

Cognitive behavioral therapy skills via a smartphone app for subthreshold depression among adults in the community: the RESiLIENT randomized controlled trial

In the format provided by the
authors and unedited

Supplements to:

Cognitive behavioural therapy skills using a smartphone app for subthreshold depression among adults in the community: the RESiLIENT randomised controlled trial

Table of Contents

1. Protocol
2. Statistical Analysis Plan (SAP)
3. Supplementary Data
4. CONSORT 2010 Checklist for nonpharmacological and factorial trials

1. Protocol

Furukawa TA, Tajika A, Sakata M, Luo Y, Toyomoto R, Horikoshi M, Akechi T, Kawakami N, Nakayama T, Kondo N, Fukuma S, Noma H, Christensen H, Kessler RC, Cuijpers P & Wason JMS (2023) Four 2x2 factorial trials of smartphone CBT to reduce subthreshold depression and to prevent new depressive episodes among adults in the community-RESiLIENT trial (Resilience Enhancement with Smartphone in Lliving ENvironmenTs): a master protocol. **BMJ Open**, 13, e067850. 10.1136/bmjopen-2022-067850

(Submitted 30 August 2022)

(Accepted 8 February 2023)

BMJ Open Four 2×2 factorial trials of smartphone CBT to reduce subthreshold depression and to prevent new depressive episodes among adults in the community—RESiLIENT trial (Resilience Enhancement with Smartphone in Living ENvironmentTs): a master protocol

Toshi A Furukawa ¹, Aran Tajika,¹ Masatsugu Sakata,¹ Yan Luo,¹ Rie Toyomoto,¹ Masaru Horikoshi,^{2,3} Tatsuo Akechi,⁴ Norito Kawakami ⁵, Takeo Nakayama,⁶ Naoki Kondo ^{7,8}, Shingo Fukuma ⁹, Hisashi Noma ¹⁰, Helen Christensen,¹¹ Ronald C Kessler,¹² Pim Cuijpers ¹³, James M S Wason ¹⁴

To cite: Furukawa TA, Tajika A, Sakata M, *et al.* Four 2×2 factorial trials of smartphone CBT to reduce subthreshold depression and to prevent new depressive episodes among adults in the community—RESiLIENT trial (Resilience Enhancement with Smartphone in Living ENvironmentTs): a master protocol. *BMJ Open* 2023;13:e067850. doi:10.1136/bmjopen-2022-067850

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2022-067850>).

Received 30 August 2022
Accepted 08 February 2023



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Professor Toshi A Furukawa;
furukawa@kuhp.kyoto-u.ac.jp

ABSTRACT

Introduction The health burden due to depression is ever increasing in the world. Prevention is a key to reducing this burden. Guided internet cognitive-behavioural therapies (iCBT) appear promising but there is room for improvement because we do not yet know which of various iCBT skills are more efficacious than others, and for whom. In addition, there has been no platform for iCBT that can accommodate ongoing evolution of internet technologies.

Methods and analysis Based on our decade-long experiences in developing smartphone CBT apps and examining them in randomised controlled trials, we have developed the Resilience Training App Version 2. This app now covers five CBT skills: cognitive restructuring, behavioural activation, problem-solving, assertion training and behaviour therapy for insomnia. The current study is designed as a master protocol including four 2×2 factorial trials using this app (1) to elucidate specific efficacies of each CBT skill, (2) to identify participants' characteristics that enable matching between skills and individuals, and (3) to allow future inclusion of new skills. We will recruit 3520 participants with subthreshold depression and ca 1700 participants without subthreshold depression, to examine the short-term efficacies of CBT skills to reduce depressive symptoms in the former and to explore the long-term efficacies in preventing depression in the total sample. The primary outcome for the short-term efficacies is the change in depressive symptoms as measured with the Patient Health Questionnaire-9 at week 6, and that for the long-term efficacies is the incidence of major depressive episodes as assessed by the computerised Composite International Diagnostic Interview by week 50.

Ethics and dissemination The trial has been approved by the Ethics Committee of Kyoto University Graduate School of Medicine (C1556).

Trial registration number UMIN000047124.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ This trial aims to examine specific efficacies of various internet cognitive-behavioural therapy (iCBT) skills and to evaluate interactions of these skills with participants' characteristics to enable optimal matching of iCBT skills and individuals.
- ⇒ The trial is designed as a master protocol involving four 2×2 factorial trials with common intervention arms, thus allowing efficient use of the sample.
- ⇒ The trial can be expanded as a platform trial to accommodate new developments in the smartphone CBT app.
- ⇒ The trial uses remote recruitment and informed consent procedures using the internet to allow accelerated recruitment.

INTRODUCTION

In 1984, a report by the US National Institute of Mental Health concluded that prevention of depression is impossible.¹ However, research over the next 30 years has shown that depression prevention is possible² and cost-effective.³ In the meantime, however, health losses due to common mental disorders are continuing to increase and depression remains the largest source of disabilities due to mental health worldwide.^{4 5} Subthreshold depression, that is, a depressive state falling short of meeting the diagnostic threshold for major depression, is more prevalent than major depression itself and leads to major losses in well-being and productivity.⁶ Prevention of depression usually takes the form of

indicated prevention, preventing subthreshold states to develop into full-blown episodes.² To maintain and improve the mental health of the entire population, it is necessary to prevent the transition from a subdepressive state to a clinically depressive level, to reduce depression in those who are currently depressed and to maintain the euthymic state in those who have once improved.⁷

Our recently published systematic reviews have shown that various psychotherapies are effective not only in treating but also in preventing depression, with cognitive-behavioural therapy (CBT)-based interventions having the strongest evidence base in both respects.^{2 8 9} CBT typically includes cognitive and behavioural skill components such as cognitive restructuring (CR), behavioural activation (BA), problem-solving (PS), assertion training (AT), behaviour therapy for insomnia (BI) and relaxation. However, classic face-to-face CBT is not widespread enough to meet the actual clinical, let alone preventive, needs, because it requires a lot of human, financial and time resources. On the other hand, since the 2000s, research on internet CBT (iCBT) using information and communication technology has seen much progress, and it is now well established that guided iCBT is as effective as standard face-to-face treatment and both are superior to usual care in treating clinical depression^{10 11} and subclinical depression.^{11–13} Evidence is also emerging that they may be able to prevent new depression onset among participants with subthreshold depression.^{12 14}

However, there is much room for improvement for these promising interventions. First, various iCBT packages shown to be effective in randomised controlled trials consist of different combinations of CBT skills and we do not yet know if all their components are necessary or if more efficient and less time-consuming combinations may be possible for individuals, both in the treatment and prevention of depression. Second, the internet and smartphone technologies are in constant evolution but no system has yet been developed to accommodate ongoing innovations in digital interventions.

We have been developing a series of smartphone CBT apps in the past decade. Our original smartphone CBT ‘Kokoro App’ (‘Kokoro’ means mind in Japanese) included CR and BA, and its efficacy and safety have been demonstrated in a randomised controlled trial involving 164 drug-refractory depressed patients.¹⁵ We then developed new iCBT modules of PS and AT. Using PS and BA significantly reduced fear of cancer recurrence and depression among 447 breast cancer survivors in a randomised controlled trial.¹⁶ Putting all these components together, we have developed the ‘Resilience Training App’ which targets subclinical distress in the general population, promotes their resilience and well-being, and aims to prevent the future incidence of clinical depression. Using this Resilience Training App Version 1, we recently completed a fully factorial trial to examine the specific efficacy of CR, BA, PS, AT and self-monitoring among 1093 college students with subthreshold depression: after 2 months, the students saw moderate to large

reductions in their depression and anxiety on average and we were unable to detect skill-specific efficacies.¹⁷ However, when we conducted a component network meta-analysis of iCBT packages to treat depression in 76 trials (18178 participants), BA proved the most efficacious, followed by BI, PS and AT.¹⁸

Based on these results, we have now added an iCBT module for BI in Version 2 of the Resilience Training App, and we hereby plan the RESiLIENT Study (Resilience Enhancement with Smartphone in Living ENvironments), a master protocol of four 2×2 factorial randomised trials, to elucidate the following clinical questions:

1. The Resilience Training App Version 2 includes CR, BA, PS, AT and BI. By examining their various combinations in factorial trials, we aim to estimate the specific efficacies of these various CBT skills.
2. We further plan to examine interactions between these CBT skills and individuals’ baseline characteristics to enable matching between interventions and individuals.
3. The trial is designed as a platform trial in which new versions of the existing CBT modules or a completely new module may be later added to allow randomised comparisons.

METHODS

This protocol has been written in accordance with the Standard Protocol Items: Recommendations for Interventional Trials guideline,¹⁹ the Consolidated Standards of Reporting Trials statement for randomised trials of non-pharmacological treatments,²⁰ and the Japanese Ethical Guidelines for Medical and Biological Research Involving Human Subjects (23 March 2021) and its Guidance (16 April 2021). This report is based on the protocol version 1.0.2, approved on 28 July 2022 by Kyoto University Graduate School of Medicine Ethics Committee (approval number C1556). The trial has been registered in UMIN-CTR (UMIN000047124).

Overall design

The RESiLIENT trial is a platform trial involving four 2×2 factorial trials examining the efficacy of five smartphone CBT skills. Analysis 1 will focus on the acute intervention effects at week 6 among participants with the Patient Health Questionnaire-9 (PHQ-9) scores of 5 or more at week 0. Analysis 2 will focus on the long-term prophylactic effects at week 50 among all participants. The trial will be powered for analysis 1 but will remain exploratory for analysis 2. The recruitment started on 1 September 2022 and is scheduled to complete by 31 March 2024, with the last follow-up planned to complete by 30 April 2024.

Participants

Study settings

Participants will be recruited in the following four fields. The potential participants will access the internet web

Table 1 Contents of the interventions and controls

		Chapter 1	Chapter 2	Chapter 3	Chapter 4	Chapter 5	Chapter 6
1	BA	4 basic lessons for BA	2 advanced lessons for BA	1 advanced lesson for BA	1 advanced lesson for BA	BA in everyday life	To sum up and to continue
2	CR	4 basic lessons for CR	2 advanced lessons for CR	1 advanced lesson for CR	1 advanced lesson for CR	CR in everyday life	To sum up and to continue
3	PS	4 basic lessons for PS	2 advanced lessons for PS	2 advanced lessons for PS	1 advanced lesson for PS	PS in everyday life	To sum up and to continue
4	AT	5 basic lessons for AT	2 advanced lessons for AT	1 advanced lesson for AT	1 advanced lesson for AT	AT in everyday life	To sum up and to continue
5	BI	5 basic lessons for BI	1 advanced lesson for BI	1 advanced lesson for BI	1 advanced lesson for BI	BI in everyday life	To sum up and to continue
6	BA+CR	4 basic lessons for BA	4 advanced lessons for BA	4 basic lessons for CR	4 advanced lessons for CR	BA and CR in everyday life	To sum up and to continue
7	BA+PS	4 basic lessons for BA	4 advanced lessons for BA	4 basic lessons for PS	5 advanced lessons for PS	BA and PS in everyday life	To sum up and to continue
8	BA+AT	4 basic lessons for BA	4 advanced lessons for BA	5 basic lessons for AT	4 advanced lessons for AT	BA and AT in everyday life	To sum up and to continue
9	BA+BI	4 basic lessons for BA	4 advanced lessons for BA	5 basic lessons for BI	3 advanced lessons for BI	BA and BI in everyday life	To sum up and to continue
10	Self-check	–	–	–	–	–	To sum up and to continue
11	Health information	Physical exercise	Nutrition	Oral health	–	–	To sum up and to continue
12	Delayed treatment	–	–	–	–	–	To sum up and to continue

All the participants will receive weekly encouragement emails, except for groups 11 and 12. AT, assertion training; BA, behavioural activation; BI, behaviour therapy for insomnia; CR, cognitive restructuring; PS, problem-solving.

page for the trial to apply for the study (the Participant timeline section).

1. Health insurance societies

There are several types of public health insurance schemes together constituting the universal health-care in Japan. We will collaborate with three large associations from them, namely, National Federation of Health Insurance Societies covering employees and their families in large corporations (ca 30 million insured), Japan Health Insurance Association covering employees and their families of small-to-middle-sized corporations (ca 35 million insured), and National Civil Engineering Health Insurance Societies covering civil engineering and construction workers and their families (ca 0.2 million insured).

2. Business companies and corporations
3. Community and local governments
4. Direct-to-consumer advertisements

Eligibility criteria

Inclusion criteria

1. People of any sex, aged 18 years or older at the time of providing informed consent.
2. In possession of their own smartphone (either iOS or Android).
3. Written informed consent for participation in the trial.
4. Completion of all the baseline questionnaires within 1 week after providing informed consent.

Exclusion criteria

1. Cannot read or write Japanese.

2. Receiving treatment from mental health professionals at the time of screening.
3. Screening PHQ-9 total scores of (1) 15 or higher, or (2) between 10 and 14, inclusive, and scoring 2 or 3 on its item 9.^{17 21}

Interventions and controls

The participants will be randomised equally to one of the following nine intervention arms and the three control arms.

Interventions

The intervention arms include:

1. BA
2. CR
3. PS
4. AT
5. BI
6. BA+CR
7. BA+PS
8. BA+AT
9. BA+BI

BA consists of psychoeducation about pleasurable activities according to the principle 'When your body moves, so does your mind'. It provides a worksheet of a personal experiment to test out a new activity and also a gamified 'action marathon' to promote such personal experiments.

