (95% confidence interval)								
	3-month follow-up				**12-month follow-up**			
	Personalised digital self-help + usual care		**Non-personalised digital self-help +usual care**		**Personalised digital self-help + usual care**		**Non-personalised digital self-help +usual care**	
ITT observed data only								
	Global p-value							

ITT: Intention to treat; CACE: Complier Average Causal Effect.

Table A.10 MINI- IPIP-N at baseline, 1-, 3- and 12-month follow-up

Descriptive statistics: mean (SD), n												
	Baseline			1-month follow-up			3-month follow-up			12-month follow-up		
	Personalised digital self-help + usual care (N=X)	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**
ITT, observed data only												
Inferential analyses: between group mean difference (95% confidence interval)												
	3-month follow-up				**12-month follow-up**							
	Personalised digital self-help + usual care vs. Usual care	**Non-personalised digital self-help vs. Usual care**	**Personalised digital self-help + usual care vs Non-personalised digital self-help**	**Global p-value**	**Personalised digital self-help + usual care vs. Usual care**	**Non-personalised digital self-help vs. Usual care**	**Personalised digital self-help + usual care vs Non-personalised digital self-help**	**Global p-value**				
ITT, observed data only												
CACE, observed data only												

ITT, observed and imputed data								
Repeated measures analyses (interaction between intervention arm (usual care as reference) and timepoint (1-month follow-up as reference)) Interaction coefficient (95% confidence interval)								
	3-month follow-up			**12-month follow-up**				
	Personalised digital self-help + usual care	**Non-personalised digital self-help +usual care**		**Personalised digital self-help + usual care**	**Non-personalised digital self-help +usual care**			
ITT observed data only								
	Global p-value							

ITT: Intention to treat; CACE: Complier Average Causal Effect.

Table A.11 Differential treatment effects across participant characteristics

Covariate	**Personalised digital self-help + usual care**	**Non-personalised digital self-help +usual care**	
	Interaction effect; mean (95% CI)	**Interaction effect; mean (95% CI)**	**Global p-value**
3-month follow-up			
Age (continuous)			
Age (18–22 years; 16–17 years reference)			
Gender (Female reference)			
Male			
Both			
Neither			
Country (UK reference)			
Germany			
Spain			
Belgium			
12-month follow-up			
Age (continuous)			
Age (18–22 years; 16–17 years reference)			
Gender (Female reference)			
Male			
Both			
Neither			
Country (UK reference)			
Germany			
Spain			
Belgium			

Table A.12 Participant demographic characteristics at baseline: participants who did not complete primary outcome data at 3-month follow-up

Characteristic	Personalised digital self-help + usual care (N=X)	Non-personalised digital self-help +usual care (N=X)	Usual care (N=x)
Age (years); mean (SD), median [25th centile; 75th centile]			
Age (years)			
16–17			
18–22			
Gender			
Male			
Female			
Both			
Neither			
Country			
UK			
Germany			
Spain			
Belgium			
Ethnic group (tbc)			
Education (tbc)			
Occupation (tbc)			

Table A.13 Participant demographic characteristics at baseline: participants who did not complete primary outcome data at 12-month follow-up

Characteristic	Personalised digital self-help + usual care (N=X)	Non-personalised digital self-help +usual care (N=X)	Usual care (N=x)
Age (years); mean (SD), median [25th centile; 75th centile]			
Age (years)			
16–17			
18–22			
Gender			
Male			
Female			
Both			
Neither			
Country			
UK			
Germany			
Spain			
Belgium			
Ethnic group (tbc)			
Education (tbc)			
Occupation (tbc)			

Table A.14 Adverse events during 12-month follow-up

	Descriptive statistics: mean (SD), n; median [min, max] or n/N (%)			Inferential statistics: IRR (95% CI) or OR (95% CI)			Global p-value
Adverse outcome (to be coded)	**Personalised digital self-help + usual care (N=X)**	**Non-personalised digital self-help +usual care (N=X)**	**Usual care (N=x)**	**Non-personalised digital self-help +usual care**	**Personalised digital self-help + usual care**	**Non-personalised digital self-help +usual care**	
Event (count)							
Number of participants experiencing event at least once during 12-month follow-up; n/N (%)							

IRR: incidence rate ratio; OR: odds ratio

Table A.15 Characteristics of people who log on to the website but decline participation

Characteristic	**Personalised digital self-help + usual care (N=X)**
Age (years); mean (SD), median [25th centile; 75th centile]	
Age (years)	
16–17	
18–22	
Gender	
Male	
Female	
Both	
Neither	
Country	
UK	
Germany	
Spain	
Belgium	
Ethnic group (tbc)	