CR consists of psychoeducation of the cognitive-behavioural model and cognitive restructuring. The participant learns how to monitor their reactions to

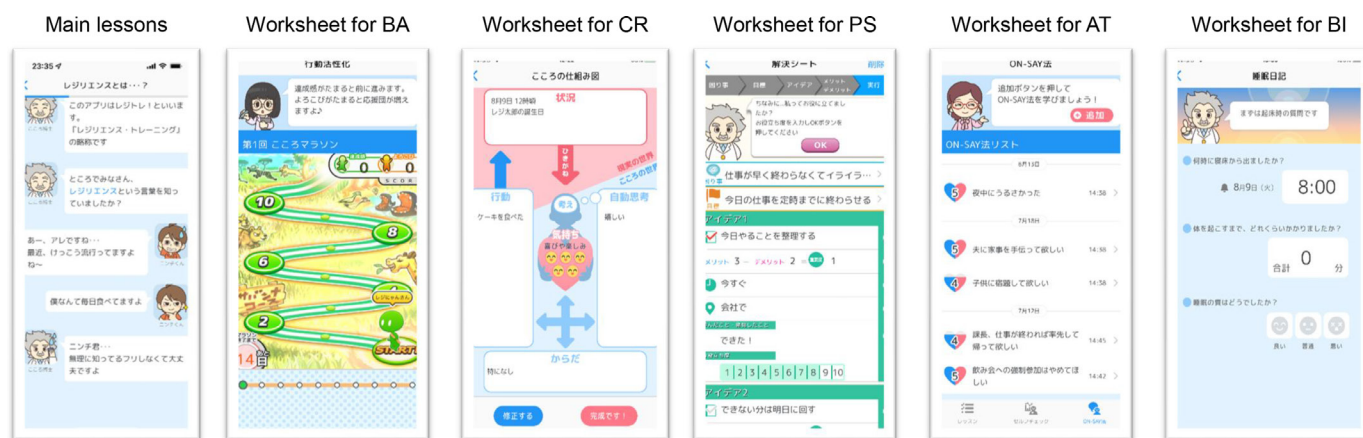


Figure 1 Sample screenshots from the Resilience Training App (copyright by Toshi A Furukawa, Masaru Horikoshi and Tatsuo Akechi). AT, assertion training; BA, behavioural activation; BI, behaviour therapy for insomnia; CR, cognitive restructuring; PS, problem-solving.

situations in terms of feelings, thoughts, body reactions and behaviours by filling in mind maps. The participant uses these mind maps to apply cognitive restructuring and find alternative thoughts. In order to help the participant broaden their thoughts, CR provides three tools, each of which guides them to alternative thoughts through interactions with the app characters.

PS teaches the participants how to break down the issue at hand, to specify a concrete and achievable objective for it, to brainstorm possible solutions, to compare their advantages and disadvantages, and finally to choose the most desirable action and act on it. A worksheet to guide the participants through this process is provided.

AT consists of psychoeducation of assertive communication in contrast to aggressive or passive communication. The participant learns how to express their true feelings and wishes without hurting others or sacrificing themselves. They fill in worksheets to construct appropriate lines in response to their own real-life interactions.

BI teaches the mechanisms of healthy sleep, invites the participant to keep daily sleep records, based on which the participant will start applying sleep restriction and stimulus control techniques, the two proven behavioural skills to increase sleep efficiency.

Each component consists of seven to nine lessons, divided into two to four chapters as appropriate (table 1), as well as worksheets to practise the skill (figure 1). The participants are expected to complete one chapter per week and can only move on to the next chapter after a week has passed since they started the previous chapter and after they have completed one worksheet. The participants receive weekly encouragement emails, templated but tailored in accordance with each participant's progress by the trial coordinators in the management team (the coordinators have no CBT background and hence are forbidden to provide any therapeutic contents) and are prompted to fill in PHQ-9 every week.

The reason why we emphasise and use BA in arms 6–9 and how we will analyse them will be explained in the Statistical methods section below.

Controls

In psychotherapy trials, there is no gold standard control condition like the pill placebo in pharmacotherapy trials. Psychotherapy controls must be designed in accordance with the participants' needs, the available resources and the clinical questions of the study,²² and may produce different effect size estimates.²³ In the RESILIENT trial, we therefore will use three different control conditions with different levels of stringency.

1. Weekly self-checks
2. Health information
3. Delayed intervention

In the weekly self-check arm, the participants will receive weekly encouragement emails and monitor their moods through the weekly PHQ-9 up to week 6 (as in the intervention arms) and then the monthly PHQ-9 thereafter up to week 50. This arm is intended as an attention control to match the attention provided through encouragement emails and self-checks but lacking in active interventions.

In the health information arm, the participants will receive URL of websites containing tips for healthy life (physical activities, nutrition and oral health, none of which focuses on mental health) for the initial 3 weeks and will be asked to answer quizzes for comprehension. They will be asked to fill in self-reports at weeks 3 and 6 (without encouragement emails), and then follow-up evaluations thereafter up to week 50. This arm is intended as a placebo intervention.

In the delayed treatment arm, the participants will be placed on the wait list and be asked to fill in self-reports at weeks 3 and 6. They will then be randomised to 1 of the 11 arms (#1 through #11) at week 6, if they so wish then, and will follow respective programmes thereafter.

Concomitant interventions

All the participants, in the active or control arms, are free to seek whatever mental health interventions they wish through the 50 weeks. The received professional interventions, either pharmacotherapy or psychotherapy, will be monitored and recorded at weeks 6 and 50.

New interventions

The RESilIENT trial is designed as a platform trial to accommodate any new psychosocial intervention. In order to examine a new skill (X), two new arms will be added, namely X and BA+X, to allow a new 2×2 factorial trial, using the BA and the control condition in common with the foregoing factorial trials. However, for the new factorial trial, we will use only those who were concurrently randomised as the new intervention arms X and BA+X. The necessary sample size for the new factorial trial will be calculated de novo, when the addition is planned, and the addition must be approved by the Ethics Committee.

Measures

Primary outcomes

The primary outcome is the change in PHQ-9 from baseline to week 6 (analysis 1).

Secondary outcomes

The secondary outcomes include:

1. Incidence of a major depressive episode by week 50 (analysis 2).
2. Changes in PHQ-9 from baseline to weeks 1, 2, 3, 4, 5; 10, 14, 18, 22, 26, 30, 34, 38, 42, 46, 50.
3. Changes in the Generalized Anxiety Disorder-7 (GAD-7) from baseline to weeks 3, 6, 26 and 50, as anxiety usually coexists with depression and we aim to examine what effects the intervention components, while mainly targeting depression, may have on anxiety. We also aim to monitor broad psychopathology of common mental disorders encompassing depression, anxiety and insomnia as well as positive mental health which are deteriorated in common mental disorders.
4. Changes in the Insomnia Severity Index (ISI) from baseline to weeks 3, 6, 26 and 50.
5. Changes in the Short Warwick Edinburgh Mental Well-Being Scale (SWEMWBS) from baseline to weeks 3, 6, 26 and 50.
6. Changes in the CBT Skills Scale from baseline to weeks 6, 26 and 50.
7. Changes in the Work and Social Adjustment Scale (WSAS), the Utrecht Work Engagement Scale (UWES), the Health and Work Performance Questionnaire (HPQ) and the EuroQOL-5D-5L (EQ-5D-5L) from baseline to weeks 6 and 50.
8. The Client Satisfaction Questionnaire-3 (CSQ-3), adherence to smartphone CBT, co-interventions, safety information at weeks 6 and 50.
9. Medical expenditures and absenteeism among those participating through their health insurance societies by week 50.

Baseline and outcome variables

We will measure the following baseline variables as potential prognostic factors (characteristics that predict the outcome regardless of the intervention) and effect modifiers (characteristics that predict the differences in outcomes between the interventions) for iCBT efficacy.

1. Demographic variables

- Age
- Sex
- Marital status
- Number of cohabitants
- Socioeconomic status of the residential area (as estimated by the zip code)
- Education
- Employment
- Physical conditions and illnesses, including past and current treatments
- Physical disability
- Mental health conditions, including past and current treatments
- Alcohol use and CAGE questionnaire^{24 25}

2. Psychosocial variables

- Short form of the Big-5 questionnaire^{26–28}
- Adverse Childhood Experiences (ACE)^{29 30}
- Assessment of Signal Cases (ASC)-Life difficulties³¹
- ASC-Social support³¹
- CBT Skills Scale³²
- Readiness for smartphone CBT

ASC-Motivation³¹

Familiarity with smartphone usage

3. Clinical variables

- PHQ-9²¹
- GAD-7³³
- ISI³⁴
- Past and current major depressive episodes according to the major depression section of the computerised WHO Composite International Diagnostic Interview (CIDI)^{35 36}
- WSAS³⁷
- HPQ^{38 39}
- UWES⁴⁰
- EQ-5D-5L^{41 42}
- SWEMWBS^{43 44}
- Co-interventions
- CSQ-3⁴⁵
- Adherence to smartphone CBT
- Safety information
- Health insurance data (claims data, absenteeism) among those participating through their health insurance societies

Table 2 shows the overall schedule of baseline and outcome measurements.

Instruments

Adverse Childhood Experiences

The ACE is one of the established and frequently used questionnaires to assess childhood maltreatment including psychological and physical abuse during childhood and other adverse environments. Each relevant experience is rated as having been present or not.²⁹ The reliability and validity of its Japanese version have been established.³⁰

Table 2 Measurement schedule

	Questionnaire part I at landing page	Questionnaire part II on smartphone app (week 0)										Weeks		
		Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Weeks 10, 14, 18, 22	Week 26	Weeks 30, 34, 38, 42, 46	Week 50			
Allowances (days)	–	+6	+6	+6	+6	+6	+14	+14	+14	+14	+28			
PHQ-9	○	Δ	Δ	○	Δ	Δ	○	○	○	○	○			
GAD-7	○			○			○		○		○			
ISI	○		○	○		○	○		○		○			
SWEMWBS	○		○	○		○	○		○		○			
CIDI	○										○			
Baseline characteristics	○													
ASC-Motivation	○		○											
ICT usage	○													
ACE	○													
Short Big-5	○													
ASC-Life difficulties	○													
ASC-Social support	○													
CBT skills	○						○		○		○			
WSAS	○						○		○		○			
UWES	□						□		□		□			
HPQ	□						□		□		□			
EQ-5D-5L	○						○		○		○			
Co-interventions							○		○		○			
CSQ-3							○		○		○			

○: to be administered to all participants; □: only to those currently employed; and Δ: all participants except for groups 10 and 11.

.ACE, Adverse Childhood Experiences; ASC, Assessment of Signal Cases; CBT, cognitive-behavioural therapy; CIDI, Composite International Diagnostic Interview; CSQ-3, Customer Satisfaction Questionnaire-3; EQ-5D-5L, EuroQOL-5D-5L; GAD-7, Generalized Anxiety Disorder-7; HPQ, Health and Work Performance Questionnaire; ICT, information and communication technologies; ISI, Insomnia Severity Index; PHQ-9, Patient Health Questionnaire-9; SWEMWBS, Short Warwick Edinburgh Mental Well-Being Scale; UWES, Utrecht Work Engagement Scale; WSAS, Work and Social Adjustment Scale.

Assessment of Signal Cases

The ASC is designed to measure factors shown to influence the course of the treatment including social support, life difficulties and motivation.³¹ Each subscale contains six to seven items, each rated on a 5-point scale for the past week. The validity and reliability of the original version have been ascertained.³¹ The Japanese version was developed through back-translation and its reliability will be examined in this current study.

CBT Skills Scale

In order to measure specific cognitive or behavioural skills, we have developed the short versions based on the following established questionnaires: CR,⁴⁶ BA,^{47 48} assertiveness,⁴⁹ PS⁵⁰ and sleep hygiene.⁵¹ Each item is rated on a Likert scale between 0=barely true of me through 3=mostly true of me, and the average of the score items constitutes the score for each skill. The reliability and validity of the CBT Skills Scale have been ascertained.³²

Composite International Diagnostic Interview

To determine the Diagnostic and Statistical Manual of Mental Disorders, 4th Ed. (DSM-IV) major depressive disorder diagnosis during the lifetime and the previous 12 months, we will use the self-administered Japanese WHO-CIDI 3.0 depression section.^{52 53} The self-administered version has been demonstrated to be reliable in a 1-year test-retest survey³⁶ and to have good concordance with the clinical diagnosis of major depressive episode.⁵⁴

Client Satisfaction Questionnaire-3

The CSQ-3 is the short version of the longer, 8-item and 18-item version to measure client satisfaction with psychotherapies. Its reliability and validity have been established.⁴⁵

EuroQOL-5D-5L

The EQ-5D is one of the most frequently and well-established instruments to describe and value health in the five domains of mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each item was rated in three levels in its original version but has now been expanded to five levels to enhance the scale's sensitivity.⁴¹ The population norms for the Japanese have been established.⁴²

Generalized Anxiety Disorder-7

The GAD-7, which consists of seven questions that depict anxiety, tension and worry, was created to gauge the severity of generalised anxiety. Each item is given a value ranging from 0=not at all to 3=almost every day, with the total score falling between 0 and 21. Its validity and reliability have been proven.³³ The PHQ-9 and GAD-7 have been recommended and used as part of a standardised battery to measure health outcomes for depression and anxiety.^{55 56}

Health and Work Performance Questionnaire

We will employ one item from the WHO HPQ's presenteeism subscale. The reliability and validity of the original scale have been established.^{38 57} The reliability and validity of the one-item version have been ascertained.³⁹

Insomnia Severity Index

The ISI was developed and validated as a brief screening measure of insomnia and as an outcome measure in treatment research. It has adequate reliability and validity to quantify perceived insomnia severity³⁴ and has been used extensively in insomnia research. The reliability and validity of the Japanese version have been confirmed.⁵⁸

Personal Health Questionnaire-9

The PHQ-9 comprises the nine DSM-5 diagnostic criteria items for a major depressive episode.²¹ Each item is given a value ranging from 0=not at all to 3=almost every day, with the overall score falling between 0 and 27. Both the original PHQ-9 and its Japanese version have satisfactory validity and reliability.^{59 60} The PHQ-9 and GAD-7 have been recommended and used as part of a standardised battery to measure health outcomes for depression and anxiety.^{55 56}

Short form of the Big-5 questionnaire

There are several validated scales to measure personality traits according to the five-factor model. We will use the Big-Five Scale of Personality Trait Adjectives,²⁶ which is widely used in Japan, in its abridged form. The short version's validity and reliability have been established.^{27 28} Each of the five personality traits—neuroticism, extraversion, openness, agreeableness and conscientiousness—is assessed using five to seven related adjectives, each of which is rated on a 5-point Likert scale from 0=untrue of me to 4=true of me.

Short Warwick Edinburgh Mental Well-Being Scale

The SWEMWBS is a short seven-item version of the Warwick Edinburgh Mental Well-Being Scale, developed to measure positive mental well-being at the general population level.⁴³ The scale is based on a conceptual framework that mental well-being consists of two key dimensions of hedonia and eudaimonia. Each item is rated on a 5-point Likert scale between 1=none of the time and 5=all of the time. The reliability and validity of the SWEMWBS have been well confirmed.⁶¹ The Japanese version has been developed using the standard back-translation method.⁶²

Utrecht Work Engagement Scale

Engagement in work has been conceptualised in three dimensions of vigour (high levels of energy while working), dedication (sense of significance and enthusiasm) and absorption (sense of concentration and engrossment). We will use the ultra-short, three-item version of the widely used UWES. The reliability and validity of the scale, including its Japanese version, have been confirmed.⁴⁰

Work and Social Adjustment Scale

We will measure the impairments in social functioning in the domains of work, home management, social leisure activities, private leisure activities and close relationships, each rated on a 9-point Likert scale between 0=not at all through 8=very severely. The reliability and validity of the scale have been established for the original³⁷ as well as the Japanese version.⁶³

Participant timeline

Figure 2 shows the participant timeline.

Statistical methods

The 2×2 factorial design

The RESiLIENT trial is a master protocol for the following four 2×2 factorial trials, each of which is to be analysed independently.

Randomisation for BA		
Randomisation for CR	BA+, CR+	BA-, CR+
	BA+, CR-	BA-, CR-
Randomisation for BA		
Randomisation for PS	BA+, PS+	BA-, PS+
	BA+, PS-	BA-, PS-
Randomisation for BA		
Randomisation for AT	BA+, AT+	BA-, AT+
	BA+, AT-	BA-, AT-
Randomisation for BA		
Randomisation for BI	BA+, BI+	BA-, BI+
	BA+, BI-	BA-, BI-

The four factorial trials will use the BA arm and the control arm in common. We will use the no treatment arm for the primary analysis, and use the self-check arm and the health information arm as sensitivity analyses using increasingly stringent comparisons.

We use the factorial design for the following reasons. First, the factorial design can increase the statistical power to examine the efficacy of the included skills. Second, in the factorial design, we can test the interaction between the iCBT skills.^{64–66} We anticipate that the optimal iCBT packages would likely include two or three skills. While the available studies indicate that there probably are no important interactions among various iCBT skills,^{17 18} it would be informative to have tested the interactions when we eventually propose combinations of iCBT skills.

We test out the combinations of various skills with BA for the following reasons. First, BA was the most efficacious and the only statistically significant component in the individual patient data component network meta-analysis of iCBT components and delivery methods.¹⁸ Second, BA

alone has often been shown to be as efficacious as the more complex CBT packages in the network meta-analysis of various psychotherapies for depression.^{8 67} Third, in trials that aimed at increasing scalability of CBT interventions, the active interventions often included BA.⁶⁸ Lastly, the foregoing studies that used different versions of the Resilience Training App included BA when they successfully differentiated the intervention from the control condition^{15 16} or suggested that BA may be the most efficacious though not statistically significantly so among the included components in the fully factorial trial examining CR, BA, PS and AT.¹⁷

Statistical analyses

All analyses will be conducted with the intention-to-treat sample.

Analysis 1

For the factorial trial, we will use the mixed-effects repeated-measures analysis (MMRM) to estimate the mean difference in change scores on PHQ-9 for each component: the model will include treatments, visit (as categorical) and treatment-by-visit interaction as fixed effects, adjusted for the baseline PHQ-9 scores, employment status, age and sex. The primary outcome is the change in PHQ-9 scores from baseline to week 6. We will use the unstructured variance-covariance matrix for the MMRM to account the within-participant correlations: in case of non-convergence, we will use the Toeplitz, heterogeneous compound asymmetry, autoregressive (1), compound symmetry, and variance components models in this order. No adjustment for multiple testing will be applied in examination of statistical significance of the main effects and the conventional threshold for statistical significance ($p < 0.05$, two sided) will be used, because conventionally, the multiple hypotheses have been tested independently in these trial designs assuming as if the tested interventions could have been assessed in separate trials.^{64 65 69 70} We will assess the interaction between the treatments in a sensitivity analysis. Previous research has not been suggestive of interactions among iCBT components,^{17 18 71} but when we identify a strong interaction, we will interpret the results considering the interaction. We will use the delayed treatment arm for the primary comparison but the health information and the self-check arms as sensitivity analyses.

We will examine the influence of baseline characteristics as prognostic factors by entering them as covariates and as effect modifiers by entering their interaction with treatment in the above MMRM.

We will use a similar MMRM model for GAD-7, ISI and SWEMWBS at week 6. For CBT skills, WSAS, UWES, HPQ and EQ-5D-5L, which are measured only at baseline and at week 6, their change scores will be assessed using the analysis of covariance: missing baseline values will be imputed through multiple imputation. For CSQ-3, which is measured only at week 6, its endpoint score will be

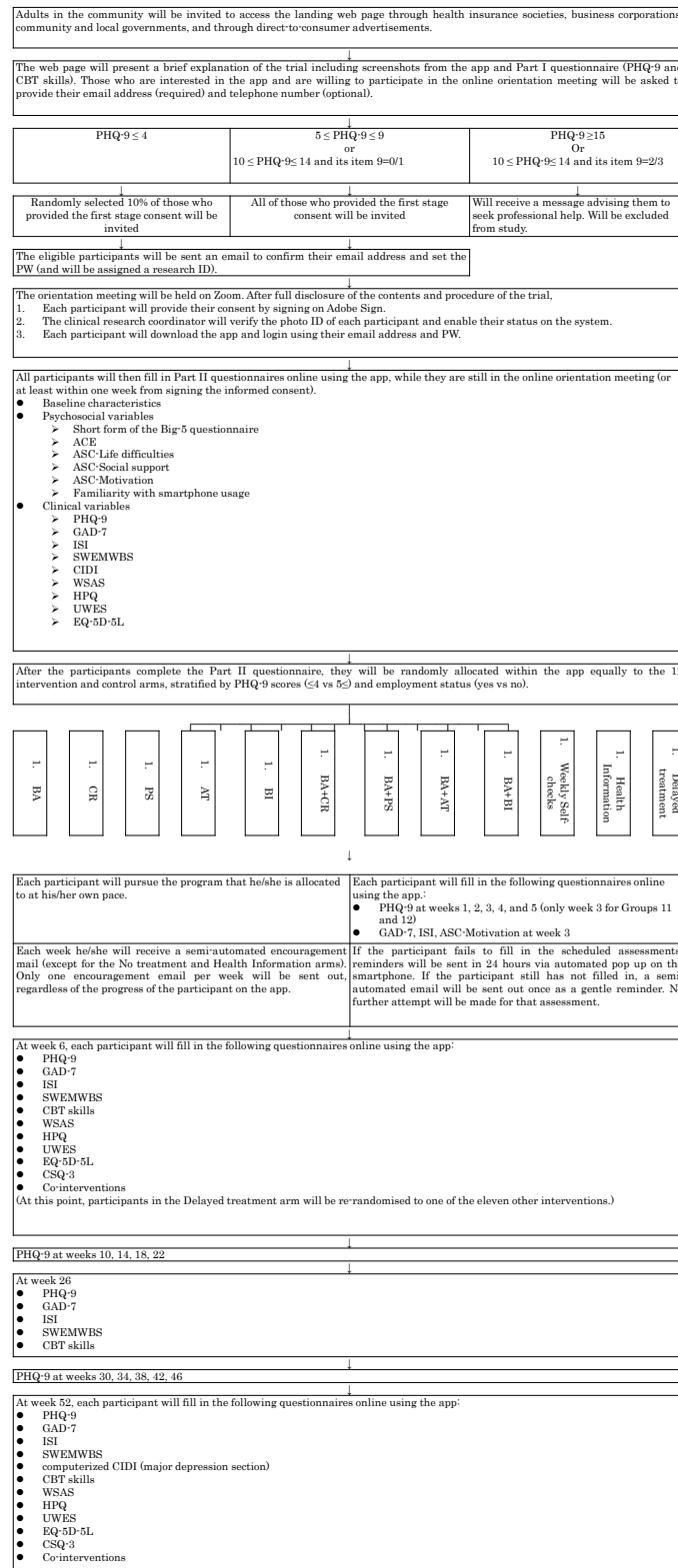


Figure 2 Participant timeline. ACE, Adverse Childhood Experiences; ASC, Assessment of Signal Cases; AT, assertion training; BA, behavioural activation; BI, behaviour therapy for insomnia; CBT, cognitive-behavioural therapy; CIDI, Composite International Diagnostic Interview; CR, cognitive restructuring; CSQ-3, Client Satisfaction Questionnaire-3; EQ-5D-5L, EuroQOL-5D-5L; GAD-7, Generalized Anxiety Disorder-7; HPQ, Health and Work Performance Questionnaire; ISI, Insomnia Severity Index; PHQ-9, Patient Health Questionnaire-9; PS, problem-solving; PW, password; SWEMWBS, Short Warwick Edinburgh Mental Well-Being Scale; UWES, Utrecht Work Engagement Scale; WSAS, Work and Social Adjustment Scale.

assessed using the analysis of variance: missing baseline values will be imputed through multiple imputation.

Analysis 2

An exploratory analysis will be conducted to compare the incidence of a major depressive episode between the proposed optimal skill(s) against the control conditions through the time-to-event analysis (Kaplan-Meier survival analysis) among those with or without baseline subthreshold depression and who did not have a major depressive episode at baseline. The influence of baseline characteristics on the time-to-event will be explored by using Cox proportional hazard analysis if the proportional hazard assumption is met.

Preplanned secondary analyses

The following secondary analyses are preplanned, to make the analyses more sensitive to changes in respective scales (SWEMWBS for well-being and ISI for insomnia) by limiting the participants to those with potentials for improvement and minimising ceiling or floor effects:

1. Efficacy of various components on SWEMWBS among those with reduced well-being at baseline (defined as scoring 20 or less on the SWEMWBS).^{44 61}
2. Efficacy of the BI component on sleep parameters among those with substantive insomnia at baseline (defined as scoring 8 or more on the ISI).³⁴ The primary outcome of this secondary analysis will be the ISI but will include other sleep parameters such as sleep efficiency, total sleep, sleep latency, etc.

Interim analysis

No interim analysis is planned.

Sample size

The RESiLIENT trial will be powered for each factorial trial for its primary outcome of the analysis 1, that is, PHQ-9 at week 6 among those with the baseline PHQ-9 scores of 5 or more. The Fun the Learn, Act and Think through Technology trial has shown a standardised mean difference of 0.3–0.4 for the earlier version of the Resilience Training App using the combination of BA+CR over the waiting list control among patients with drug-refractory depression.¹⁵ In order to detect a standardised mean difference of 0.2 at $\alpha=0.05$ and $\beta=0.10$, each factorial trial in the current master protocol would need in total 1053 participants (FactorialPowerPlan, <http://methodology.psu.edu>), hence 264 in each arm. Altogether the RESiLIENT trial will require $264 \times 12 = 3168$ participants with baseline PHQ-9 scores of 5 or more. The Healthy Campus Trial (HCT) had 5% dropout rates for its eighth week outcome.¹⁷ We may expect somewhat higher dropout rates in the RESiLIENT trial, which targets the general adult population in the community. Assuming a 10% dropout rate, the required sample size will be 3520.

We will include all participants with any baseline PHQ-9 scores for analysis 2, but the power for analysis 2 may be low due to the low event rates among the participants. The post hoc power calculation will be conducted for the

expected 20% reduction in terms of HR.² In the HCT participants, which used the same recruitment procedure and entered all those with PHQ-9 scores of 5 or more but entered only a random one-tenth of those scoring 4 or less, the latter were approximately half of the former: we therefore anticipate approximately 5200 participants altogether for the RESiLIENT trial.

Allocation and blinding

We will use permuted block randomisation, stratified by PHQ-9 scores (≤ 4 vs $5 \leq$) and employment status (yes vs no). The block size will be known to the study statistician only, who will generate the random sequence by using SAS PROC PLAN.

The above sequence will be implemented in the app and participants will be allocated to 1 of the 12 intervention or control arms automatically by the app when they complete the part I questionnaires.

Neither the participants nor the trial management team will be blinded to the interventions. However, the trial statistician will be blinded to the interventions until after they sign off the completed statistical reports.

Data collection and management

All the data will be entered into the web or the smart-phone app by the participants. The data will be transferred securely to the server using the Secure Socket Layer. The server program behind the web and the smart-phone app has built-in automated checks for the inappropriate or missing data entries.

The trial management team will send out encouragement emails weekly (up to week 6) and at weeks 26 and 50, to compliment the participants' progress and also to prompt their responses to the questionnaires. The encouragement emails will not contain any CBT-specific instructions. The trial management team will be based at Kyoto University Department of Health Promotion and Human Behavior.

Data monitoring

The trial management team will submit a report to the independent Data and Safety Monitoring Board (DSMB) every 6 months after enrolment of the first participant. The report will contain information on the number of participants enrolled, their follow-up status and serious adverse events. Based on this report, the DSMB will provide recommendations to the Steering Committee.

The Steering Committee will discontinue the trial when they receive a recommendation or an instruction to the effect from the DSMB or the overseeing Ethics Committee.

Harms

All the serious adverse events (defined as those leading to death, life-threatening, necessitating hospitalisation or leading to permanent serious disabilities) will be handled according to the standard operationalised procedure #10 by Kyoto University Graduate School of Medicine (<http://www.ec.med.kyoto-u.ac.jp/doc/SOP10.pdf>).

In addition, when any rise in suicidality (defined as scoring 10 or more on the total PHQ-9 scores and scoring 2 or more on item 9 of PHQ-9 for 2 consecutive weeks) is noted, the participant will receive an email recommending visiting the consultation services listed by the Japanese Ministry of Health, Labour and Welfare for the general public (<https://kokoro.mhlw.go.jp/agency/>) or other mental health professionals.

Ethics

The protocol has been approved by the Kyoto University Graduate School of Medicine Ethics Committee (approval number C1556). Any amendment to the protocol will be reviewed by the Ethics Committee.

Consent

The participant will provide written informed consent through Adobe Sign after full disclosure of the purpose and procedures of the study based on the written materials as well as through the online orientation meeting.

Confidentiality and access to data

All the participants will be assigned a research ID upon registration and their data will be managed by this ID. No personally identifiable information will be collected except for the email address and telephone number (optional), which will be linked to the research ID. The correspondence table between the email address and the research ID as well as the signed forms will be stored in a separate server different from the main server for the trial data. Only the trial management team will have access to the correspondence table, while the Steering Committee members will have access to the full trial data including the research ID without the correspondence table.

Ancillary and post-trial care

No ancillary or post-trial care is planned.

Patient and public involvement

People with lived experiences and members from the general community have joined the Steering Committee after the study protocol had been semifinalised and the funding had been secured. Since then they have participated in the improvement and finalisation of the smartphone apps as well as the contents and quantity of questionnaires, and also in the preparation of the materials to be shared with the potential participants in the orientation meeting for informed consent.

They will continue to serve in the Steering Committee throughout the current trial and will provide advice as needed on the conduct of the trial as well as possible future additions/modifications to the study.

DISCUSSION

This protocol paper described the study protocol for the master trial involving four 2×2 factorial trials examining five CBT skills (CR, BA, PS, AT and BI) implemented on the smartphone platform among 5200 participants

with or without baseline subthreshold depression. We plan to recruit 3520 adults with subthreshold depression to examine the acute phase effectiveness to reduce depressive symptoms, and further ca 1700 adults without subthreshold depression to examine the long-term effectiveness to prevent depression among those with or without baseline subthreshold depression but without baseline major depressive episode.

The trial has several novel design and implementation features. First, it aims to elucidate specific efficacy of the included CBT skills by way of the factorial experiments. The factorial design also enables examination of interactions among the CBT skills. Second, it is a master protocol to conduct four factorial trials simultaneously and makes efficient use of the study participants across the planned four factorial trials. Third, the master protocol allows room for addition of a new arm in the course of the trial if necessary. Fourth, it uses remote recruitment and informed consent procedures taking full advantage of the internet. We expect that this will allow us to recruit a fairly large sample within relatively short periods. Fifth, the large sample size will allow exploration of prognostic factors and effect modifiers to match individual characteristics with each CBT skill. Lastly, we believe that the patient and public involvement in the preparatory phases of this study has not only made the study more relevant to their needs but also increased its feasibility.

Given the novelty and scale of the study, we anticipate the following difficulties. First, the smartphone app and the backend server system controlling the smartphone app for this study are bound to be complex and sophisticated, due to the multiplicities of the interventions and the complexities of the study design. We have dedicated months and years to develop the platform, have pilot tested it before the start of the enrolment and plan to implement the initial phase 3–6 months where we only slowly recruit the participants to adjust remaining glitches in the system if any and to ascertain that the interventions and assessments are delivered as planned. Second, we can never be sure of the success of the recruitment. Our past experiences suggest that we may have to advertise our study to some 200 000 people to ensure the sample size of 5200 eligible participants. At the time of the writing of this protocol paper (July 2022), the total population accessible to our collaborators sums up to 400 000. Should this turn out to be insufficient, we will seek additional collaborators.

In conclusion, we expect our study to provide valuable information on the specific efficacies of individual CBT skills, the prognostic factors and effect modifiers for them to enable matching between individual characteristics and CBT components, and finally to build more efficient and effective smartphone CBT apps to promote mental health and reduce depression burden in the general population.

Ethics and dissemination

The protocol has been approved by the Kyoto University Graduate School of Medicine Ethics Committee (approval number C1556).

We will disseminate the results from the study in scientific meetings and in scientific journals. The main results will be posted in UMIN-CTR and on the home page of the trial.

The deidentified data will be uploaded to UMIN-ICDR (<http://www.umin.ac.jp/icdr/index-j.html>) after the main results have been published, in order to enable secondary use of the data in the interest of public health and medicine.

Author affiliations

¹Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine/School of Public Health, Kyoto, Japan

²National Center of Neurology and Psychiatry, Tokyo, Japan

³Center for Cognitive Behavior Therapy and Research, National Center of Neurology and Psychiatry, Tokyo, Japan

⁴Department of Psychiatry and Cognitive-Behavioral Medicine, Nagoya City University Graduate School of Medical Sciences, Nagoya, Aichi, Japan

⁵Department of Digital Mental Health, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

⁶Health Informatics, Kyoto University School of Public Health, Kyoto, Japan

⁷Department of Social Epidemiology, Kyoto University Graduate School of Medicine/School of Public Health, Kyoto, Japan

⁸Institute for Future Initiative, The University of Tokyo, Tokyo, Japan

⁹Human Health Sciences, Kyoto University Graduate School of Medicine, Kyoto, Japan

¹⁰Department of Data Science, The Institute of Statistical Mathematics, Tachikawa, Tokyo, Japan

¹¹Black Dog Institute, Randwick, New South Wales, Australia

¹²Department of Health Care Policy, Harvard Medical School, Boston, Massachusetts, USA

¹³Department of Clinical, Neuro and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

¹⁴Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne, UK

Twitter Toshi A Furukawa @Toshi_FRKW

Contributors TAF conceived the study. TAF, HN, PC and JW designed the study, in repeated consultations with AT, MS, YL, RT, MH, TA, NKawakami, TN, NKondo, SF, HC and RCK. TAF drafted the first manuscript, while HN, PC, JW, AT, MS, YL, RT, MH, TA, NKawakami, TN, NKondo, SF, HC and RCK made critical contributions to the manuscript. All the authors have seen the final manuscript and approved it for submission.

Funding The trial is funded by the Japan Agency for Medical Research and Development (AMED) (grant number JP22de0107005).

Disclaimer The funder has no role in study design, data collection, management, analysis and interpretation of data, nor in writing or submission of this manuscript.

Competing interests TAF, MH and TA jointly possess intellectual properties of the Resilience Training App. TAF reports personal fees from Boehringer-Ingelheim, DT Axis, Kyoto University Original, Shionogi and SONY, and a grant from Shionogi, outside the submitted work; In addition, TAF has patents 2020-548587 and 2022-082495 pending, and intellectual properties for Kokoro-app licensed to Mitsubishi-Tanabe. AT received lecture fees from Sumitomo Dainippon Pharma, Eisai, Janssen Pharmaceutical, Meiji-Seika Pharma, Mitsubishi Tanabe Pharma, Otsuka, and Takeda Pharmaceutical. MS reports personal fee from SONY outside the submitted work. TA has received lecture fees from Chugai, Daiichi-Sankyo, Dainippon-Sumitomo, Eisai, Janssen, Kyowa, Lilly, MSD, Meiji-Seika Pharma, Mochida, Nipro, Nihon-Zoki, Otsuka, Pfizer, Takeda and Viatrix; and royalties from Igaku-shoin, outside the submitted work. TA has patents (2019-01749 and 2020-135195&2022-069057) pending. NKA reports personal fees from Junpukai Foundation, Riken Institute, JAXA, Sekisui Chemical Co and SB@Work, outside

the submitted work; and Department of Digital Mental Health is an endowment department, supported with an unrestricted grant from 15 enterprises (<https://dmh.m.u-tokyo.ac.jp/c/>). TN reports personal fees from Pfizer, MSD, Ohtsuka, Chuga, Dentsu, Takeda, Janssen, Boehringer Ingelheim, Eli Lilly, Mitsubishi Tanabe, Novartis, Allergan, Maruho, Novo Nordisk, TOA EIYO; and research grants from I&H Co, Cocokaraffine Group Co and Konica Minolta, all outside the submitted work. HN received a research fund from GlaxoSmithKline and consulting fees from Kyowa Kirin, GlaxoSmithKline, Ono Pharmaceutical, Sony, Terumo, outside the submitted work. JMSW reports research funding from Boehringer-Ingelheim, GSK, Intercept, UCB, all outside the current work. All the other coauthors declare no conflicts of interests.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Toshi A Furukawa <http://orcid.org/0000-0003-2159-3776>

Norito Kawakami <http://orcid.org/0000-0003-1080-2720>

Naoki Kondo <http://orcid.org/0000-0002-6425-6844>

Shingo Fukuma <http://orcid.org/0000-0002-8379-8761>

Hisashi Noma <http://orcid.org/0000-0002-2520-9949>

Pim Cuijpers <http://orcid.org/0000-0001-5497-2743>

James M S Wason <http://orcid.org/0000-0002-4691-126X>

REFERENCES

- 1 Lobel B, Hirschfeld RM. *Depression: what we know (DHHS publication no. ADM 84-1318)*. Rockville, MD: National Institute of Mental Health, 1984.
- 2 Cuijpers P, Pineda BS, Quero S, *et al.* Psychological interventions to prevent the onset of depressive disorders: a meta-analysis of randomized controlled trials. *Clin Psychol Rev* 2021;83:101955.
- 3 Mihalopoulos C, Vos T, Pirkis J, *et al.* The economic analysis of prevention in mental health programs. *Annu Rev Clin Psychol* 2011;7:169–201.
- 4 Ormel J, Hollon SD, Kessler RC, *et al.* More treatment but no less depression: the treatment-prevalence paradox. *Clin Psychol Rev* 2022;91:102111.
- 5 Furukawa TA. An epidemic or a plague of common mental disorders? *Acta Psychiatr Scand* 2019;140:391–2.
- 6 Judd LL, Schettler PJ, Akiskal HS. The prevalence, clinical relevance, and public health significance of subthreshold depressions. *Psychiatr Clin North Am* 2002;25:685–98.
- 7 Herrman H, Patel V, Kieling C, *et al.* Time for united action on depression: a lancet-world psychiatric association commission. *Lancet* 2022;399:957–1022.
- 8 Cuijpers P, Quero S, Noma H, *et al.* Psychotherapies for depression: a network meta-analysis covering efficacy, acceptability and long-term outcomes of all main treatment types. *World Psychiatry* 2021;20:283–93.
- 9 Furukawa TA, Shinohara K, Sahker E, *et al.* Initial treatment choices to achieve sustained response in major depression: a systematic review and network meta-analysis. *World Psychiatry* 2021;20:387–96.
- 10 Cuijpers P, Noma H, Karyotaki E, *et al.* Effectiveness and acceptability of cognitive behavior therapy delivery formats in adults with depression: a network meta-analysis. *JAMA Psychiatry* 2019;76:700–7.
- 11 Karyotaki E, Efthimiou O, Miguel C, *et al.* Internet-Based cognitive behavioral therapy for depression: a systematic review and individual patient data network meta-analysis. *JAMA Psychiatry* 2021;78:361–71.
- 12 Reins JA, Buntrock C, Zimmermann J, *et al.* Efficacy and moderators of Internet-based interventions in adults with subthreshold depression: an individual participant data meta-analysis of randomized controlled trials. *Psychother Psychosom* 2021;90:94–106.

- 13 NICE. NICE guideline [NG222]. In: *Depression in adults: treatment and management*. 2022.
- 14 Cook L, Mostazir M, Watkins E. Reducing stress and preventing depression (respond): randomized controlled trial of web-based rumination-focused cognitive behavioral therapy for high-ruminating university students. *J Med Internet Res* 2019;21:e11349.
- 15 Mantani A, Kato T, Furukawa TA, et al. Smartphone cognitive behavioral therapy as an adjunct to pharmacotherapy for refractory depression: randomized controlled trial. *J Med Internet Res* 2017;19:e373.
- 16 Akechi T, Yamaguchi T, Uchida M, et al. Smartphone psychotherapy reduces fear of cancer recurrence among breast cancer survivors: a fully decentralized randomized controlled clinical trial. *J Clin Oncol in Press* 2023;41:1069–78.
- 17 Sakata M, Toyomoto R, Yoshida K, et al. Components of smartphone cognitive-behavioural therapy for subthreshold depression among 1093 university students: a factorial trial. *Evid Based Ment Health* 2022;25:e18–25.
- 18 Furukawa TA, Suganuma A, Ostinelli EG, et al. Dismantling, optimising, and personalising internet cognitive behavioural therapy for depression: a systematic review and component network meta-analysis using individual participant data. *Lancet Psychiatry* 2021;8:500–11.
- 19 Chan A-W, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med* 2013;158:200–7.
- 20 Boutron I, Altman DG, Moher D, et al. Consort statement for randomized trials of nonpharmacologic treatments: a 2017 update and a consort extension for nonpharmacologic trial Abstracts. *Ann Intern Med* 2017;167:40–7.
- 21 Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001;16:606–13.
- 22 Gold SM, Enck P, Hasselmann H, et al. Control conditions for randomised trials of behavioural interventions in psychiatry: a decision framework. *Lancet Psychiatry* 2017;4:725–32.
- 23 Michopoulos I, Furukawa TA, Noma H, et al. Different control conditions can produce different effect estimates in psychotherapy trials for depression. *J Clin Epidemiol* 2021;132:59–70.
- 24 Fiellin DA, Reid MC, O'Connor PG. Screening for alcohol problems in primary care: a systematic review. *Arch Intern Med* 2000;160:1977–89.
- 25 Bush B, Shaw S, Cleary P, et al. Screening for alcohol abuse using the cage questionnaire. *Am J Med* 1987;82:231–5.
- 26 Wada S. Construction of the big five scales of personality trait terms and concurrent validity with NPI [in Japanese]. *The Japanese Journal of Psychology* 1996;67:61–7.
- 27 Namikawa T, Tani I, Wakita T, et al. Development of a short form of the Japanese big-five scale, and a test of its reliability and validity. *Shinrigaku Kenkyu* 2012;83:91–9.
- 28 Toyomoto R, Sakata M, Yoshida K, et al. Validation of the Japanese big five scale short form in a university student sample. *Front Psychol* 2022;13:862646.
- 29 Felitti VJ, Anda RF, Nordenberg D, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The adverse childhood experiences (ACE) study. *Am J Prev Med* 1998;14:245–58.
- 30 Tsuboi S. Creating a Japanese version adverse childhood experiences score calculator. 科学研究費助成事業 研究成果報告書 2014.
- 31 Probst T, Lambert MJ, Dahlbender RW, et al. Providing patient progress feedback and clinical support tools to therapists: is the therapeutic process of patients on-track to recovery enhanced in psychosomatic in-patient therapy under the conditions of routine practice? *J Psychosom Res* 2014;76:477–84.
- 32 Sakata M, Toyomoto R, Yoshida K, et al. Development and validation of the cognitive behavioural therapy skills scale among college students. *Evid Based Ment Health* 2021;24:70–6.
- 33 Spitzer RL, Kroenke K, Williams JBW, et al. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006;166:1092–7.
- 34 Bastien CH, Vallières A, Morin CM. Validation of the insomnia severity index as an outcome measure for insomnia research. *Sleep Med* 2001;2:297–307.
- 35 Peters L, Andrews G. Procedural validity of the computerized version of the composite international diagnostic interview (CIDI-auto) in the anxiety disorders. *Psychol Med* 1995;25:1269–80.
- 36 Shimoda H, Inoue A, Tsuno K, et al. One-Year test-retest reliability of a Japanese web-based version of the who composite international diagnostic interview (CIDI) for major depression in a working population. *Int J Methods Psychiatr Res* 2015;24:204–12.
- 37 Mundt JC, Marks IM, Shear MK, et al. The work and social adjustment scale: a simple measure of impairment in functioning. *Br J Psychiatry* 2002;180:461–4.
- 38 Kessler RC, Barber C, Beck A, et al. The world Health organization health and work performance questionnaire (HPQ). *J Occup Environ Med* 2003;45:156–74.
- 39 Kawakami N, Inoue A, Tsuchiya M, et al. Construct validity and test-retest reliability of the world mental health Japan version of the world Health organization health and work performance questionnaire short version: a preliminary study. *Ind Health* 2020;58:375–87.
- 40 Schaufeli WB, Shimazu A, Hakanen J, et al. An ultra-short measure for work engagement. *European Journal of Psychological Assessment* 2019;35:577–91.
- 41 Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;20:1727–36.
- 42 Ikeda S, Shiroya T, Igarashi A, et al. Developing a Japanese version of the EQ-5D-5L value set. *Journal of National Institute of Public Health* 2015;64:47–55.
- 43 Bartam DJ, Sinclair JM, Baldwin DS. Further validation of the warwick-edinburgh mental well-being scale (WEMWBS) in the UK veterinary profession: Rasch analysis. *Qual Life Res* 2013;22:379–91.
- 44 Shah N, Cader M, Andrews B, et al. Short warwick-edinburgh mental well-being scale (SWEMWBS): performance in a clinical sample in relation to PHQ-9 and GAD-7. *Health Qual Life Outcomes* 2021;19:260.
- 45 Attkisson CC, Zwick R. The client satisfaction questionnaire. psychometric properties and correlations with service utilization and psychotherapy outcome. *Eval Program Plann* 1982;5:233–7.
- 46 Strunk DR, Hollars SN, Adler AD, et al. Assessing patients' cognitive therapy skills: initial evaluation of the competencies of cognitive therapy scale. *Cognit Ther Res* 2014;38:559–69.
- 47 Manos RC, Kanter JW, Luo W. The behavioral activation for depression scale-short form: development and validation. *Behav Ther* 2011;42:726–39.
- 48 Yamamoto T, Shudo Y, Sakai M. Development of the Japanese version of behavioral activation for depression scale-short form (BADs-SF) and examination of its reliability and validity. *Japanese Journal of Cognitive Therapy* 2015;8:96–105.
- 49 Aikawa A, Fujita M. An attempt to construct a social skills self-rating scale for adults. *Bulletin of Tokyo Gakugei University Part I* 2005;56:87–93.
- 50 Heppner PP, Petersen CH. The development and implications of a personal problem-solving inventory. *Journal of Counseling Psychology* 1982;29:66–75.
- 51 Mastin DF, Bryson J, Corwyn R. Assessment of sleep hygiene using the sleep hygiene index. *J Behav Med* 2006;29:223–7.
- 52 Kessler RC, Üstün TB. The world mental health (WMH) survey initiative version of the world Health organization (who) composite international diagnostic interview (CIDI). *Int J Methods Psychiatr Res* 2004;13:93–121.
- 53 Kawakami N, Takeshima T, Ono Y, et al. Twelve-month prevalence, severity, and treatment of common mental disorders in communities in Japan: preliminary finding from the world mental health Japan survey 2002–2003. *Psychiatry Clin Neurosci* 2005;59:441–52.
- 54 Inoue A, Kawakami N, Tsuno K, et al. Organizational justice and major depressive episodes in Japanese employees: a cross-sectional study. *J Occup Health* 2013;55:47–55.
- 55 Obbarius A, van Maasakkers L, Baer L, et al. Standardization of health outcomes assessment for depression and anxiety: recommendations from the ICHOM depression and anxiety Working group. *Qual Life Res* 2017;26:3211–25.
- 56 Bone C, Simmonds-Buckley M, Thwaites R, et al. Dynamic prediction of psychological treatment outcomes: development and validation of a prediction model using routinely collected symptom data. *Lancet Digit Health* 2021;3:e231–40.
- 57 Kessler RC, Ames M, Hymel PA, et al. Using the world Health organization health and work performance questionnaire (HPQ) to evaluate the indirect workplace costs of illness. *J Occup Environ Med* 2004;46(6 Suppl):S23–37.
- 58 Munezawa T, Morin CM, Inoue Y, et al. Development of the Japanese version of the insomnia severity index (ISI-J). *Jpn J Psychiatr Treat* 2009;24:219–25.
- 59 Furukawa TA. Assessment of mood: guides for clinicians. *J Psychosom Res* 2010;68:581–9.
- 60 Muramatsu K, Miyaoka H, Kamijima K, et al. Performance of the Japanese version of the patient health questionnaire-9 (J-PHQ-9) for depression in primary care. *Gen Hosp Psychiatry* 2018;52:64–9.
- 61 Shah N, Cader M, Andrews WP, et al. Responsiveness of the short warwick Edinburgh mental well-being scale (SWEMWBS): evaluation a clinical sample. *Health Qual Life Outcomes* 2018;16:239.

- 62 Suganuma S, Hirano M, Nakano M, *et al.* Development of the japanese version of warwick-edinburgh mental well-being scale (WEMWBS): the hedonic/eudaimonic state index. *Japanese Journal of Clinical Psychology* 2016;16:471–5.
- 63 Yamamoto T, Koga Y, Sakai M. Development of the japanese version of the work and social adjustment scale and examination of its reliability and validity. *Clinical Psychiatry (Seishin Igaku)* 2019;61:1207–14.
- 64 Montgomery AA, Astin MP, Peters TJ. Reporting of factorial trials of complex interventions in community settings: a systematic review. *Trials* 2011;12:179.
- 65 Mdege ND, Brabyn S, Hewitt C, *et al.* The 2 × 2 cluster randomized controlled factorial trial design is mainly used for efficiency and to explore intervention interactions: a systematic review. *J Clin Epidemiol* 2014;67:1083–92.
- 66 McAlister FA, Straus SE, Sackett DL, *et al.* Analysis and reporting of factorial trials: a systematic review. *JAMA* 2003;289:2545–53.
- 67 Ciharova M, Furukawa TA, Efthimiou O, *et al.* Cognitive restructuring, behavioral activation and cognitive-behavioral therapy in the treatment of adult depression: a network meta-analysis. *J Consult Clin Psychol* 2021;89:563–74.
- 68 Karyotaki E, Araya R, Kessler RC, *et al.* Association of task-shared psychological interventions with depression outcomes in low- and middle-income countries: a systematic review and individual patient data meta-analysis. *JAMA Psychiatry* 2022;79:430–43.
- 69 Fleiss JL. *The design and analysis of clinical experiments*. Hoboken, NJ, USA: John Wiley & Sons, 8 February 1999.
- 70 Balakrishnan N. *Methods and applications of statistics in clinical trials*. Hoboken, NJ, USA: Concepts, Principles, 29 August 2014.
- 71 Efthimiou O, Seo M, Karyotaki E, *et al.* Bayesian models for aggregate and individual patient data component network meta-analysis. *Stat Med* 2022;41:2586–601.

2. Statistical Analysis Plan (SAP)

Noma H, Furukawa TA, Tajika A, Sakata M, Luo Y, Toyomoto R, Horikoshi M, Akechi T, Kawakami N, Nakayama T, Kondo N, Fukuma S, Christensen H, Kessler RC, Cuijpers P & Wason J (2024) RESILIENT (Resilience Enhancement with Smartphone in Lliving ENvironmentTs) Trial: The Statistical Analysis Plan. **medRxiv**, 2024.2003.2016.24304261. 10.1101/2024.03.16.24304261

(Posted 8 March 2024)

RESiLIENT (Resilience Enhancement with Smartphone in Living ENvironmentS) Trial: The Statistical Analysis Plan

March 16, 2024

Hisashi Noma*	PhD	Department of Interdisciplinary Statistical Mathematics, The Institute of Statistical Mathematics, Tokyo, Japan
Toshi A. Furukawa	MD, PhD	Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine/School of Public Health, Kyoto, Japan
Aran Tajika	MD, PhD	
Masatsugu Sakata	PhD	
Yan Luo	MD, PhD	
Rie Toyomoto	RN, MS	
Masaru Horikoshi	PhD	CBT Center, National Center for Neurology and Psychiatry, Kodaira, Japan
Tatsuo Akechi	MD, PhD	Department of Psychiatry and Cognitive-Behavioral Medicine, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan
Norito Kawakami	MD, PhD	Department of Digital Mental Health, Graduate School of Medicine, University of Tokyo, Tokyo, Japan
Takeo Nakayama	MD, PhD	Department of Health Informatics, Kyoto University Graduate School of Medicine/School of Public Health, Kyoto, Japan
Naoki Kondo	MD, PhD	Department of Social Epidemiology, Kyoto University Graduate School of Medicine/School of Public Health, Kyoto, Japan
Shingo Fukuma	MD, PhD	School of Human Health Sciences, Kyoto University Graduate School of Medicine, Kyoto, Japan
Helen Christensen	PhD	Black Dog Institute, University of New South

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

		Wales, Sydney, Australia
Ronald C. Kessler	PhD	Department of Health Care Policy, Harvard Medical School, Boston, USA
Pim Cuijpers	PhD	Department of Clinical, Neuro and Developmental Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
James Wason	PhD	Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne, UK

*Correspondence: Hisashi Noma, PhD
 Department of Interdisciplinary Statistical Mathematics
 The Institute of Statistical Mathematics
 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan
 Email: noma@ism.ac.jp
 Telephone: +81-50-5533-8440

RESiLIENT (Resilience Enhancement with Smartphone in LIving ENvironmentTs) trial

STATISTICAL ANALYSIS PLAN

Version: 1.1

Date: March 16, 2024

Principal Investigator: Toshi A. Furukawa, M.D., Ph.D.
Professor
Departments of Health Promotion and Human Behavior
Kyoto University Graduate School of Medicine / School of
Public Health
Yoshida Konoe-cho, Sakyo-ku
Kyoto 606-8501, Japan
Phone: +81-75-753-9491
Email: furukawa@kuhp.kyoto-u.ac.jp

Trial Statistician: Hisashi Noma, Ph.D.
Professor
Department of Interdisciplinary Statistical Mathematics
The Institute of Statistical Mathematics
10-3, Midori-cho, Tachikawa
Tokyo 190-8562, JAPAN
Phone: +81-50-5533-8440
Email: noma@ism.ac.jp

ABSTRACT

This document gives a statistical analysis plan of the acute intervention effects of the RESiLIENT (Resilience Enhancement with Smartphone in LIving ENvironmenTs) trial, designed as a master protocol including four 2×2 factorial trials to elucidate specific efficacies of five cognitive-behavioural therapy (CBT) skills (cognitive restructuring, behavioural activation, problem-solving, assertion training, and behavior therapy for insomnia) using the internet CBT program. This trial was registered within the UMIN Clinical Trials Registry in Japan (UMIN000047124). The study protocol has previously been published (BMJ Open 2023;13:e067850).

CONTENTS

1. Scope	1
2. Background of RESiLIENT (Resilience Enhancement with Smartphone in Living ENvironmenTs) trial	2
3. General Considerations	8
3.1 Software.....	8
3.2 Significance Level and Reliability Coefficient	8
3.3 Numerical Display of Results.....	8
4. Data Handling and Analysis Sets	9
4.1 Handling of Data	9
4.2 Analysis Sets.....	9
5. Descriptive Analyses	10
5.1 Patient Flow and Reliability of Trial	10
5.2 Baseline Characteristics.....	12
6. Efficacy Analyses	15
6.1 Primary Efficacy Analyses	15
6.2 Secondary Efficacy Analyses	19
6.2.1 Combined Master Protocol Analysis via Pooling the Four Factorial Trials..	19
6.2.2 Efficacy Analyses for Secondary Outcomes	21
6.2.3 Sensitivity analyses	21
7. Safety Analyses	22
8. Revision History	23
References	24

Version 1.1

ABBREVIATIONS AND DEFINITIONS

CBT	cognitive-behavioral therapy
CR	cognitive restructuring
BA	behavioural activation
PS	problem-solving
AT	assertion training
BI	behavior therapy for insomnia
PHQ-9	Patient Health Questionnaire-9
GAD-7	Generalized Anxiety Disorder-7
ISI	Insomnia Severity Index
SWEMWBS	Short Warwick Edinburgh Mental Well-Being Scale
WSAS	Work and Social Adjustment Scale
UWES	Utrecht Work Engagement Scale
HPQ	Health and Work Performance Questionnaire
EQ-5D-5L	EuroQOL-5D-5L
CAQ-3	Client Satisfaction Questionnaire-3
FAS	full analysis set
MMRM	mixed-effects models for repeated measures

1. Scope

This document gives a detailed statistical analysis plan of the acute intervention effects of the "RESiLIENT (Resilience Enhancement with Smartphone in LIving ENvironmenTs)" trial and should be read in conjunction with the trial protocol ¹. Portions of this document are replicated from the study protocol and supplemented with additional detail as appropriate.

2. Background of RESiLIENT (Resilience Enhancement with Smartphone in Living ENvironmentTs) trial

Objective

Cognitive-behavioural therapy (CBT) is an empirically supported intervention to reduce depression severity and to prevent depression onset²⁻⁴. However, it has been traditionally examined as a package of various skills, specific efficacy of which remains largely unexplored. To scale up its application and to combat the rising burden of depression across the globe, we need to assess the specific efficacy of each available CBT skill so that we can inform the most efficacious and efficient intervention package(s). The RESiLIENT (Resilience Enhancement with Smartphone in Living ENvironmentTs) trial¹ was designed as a master protocol including four 2 × 2 factorial trials to elucidate specific efficacies of five CBT skills (cognitive restructuring [CR], behavioural activation [BA], problem-solving [PS], assertion training [AT], and behavior therapy for insomnia [BI]) using the internet CBT (iCBT) program.

Design

The RESiLIENT trial was designed as a master protocol involving four 2×2 factorial trials examining the efficacy of five smartphone CBT skills. The recruitment started on 1 September 2022 and was completed on 21 February 2024, with the last follow-up of the acute intervention phase planned to complete by 19 April 2024.

This SAP focuses on the effects of iCBT skills up to week 6.

Study setting

Participants were recruited in the following four fields. The potential participants will access the internet web page for the trial to apply for the study, receive the explanations about the study from the trial coordinators online remotely, and provide their informed consent by electronic signature.

1. Health insurance societies: There are several types of public health insurance schemes together constituting the universal healthcare in Japan. We will collaborate with two large associations, namely, National Federation of Health Insurance Societies covering employees and their families in large corporations (ca 30 million insured), and Japan Health Insurance Association covering employees and their families of small-to-middle-sized corporations (ca 35 million insured).
2. Business companies and corporations
3. Community and local governments

4. Direct-to-consumer advertisements

Inclusion criteria

1. People of any sex, aged 18 years or older at the time of providing informed consent.
2. In possession of their own smartphone (either iOS or Android).
3. Written informed consent for participation in the trial.
4. Completion of all the baseline questionnaires within 1 week after providing informed consent.
5. Screening PHQ-9 total scores of (1) between 5 and 9, inclusive, or (2) between 10 and 14, inclusive, but not scoring 2 or 3 on its item 9 (suicidality).

Exclusion criteria

1. Cannot read or write Japanese.
2. Receiving treatment from mental health professionals at the time of screening.

Interventions

The participants have been randomised equally to one of the following nine intervention arms or three control arms, stratified by employment status (yes or no) using a computer-generated permuted block randomization table implemented in the app.

The intervention arms include:

1. BA
2. CR
3. PS
4. AT
5. BI
6. BA+CR
7. BA+PS
8. BA+AT
9. BA+BI

BA consists of psychoeducation about pleasurable activities according to the principle of “outside-in” under the catchphrase ‘When your body moves, so does your mind’. It provides a worksheet of a personal experiment to test out a new activity and also a gamified ‘action marathon’ to promote such personal experiments.

CR consists of psychoeducation of the cognitive-behavioural model and cognitive restructuring. The participant learns how to monitor their reactions to situations in terms of feelings, thoughts, body reactions and behaviours by filling in mind maps. The

participant uses these mind maps to apply cognitive restructuring and find alternative thoughts. In order to help the participant broaden their thoughts, CR provides three tools, each of which guides them to alternative thoughts through interactions with the app characters. PS teaches the participants how to break down the issue at hand, to specify a concrete and achievable objective for it, to brainstorm possible solutions, to compare their advantages and disadvantages, and finally to choose the most desirable action and act on it. A worksheet to guide the participants through this process is provided.

AT consists of psychoeducation of assertive communication in contrast to aggressive or passive communication. The participant learns how to express their true feelings and wishes without hurting others or sacrificing themselves. They fill in worksheets to construct appropriate lines in response to their own real-life interactions.

BI teaches the mechanisms of healthy sleep, invites the participant to keep daily sleep records, based on which the participant will start applying sleep restriction and stimulus control techniques, the two proven behavioural skills to increase sleep efficiency.

Each component consists of seven to nine chapters, each divided into two to four lessons as appropriate, as well as worksheets to practice the skill. The participants are expected to complete one chapter per week and can only move on to the next chapter after a week has passed since they started the previous chapter and after they have completed one worksheet. The participants receive weekly encouragement emails, templated but tailored in accordance with each participant's progress by the trial coordinators in the management team (all the coordinators do not have a CBT background and are forbidden to provide any therapeutic contents) and are prompted to fill in PHQ-9 every week.

Controls

In psychotherapy trials, there is no gold standard control condition like the pill placebo in pharmacotherapy trials. Psychotherapy controls must be designed in accordance with the clinical questions of the study as well as the participants' needs and the available resources, and may produce different effect size estimates (ref). In the RESiLIENT trial, we therefore will use three different control conditions with different levels of stringency.

1. Weekly self-checks
2. Health information
3. Delayed intervention

In the weekly self-check arm, the participants will receive weekly encouragement emails and monitor their moods through the weekly PHQ-9 measurements up to week 6 (as in the intervention arms) and then the monthly PHQ-9 measurements thereafter up to week

50. This arm is intended as an attention control to match the attention provided through encouragement emails and self-checks but lacking in active interventions.

In the health information arm, the participants will receive the URLs of websites containing tips for healthy life (physical activities, nutrition and oral health, none of which focuses on mental health) for the initial 3 weeks and will be asked to answer quizzes for comprehension. They will be asked to fill in self-reports at weeks 3 and 6 (without encouragement emails), and then follow-up evaluations thereafter up to week 50. This arm is intended as a placebo intervention.

In the delayed treatment arm, the participants will be placed on the wait list and be asked to fill in self-reports at weeks 3 and 6. They will then be randomised to 1 of the 11 arms (#1 through #11) at week 6, if they so wish then, and will follow respective programs thereafter.

Concomitant interventions

All the participants, in the active and control arms, are free to seek whatever mental health interventions they wish through the 50 weeks. The received professional interventions, either pharmacotherapy or psychotherapy, will be monitored and recorded at weeks 6 and 50.

Primary outcome

The primary outcome is the change in the PHQ-9 from baseline to week 6.

Secondary outcome

The secondary outcomes include:

1. Changes in PHQ-9 from baseline to weeks 1, 2, 3, 4, and 5.
2. Changes in the Generalized Anxiety Disorder-7 (GAD-7) from baseline to weeks 3 and 6.
3. Changes in the Insomnia Severity Index (ISI) from baseline to weeks 3 and 6.
4. Changes in the Short Warwick Edinburgh Mental Well-Being Scale (SWEMWBS) from baseline to weeks 3 and 6.

Anxiety usually coexists with depression and we aim to examine what effects the intervention components, while mainly targeting depression, may have on anxiety. We also aim to monitor broad psychopathology of common mental disorders encompassing depression, anxiety and insomnia as well as positive mental health which are deteriorated in common mental disorders.

5. Changes in the CBT Skills Scale from baseline to week 6.

Table 1. Contents of the interventions and controls.

		Introduction	Chapter 1	Chapter 2	Chapter 3	Chapter 4	Chapter 5	Chapter 6
1	BA	3 lessons for psychoeducation	4 basic lessons for BA	2 advanced lessons for BA	1 advanced lesson for BA	1 advanced lesson for BA	BA in everyday life	Epilogue
2	CR	3 lessons for psychoeducation	4 basic lessons for CR	2 advanced lessons for CR	1 advanced lesson for CR	1 advanced lesson for CR	CR in everyday life	Epilogue
3	PS	3 lessons for psychoeducation	4 basic lessons for PS	2 advanced lessons for PS	2 advanced lessons for PS	1 advanced lesson for PS	PS in everyday life	Epilogue
4	AT	3 lessons for psychoeducation	5 basic lessons for AT	2 advanced lessons for AT	1 advanced lesson for AT	1 advanced lesson for AT	AT in everyday life	Epilogue
5	BI	3 lessons for psychoeducation	5 basic lessons for BI	1 advanced lesson for BI	1 advanced lesson for BI	1 advanced lesson for BI	BI in everyday life	Epilogue
6	BA+CR	3 lessons for psychoeducation	4 basic lessons for BA	4 advanced lessons for BA	4 basic lessons for CR	4 advanced lessons for CR	BA&CR in everyday life	Epilogue
7	BA+PS	3 lessons for psychoeducation	4 basic lessons for BA	4 advanced lessons for BA	4 basic lessons for PS	5 advanced lessons for PS	BA&PS in everyday life	Epilogue
8	BA+AT	3 lessons for psychoeducation	4 basic lessons for BA	4 advanced lessons for BA	5 basic lessons for AT	4 advanced lessons for AT	BA&AT in everyday life	Epilogue
9	BA+BI	3 lessons for psychoeducation	4 basic lessons for BA	4 advanced lessons for BA	5 basic lessons for BI	3 advanced lessons for BI	BA&BI in everyday life	Epilogue
10	Self-check	1 lesson for psychoeducation	-	-	-	-	-	Epilogue
11	Health information	1 lesson for psychoeducation	Physical exercise	Nutrition	Oral health	-	-	Epilogue
12	Delayed treatment	1 lesson for psychoeducation	-	-	-	-	-	Epilogue

6. Changes in the Work and Social Adjustment Scale (WSAS), the Utrecht Work Engagement Scale (UWES), the Health and Work Performance Questionnaire (HPQ) and the EuroQOL-5D-5L (EQ-5D-5L) from baseline to week 6.

7. The Client Satisfaction Questionnaire-3 (CSQ-3), adherence to smartphone CBT, co-interventions, safety information at week 6.

Table 2. Measurement schedule.

	Questionnaire Part I at landing page	Questionnaire Part II on smartphone app (Week 0)	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6
Allowances (days)	-		+6	+6	+6	+6	+6	+14
PHQ-9	○	○	△	△	○	△	△	○
GAD-7		○			○			○
ISI		○			○			○
SWEMWBS		○			○			○
CIDI		○						
Baseline characteristics		○						
ASC-Motivation		○			○			
ICT usage		○						
ACE		○						
Short Big 5		○						
ASC-Life difficulties		○						
ASC-Social support		○						
CBT skills	○							○
WSAS		○						○
UWES		□						□
HPQ		□						□
EQ-5D-5L		○						○
Co-interventions								○
CSQ-3								○

○: to be administered to all participants, □: only to those currently employed, and △: all participants except for Groups 11 and 12.

ACE: Adverse Childhood Experiences, ASC: Assessment of Signal Cases, CIDI: Composite International Diagnostic Interview, CSQ-3: Customer Satisfaction Questionnaire-3, EQ-5D-5L: EuroQOL-5D-5L, GAD-7: General Anxiety Disorder-7, HPQ: Health and Work Performance Questionnaire, ICT: Information and communication technologies, ISI: Insomnia Severity Index, PHQ-9: Patient Health Questionnaire-9, SWEMWBS: Short Warwick Edinburgh Mental Well-Being Scale, UWES: Utrecht Work Engagement Scale, WSAS: Work and Social Adjustment Scale

3. General Considerations

3.1 Software

We will use Statistical Analysis System (SAS) Ver. 9.4. (SAS Institute, Inc., Cary, NC) and R Ver. 4.3.2 (R Foundation for Statistical Computing, Vienna, Austria) for all statistical analyses.

3.2 Significance Level and Reliability Coefficient

Unless otherwise specified, estimates of treatment effects will be presented with 95% confidence intervals. P-values will be 2-tailed. A P-value <0.05 will be considered statistically significant. No correction for multiple tests will be made.

3.3 Numerical Display of Results

As a general rule, the analysis results will be indicated by the number of digits shown below. The data will be rounded to the nearest whole number.

Min, Max	Number of significant digits of data
Mean, Quantile, Median	Number of significant digits of data
Standard Deviation (SD)	Number of significant digits of data plus one
Percent (%)	Integer value
P-value	Rounded to the nearest fourth decimal place, with three decimal places, unless $P < 0.0001$

4. Data Handling and Analysis Sets

4.1 Handling of Data

The handling of individual cases and data will follow the results of examinations by the data safety and monitoring committee.

4.2 Analysis Sets

The analysis will follow the intention-to-treat principle and include all the participants randomized.

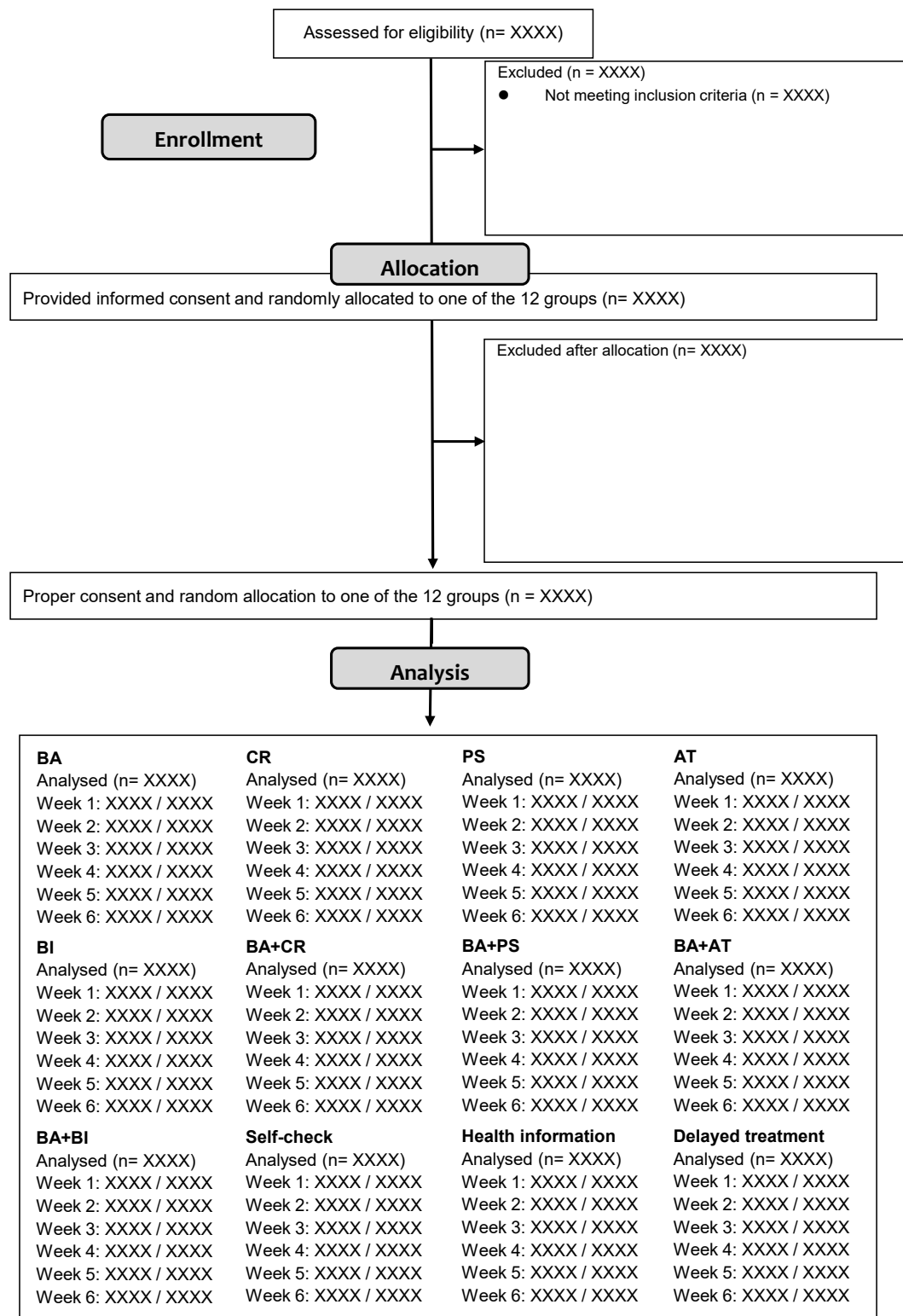
5. Descriptive Analyses

5.1 Patient Flow and Reliability of Trial

We will make a patient flow diagram as shown in Figure 1, following the CONSORT statement ⁵, to describe the progress of trial (enrollment, intervention, follow-up, and data analysis).

We will report the total number of patients assessed for eligibility, enrolled, and withdrawn/dropping out and their reasons, as well as frequency of inclusion/exclusion from the analysis and the reasons for exclusion.

Figure 1. CONSORT diagram



PHQ-9: Patient Health Questionnaire-9 IC: informed consent

5.2 Baseline Characteristics

We will summarize the participants' baseline demographic, psychosocial and clinical characteristics in the ITT sample as shown in Table 3.

Continuous variables will be summarized using the number of observations (n), mean, standard deviation, median, minimum, and maximum. The number of missing observations will be reported.

For categorical variables, the frequencies and percentages (based on the non-missing sample size) as observed will be reported. The number of missing observations will also be reported.

Table 3. Baseline characteristics of total participants by each component (N= XXXX).

	Components																					
	Total		BA		CR		PS		AT		BI											
			Presence	Absence	Presence	Absence	Presence	Absence	Presence	Absence	Presence	Absence										
N	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%		
Demographic variables																						
Age (Mean, SD)																						
Sex (Female)																						
Marital status																						
Number of cohabitants (Mean, SD)																						
Socioeconomic status of the residential area																						
Education																						
Employment																						
Psychosocial variables	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD		
Personality trait (Big five)																						
Neuroticism																						
Extraversion																						
Openness																						
Conscientiousness																						
Agreeableness																						
Adverse childhood experiences (ACE)																						
Assessment of signal cases (ASC)-Life difficulties																						
ASC-Social support																						
ASC-Motivation																						
CBT skills																						
BA skills																						
CR skills																						
PS skills																						
AT skills																						
BI skills																						

Readiness for smartphone CBT																								
Familiarity with smartphone usage																								
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Clinical variables																								
Physical conditions and illnesses, including past and current treatments																								
Physical disability																								
Mental health conditions, including past and current treatments																								
Alcohol use and CAGE questionnaire																								
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
PHQ-9																								
GAD-7																								
ISI																								
WSAS																								
HPQ																								
UWES																								
EQ-5D-5L																								
SWEMWBS																								

6. Efficacy Analyses

6.1 Primary Efficacy Analyses

In the preplanned primary analysis, each of the four 2*2 factorial trials is to be analyzed independently. The four trials will use the BA arm (intervention arm #1 in Table 1) and the control arm (no treatment, intervention arm #12 in Table 1) in common. Intervention arms #10 and #11, with increasing stringency as control, will be used in sensitivity analyses, i.e., we will perform additional analyzes changing the control arm to the arm #10 or #11 respectively. For the fourth trial involving BA and BI, the participants who work in three shifts will be excluded, as shift workers cannot be expected to make optimal use of BI.

Table 4. Four 2×2 factorial trials.

	Randomization for BA	
Randomization for CR	BA+, CR+	BA-, CR+
	BA+, CR-	BA-, CR-

	Randomization for BA	
Randomization for PS	BA+, PS+	BA-, PS+
	BA+, PS-	BA-, PS-

	Randomization for BA	
Randomization for AT	BA+, AT+	BA-, AT+
	BA+, AT-	BA-, AT-

	Randomization for BA	
Randomization for BI	BA+, BI+	BA-, BI+
	BA+, BI-	BA-, BI-

We will use the mixed-effects models for repeated-measures (MMRM)^{6,7} to estimate the mean difference in change scores on the PHQ-9 for each component with adjusting informative missing of the outcomes. The model will include fixed effects of treatment, visit (as categorical), and treatment-by-visit interactions, adjusted for baseline PHQ-9

scores, employment status, age and sex. Each of the experimental factors will be coded at two levels (presence coded as 1 and absence coded as 0). The covariance matrix structure of outcome variables will be set to unstructured; if the estimating algorithm does not converge, we will use the matrix structure in the following order: Toeplitz, heterogeneous compound symmetry, autoregressive (1), compound symmetry, and variance components models. The standard error estimates and degree-of-freedom will be adjusted by the Kenward-Roger method ^{8,9}. We will also estimate least squares mean change scores of the PHQ-9 at week 1, 2, 3, 4, and 5 based on the estimated models of the MMRM analysis. The results will be presented as Table 5. For the effect of BA, which will be estimated in all of the four 2*2 trials, we will present the effect size estimate of the first trial, i.e. the 2*2 trial involving BA and CR, in Table 5 as the primary result because BA and CR are the most typical ingredients across broadly conceived CBT interventions ¹⁰. However, if the effect size estimates of BA are considerably heterogeneous among the four trials, we will present all of these results, and will discuss the sources of the heterogeneity.

No adjustment for multiple testing will be applied in examination of statistical significance of the main effects and the conventional threshold for statistical significance ($p < 0.05$, two sided) will be used, because conventionally, the multiple hypotheses have been tested independently in these trial designs assuming as if the tested interventions could have been assessed in separate trials ^{11,12}. We will perform another MMRM analysis involving a two-way interaction term of each two components to assess the interaction between the treatments as a sensitivity analysis (Table 6). Previous research has not been suggestive of interactions among iCBT components ^{10,13}, but when we identify a strong interaction, we will interpret the results considering the interaction. We also perform linear regression analyses using only the week 6 data as sensitivity analyzes.

Table 5. Analysis of main effects using the mixed-model repeated measures analysis (MMRM) for PHQ-9 score

Component	Estimated least squares mean change scores of the PHQ-9 [†]							Estimated differences of change scores (95% CI)	Effect size for Baseline-Week 6 change scores [‡] (95% CI)	Effect size for differences of change scores and their p-values [‡] (95% CI)	P-value
	N	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6				
CR											
Present											
Absent											
PS											
Present											
Absent											
AT											
Present											
Absent											
BI											
Present											
Absent											
BA*											
Present											
Absent											

[†] Minus value indicates depression reduction.

[‡] We will calculate effect sizes (standardised mean differences) using the observed baseline SD for within-group change score and the observed week 6 SD for group differences.

* The effect size estimate of BA obtained from the first 2*2 trial involving BA and CR.

Table 6. Analysis of main and interaction effects using the mixed-model repeated measures analysis (MMRM) for PHQ-9 score

Arm	Estimated least squares mean change scores of the PHQ-9 [†]							Estimated differences of change scores (95% CI)	Effect size for Baseline-Week 6 change scores [‡] (95% CI)	Effect size for differences of change scores and their p-values [‡] (95% CI)	P-value
	N	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6				
1st trial											
CR + BA											
CR											
BA											
No treatment											
2nd trial											
PS + BA											
PS											
BA											
No treatment											
3rd trial											
AT + BA											
AT											
BA											
No treatment											
4th trial											
BI + BA											
BI											
BA											
No treatment											

[†] Minus value indicates depression reduction.

[‡] We will calculate effect sizes (standardised mean differences) using the observed baseline SD for within-group change score and the observed week 6 SD for group differences.

6.2 Secondary Efficacy Analyses

6.2.1 Combined Master Protocol Analysis via Pooling the Four Factorial Trials

To assess the comparative efficacy of the five CBT skills and their combinations, we will additionally perform a combined master protocol analysis via pooling the four factorial trials along with all the control arms. We will perform the MMRM analysis^{6,7} for the pooled dataset, and will provide comparative effect size information for all the CBT skills and their combinations evaluated in the four trials. We will perform two combined master protocol analyses involving (1) main effects of 5 CBT skills and (2) main effects and 2 way interaction terms of 5 CBT skills, while setting the control arm to the arm #12. Control arms #10 and #11, whose outcomes are assumed to be different, will also be included and compared comprehensively. The results will be presented as in Table 7. The combined master protocol analysis allows direct comparisons among all active components, thus informing (1) what component is the most efficacious and (2) whether the effects of individual CBT skills are different (via significance tests). In this combined master protocol analysis, we will include shift-workers, although such inclusion may lead to underestimating the effect of BI on depression, to maintain the statistical power for the analysis and make the results as comparable as possible to those in the primary 2*2 trial analyses.

Table 7. Combined master protocol analysis of main and interaction effects using the mixed-model repeated measures analysis (MMRM) for PHQ-9 score

Component	Estimated least squares mean change scores of the PHQ-9 [†]							Estimated differences of change scores (95% CI)	Effect size for Baseline-Week 6 change scores [‡] (95% CI)	Effect size for differences of change scores and their p-values [‡] (95% CI)	P-value
	N	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6				
BA											
CR											
PS											
AT											
BI											
BA + CR											
BA + PS											
BA + AT											
BA + BI											
Self-check											
Health information											
Delayed treatment [reference]											

[†] Minus value indicates depression reduction.

[‡] We will calculate effect sizes (standardised mean differences) using the observed baseline SD for within-group change score and the observed week 6 SD for group differences.

6.2.2 Efficacy Analyses for Secondary Outcomes

We will perform MMRM analyzes for the repeated measurement outcomes of GAD-7, ISI and SWEMWBS until week 6. The same modelling strategy with the analysis of the primary outcome in Section 6.1 will be adopted.

6.2.3 Sensitivity analyses

We will evaluate the robustness of the foregoing analyses by conducting the following sensitivity analyses.

For the 2*2 trials

- (1) Using the control arm #10 (self-check) or #11 (health information) instead of #12 (delayed treatment) in each of the 2*2 trial
- (2) We will examine interactions among the skills as a sensitivity analysis. In the examination of interactions among components, we will also perform linear regression analyses using only the week 6 data for each of the 2*2 trials.
- (3) For the factorial trial examining BA and BI, subgrouping participants to those who scored 8 or more vs those who scored 7 or less on the ISI at baseline. Here the primary outcome is the PHQ-9, and this subgroup analysis will confirm if the effect of BI is similar regardless of the baseline insomnia severity.

For the secondary outcomes

- (4) For the secondary outcome of the GAD-7, limiting participants to those who scored 5 or more on the GAD-7 at baseline (i.e. those who presented with at least some anxiety symptoms at baseline). This will mitigate the floor effect on the GAD-7.
- (5) For the secondary outcome of the ISI, limiting participants to those who scored 8 or more on the ISI at baseline (i.e. those who presented with at least some insomnia symptoms at baseline). This will mitigate the floor effect on the ISI.
- (6) For the secondary outcome of the SWEMWEBS in the four 2*2 factorial trials, limiting participants to those who scored 20 or less on the SWEMWBS at baseline (i.e. those who showed reduced well-being at baseline). This will mitigate the ceiling effect on the SWEMWEBS.

7. Safety Analyses

As this trial is a non-invasive interventional study and no adverse events due to the intervention are expected. We will only descriptively report serious adverse events as reported during the 6-week intervention in each arm.

8. Revision History

Version	Date	Changes
1.1	03-16-2024	The first version.

References

1. Furukawa TA, Tajika A, Sakata M, et al. Four 2x2 factorial trials of smartphone CBT to reduce subthreshold depression and to prevent new depressive episodes among adults in the community-RESiLIENT trial (Resilience Enhancement with Smartphone in LIVING ENvironmentTs): a master protocol. *BMJ Open*. 2023;13(2):e067850.
2. Cuijpers P, Pineda BS, Quero S, et al. Psychological interventions to prevent the onset of depressive disorders: A meta-analysis of randomized controlled trials. *Clin Psychol Rev*. 2021;83:101955.
3. Cuijpers P, Quero S, Noma H, et al. Psychotherapies for depression: a network meta-analysis covering efficacy, acceptability and long-term outcomes of all main treatment types. *World Psychiatry*. 2021;20(2):283-293.
4. Furukawa TA, Shinohara K, Sahker E, et al. Initial treatment choices to achieve sustained response in major depression: a systematic review and network meta-analysis. *World Psychiatry*. 2021;20(3):387-396.
5. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c332.
6. Mallinckrodt CH, Clark WS, David SR. Accounting for dropout bias using mixed-effects models. *J Biopharm Statist*. 2001;11(1-2):9-21.
7. Mallinckrodt CH, Kaiser CJ, Watkin JG, Molenberghs G, Carroll RJ. The effect of correlation structure on treatment contrasts estimated from incomplete clinical trial data with likelihood-based repeated measures compared with last observation carried forward ANOVA. *Clin Trials*. 2004;1(6):477-89.
8. Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*. 1997;53(3):983-997.
9. Kenward MG, Roger JH. An improved approximation to the precision of fixed effects from restricted maximum likelihood. *Comput Stat Data Anal*. 2009;53(7):2583-2595.
10. Furukawa TA, Suganuma A, Ostinelli EG, et al. Dismantling, optimising, and personalising internet cognitive behavioural therapy for depression: a systematic review and component network meta-analysis using individual participant data. *Lancet Psychiatry*. 2021;8(6):500-511.
11. Balakrishnan N. *Methods and Applications of Statistics in Clinical Trials: Concepts, Principles, Trials, and Designs*. Wiley, 2014.
12. Fleiss JL. *The Design and Analysis of Clinical Experiments*. John Wiley & Sons; 1999.
13. Sakata M, Toyomoto R, Yoshida K, et al. Components of smartphone cognitive-behavioural therapy for subthreshold depression among 1093 university students: a factorial trial. *Evid Based Ment Health*. 2022;25(e1):e18-e25.

3. Supplementary Data

Supplementary Data. Table 1. Comparison of baseline demographic, psychosocial and clinical characteristics of the participants according to the presence or absence of each intervention

	Components									
	BA*		CR		PS		AT		BI	
	Present	Absent	Present	Absent	Present	Absent	Present	Absent	Present	Absent
N	656	656	657	655	657	655	656	655	655	655
Demographic variables										
Age	43.8 (10.8)	43.4 (10.5)	43.3 (10.6)	43.9 (10.7)	43.1 (10.7)	43.9 (10.7)	43.1 (10.7)	43.9 (10.7)	42.4 (10.8)	43.9 (10.7)
Gender										
Male	332 (51%)	320 (49%)	322 (49%)	330 (50%)	334 (51%)	330 (50%)	346 (53%)	330 (50%)	339 (52%)	330 (50%)
Female	321 (49%)	334 (51%)	333 (51%)	322 (49%)	320 (49%)	322 (49%)	309 (47%)	322 (49%)	311 (47%)	322 (49%)
Others	3 (0.5%)	2 (0.3%)	2 (0.3%)	3 (0.5%)	3 (0.5%)	3 (0.5%)	1 (0.2%)	3 (0.5%)	5 (0.8%)	3 (0.5%)
Marital status										
Single	147 (22%)	147 (22%)	140 (21%)	154 (24%)	165 (25%)	154 (24%)	158 (24%)	154 (24%)	176 (27%)	154 (24%)
Married	461 (70%)	466 (71%)	472 (72%)	455 (69%)	451 (69%)	455 (69%)	460 (70%)	455 (69%)	443 (68%)	455 (69%)
Divorced/Widowed	48 (7%)	43 (7%)	45 (7%)	46 (7%)	41 (6%)	46 (7%)	38 (6%)	46 (7%)	36 (5%)	46 (7%)
Number of cohabitants including self	2.75 (1.31)	2.73 (1.32)	2.76 (1.30)	2.72 (1.34)	2.68 (1.37)	2.72 (1.34)	2.70 (1.30)	2.72 (1.34)	2.61 (1.33)	2.72 (1.34)
Area Deprivation Index (Range: 0-15.9; the higher, the more deprived)	5.57 (0.99)	5.60 (1.09)	5.56 (1.04)	5.60 (1.05)	5.71 (1.10)	5.60 (1.05)	5.61 (1.07)	5.60 (1.05)	5.67 (0.99)	5.60 (1.05)
Education, university or higher	470 (72%)	475 (72%)	486 (74%)	459 (70%)	474 (72%)	459 (70%)	465 (71%)	459 (70%)	475 (73%)	459 (70%)
Employed	603 (92%)	604 (92%)	604 (92%)	603 (92%)	604 (92%)	603 (92%)	604 (92%)	603 (92%)	603 (92%)	603 (92%)
Psychosocial variables										
Big 5 personality traits (Range: 1-5)										
Neuroticism	3.64 (0.77)	3.65 (0.76)	3.69 (0.75)	3.61 (0.78)	3.65 (0.79)	3.61 (0.78)	3.66 (0.75)	3.61 (0.78)	3.69 (0.79)	3.61 (0.78)
Extraversion	2.98 (0.93)	3.00 (0.89)	3.02 (0.92)	2.95 (0.89)	2.99 (0.93)	2.95 (0.89)	2.99 (0.91)	2.95 (0.89)	2.96 (0.89)	2.95 (0.89)
Openness	3.16 (0.77)	3.17 (0.76)	3.16 (0.77)	3.17 (0.76)	3.13 (0.74)	3.17 (0.76)	3.18 (0.75)	3.17 (0.76)	3.12 (0.76)	3.17 (0.76)
Conscientiousness	3.01 (0.80)	3.03 (0.76)	3.02 (0.79)	3.02 (0.77)	3.04 (0.82)	3.02 (0.77)	3.03 (0.77)	3.02 (0.77)	3.05 (0.75)	3.02 (0.77)
Agreeableness	3.20 (0.85)	3.17 (0.89)	3.17 (0.86)	3.19 (0.88)	3.15 (0.86)	3.19 (0.88)	3.15 (0.83)	3.19 (0.88)	3.23 (0.85)	3.19 (0.88)
Adverse Childhood Experiences (≥4 points)	56 (9%)	66 (10%)	53 (8%)	69 (11%)	65 (10%)	69 (11%)	65 (10%)	69 (11%)	69 (11%)	69 (11%)
Assessment of Signal Cases (Range:0-3)										
Life difficulties	0.94 (0.59)	0.94 (0.56)	0.93 (0.58)	0.94 (0.57)	0.95 (0.55)	0.94 (0.57)	0.96 (0.56)	0.94 (0.57)	0.94 (0.57)	0.94 (0.57)
Social support	1.33 (0.75)	1.34 (0.75)	1.33 (0.75)	1.34 (0.74)	1.31 (0.73)	1.34 (0.74)	1.28 (0.77)	1.34 (0.74)	1.29 (0.73)	1.34 (0.74)
Motivation	2.20 (0.39)	2.22 (0.38)	2.21 (0.39)	2.21 (0.38)	2.18 (0.39)	2.21 (0.38)	2.23 (0.40)	2.21 (0.38)	2.21 (0.39)	2.21 (0.38)
CBT skills (Range: 0-3)										
BA skills	1.42 (0.46)	1.42 (0.46)	1.42 (0.46)	1.43 (0.46)	1.44 (0.48)	1.43 (0.46)	1.42 (0.45)	1.43 (0.46)	1.41 (0.44)	1.43 (0.46)
CR skills	1.37 (0.57)	1.29 (0.61)	1.32 (0.59)	1.35 (0.59)	1.34 (0.58)	1.35 (0.59)	1.35 (0.60)	1.35 (0.59)	1.36 (0.58)	1.35 (0.59)
PS skills	1.76 (0.56)	1.73 (0.58)	1.73 (0.57)	1.77 (0.56)	1.75 (0.56)	1.77 (0.56)	1.74 (0.57)	1.77 (0.56)	1.76 (0.55)	1.77 (0.56)
AT skills	1.22 (0.53)	1.24 (0.51)	1.22 (0.52)	1.24 (0.53)	1.19 (0.53)	1.24 (0.53)	1.20 (0.52)	1.24 (0.53)	1.21 (0.51)	1.24 (0.53)
BI skills	1.90 (0.57)	1.89 (0.55)	1.89 (0.56)	1.89 (0.56)	1.94 (0.56)	1.89 (0.56)	1.91 (0.54)	1.89 (0.56)	1.91 (0.55)	1.89 (0.56)
Familiarity with smartphone (Range: 0-3)	1.98 (0.68)	1.97 (0.67)	1.99 (0.67)	1.97 (0.68)	1.93 (0.65)	1.97 (0.68)	2.00 (0.69)	1.97 (0.68)	1.97 (0.69)	1.97 (0.68)
Clinical variables										
PHQ-9 (Range: 0-27)	8.02 (2.72)	8.15 (2.69)	8.14 (2.82)	8.02 (2.58)	8.04 (2.79)	8.02 (2.58)	8.18 (2.84)	8.02 (2.58)	8.00 (2.66)	8.02 (2.58)
GAD-7 (Range: 0-21)	5.79 (3.43)	5.75 (3.44)	5.84 (3.43)	5.70 (3.45)	5.70 (3.30)	5.70 (3.45)	5.87 (3.43)	5.70 (3.45)	5.83 (3.36)	5.70 (3.45)
ISI (Range: 0-28)	10.22 (4.56)	10.02 (4.52)	10.01 (4.54)	10.23 (4.54)	10.23 (4.54)	10.15 (4.47)	10.23 (4.54)	10.15 (4.47)	9.85 (4.28)	10.24 (4.55)
WSAS (Range: 0-8)	1.74 (1.70)	1.80 (1.76)	1.75 (1.72)	1.79 (1.74)	1.82 (1.77)	1.79 (1.74)	1.80 (1.78)	1.79 (1.74)	1.72 (1.65)	1.79 (1.74)
HPQ (Range: 0-10)	6.21 (1.99)	5.95 (2.07)	6.13 (2.01)	6.03 (2.05)	5.99 (2.08)	6.03 (2.05)	6.04 (2.16)	6.03 (2.05)	5.93 (2.03)	6.03 (2.05)
UWES (Range: 0-6)	2.79 (1.37)	2.83 (1.27)	2.75 (1.34)	2.87 (1.29)	2.92 (1.29)	2.87 (1.29)	2.79 (1.31)	2.87 (1.29)	2.82 (1.25)	2.87 (1.29)
EQ-5D-5L utility (Range: -0.59 to 1)	0.90 (0.09)	0.91 (0.09)	0.90 (0.09)	0.91 (0.09)	0.91 (0.10)	0.91 (0.09)	0.90 (0.09)	0.91 (0.09)	0.91 (0.09)	0.91 (0.09)
SWEMWBS (Range: 7-35)	22.19 (3.52)	22.04 (3.64)	22.07 (3.59)	22.17 (3.57)	21.99 (3.62)	22.17 (3.57)	22.00 (3.77)	22.17 (3.57)	21.88 (3.78)	22.17 (3.57)
CAGE (≥2 points)	117 (18%)	105 (16%)	111 (17%)	111 (17%)	102 (16%)	111 (17%)	109 (17%)	111 (17%)	102 (16%)	111 (17%)
Past or current treatments for mental health	217 (33%)	183 (28%)	205 (31%)	195 (30%)	196 (30%)	195 (30%)	177 (27%)	195 (30%)	198 (30%)	195 (30%)
Hypertension	112 (17%)	91 (14%)	108 (16%)	95 (15%)	90 (14%)	95 (15%)	103 (16%)	95 (15%)	104 (16%)	95 (15%)
Diabetes mellitus	27 (4%)	15 (2%)	20 (3%)	22 (3%)	22 (3%)	22 (3%)	19 (3%)	22 (3%)	11 (2%)	22 (3%)
Chronic pain	106 (16%)	111 (17%)	109 (17%)	108 (16%)	115 (18%)	108 (16%)	107 (16%)	108 (16%)	93 (14%)	108 (16%)
Any physical disease	223 (34%)	229 (35%)	228 (35%)	224 (34%)	221 (34%)	224 (34%)	215 (33%)	224 (34%)	201 (31%)	224 (34%)
Physical disability	19 (3%)	24 (4%)	15 (2%)	28 (4%)	22 (3%)	28 (4%)	29 (4%)	28 (4%)	19 (3%)	28 (4%)

Categorical variables are presented as the number of participants with percentage in parentheses. Continuous variables are presented as the mean with SD in parentheses.

AT: Assertion training, BA: Behavioural activation, BI: Behaviour therapy for insomnia, CR: Cognitive restructuring, EQ-5D-5L: EuroQOL-5D-5L, GAD-7: Generalized Anxiety Disorder-7, HPQ: Health and Work Performance Questionnaire, ISI: Insomnia Severity Index, PHQ-9: Patient Health Questionnaire-9, PS: Problem solving, SWEMWBS: Short Warwick Edinburgh Mental Well-Being Scale, UWES: Utrecht Work Engagement Scale, WSAS: Work and Social Adjustment Scale,

* This contrast is obtained from the first 2*2 trial involving BA and CR.

Supplementary Data. Table 2. Analysis of main and interaction effects using the mixed-model repeated measures analysis (MMRM) for GAD-7 score at week 6

Arm	N	Estimated differences of change scores at Week 6 (95% CI)	Effect size for differences of change scores at Week 6 [‡] (95% CI)	P-value (Main effect)	P-value (Interaction of two components)
Trial 1					
BA + CR	329	-1.34 (-1.77 to -0.90)	-0.41 (-0.54 to -0.28)	1.9×10^{-9}	6.2×10^{-6}
BA	327	-1.67 (-2.10 to -1.24)	-0.51 (-0.64 to -0.38)	7.5×10^{-14}	
CR	328	-1.09 (-1.52 to -0.65)	-0.33 (-0.46 to -0.20)	8.7×10^{-7}	
Delayed treatment	328	—	—	—	
Trial 2					
BA + PS	329	-1.40 (-1.83 to -0.96)	-0.42 (-0.55 to -0.29)	3.1×10^{-10}	3.8×10^{-4}
BA	327	-1.66 (-2.09 to -1.22)	-0.50 (-0.63 to -0.37)	1.5×10^{-13}	
PS	328	-0.86 (-1.30 to -0.43)	-0.26 (-0.39 to -0.13)	1.1×10^{-4}	
Delayed treatment	328	—	—	—	
Trial 3					
BA + AT	328	-1.27 (-1.7 to -0.84)	-0.38 (-0.51 to -0.25)	7.3×10^{-9}	1.1×10^{-5}
BA	327	-1.65 (-2.08 to -1.22)	-0.49 (-0.62 to -0.36)	9.4×10^{-14}	
AT	328	-0.99 (-1.42 to -0.56)	-0.30 (-0.42 to -0.17)	6.2×10^{-6}	
Delayed treatment	328	—	—	—	
Trial 4					
BA + BI	311	-1.24 (-1.68 to -0.80)	-0.37 (-0.50 to -0.24)	3.1×10^{-8}	3.2×10^{-10}
BA	314	-1.66 (-2.10 to -1.22)	-0.49 (-0.63 to -0.36)	2.4×10^{-13}	
BI	313	-1.59 (-2.02 to -1.15)	-0.47 (-0.60 to -0.34)	1.8×10^{-12}	
Delayed treatment	313	—	—	—	

The least squares mean changes and their differences of the PHQ-9 scores were estimated with the mixed-effects models for repeated-measures (MMRM). The regression models included fixed effects of treatment, visit (as categorical), and treatment-by-visit interactions, adjusted for baseline PHQ-9 scores, employment status, age and gender. The interaction P-values are calculated from the MMRM models involving interaction terms of the corresponding treatments. P-values are two-sided and no adjustments are made for multiple comparisons.

† Minus value indicates anxiety reduction. () indicates 95% confidence intervals.

‡ Effect sizes (standardised mean differences) were calculated using the observed week 6 SD for group differences.

AT: Assertion training, BA: Behavioural activation, BI: Behaviour therapy for insomnia, CR: Cognitive restructuring, GAD-7: General Anxiety Disorder-7, PS: Problem solving

Supplementary Data. Table 3. Analysis of main and interaction effects using the mixed-model repeated measures analysis (MMRM) for ISI score at week 6

Arm	N	Estimated differences of change scores at Week 6 (95% CI)	Effect size for differences of change scores at Week 6 [‡] (95% CI)	P-value (Main effect)	P-value (Interaction of two components)
Trial 1					
BA + CR	329	-1.16 (-1.71 to -0.61)	-0.25 (-0.37 to -0.13)	3.7×10 ⁻⁵	0.019
BA	327	-1.19 (-1.74 to -0.64)	-0.25 (-0.37 to -0.14)	2.5×10 ⁻⁵	
CR	328	-0.90 (-1.45 to -0.36)	-0.19 (-0.31 to -0.08)	0.001	
Delayed treatment	328	—	—	—	
Trial 2					
BA + PS	329	-1.15 (-1.72 to -0.58)	-0.24 (-0.36 to -0.12)	7.2×10 ⁻⁵	0.069
BA	327	-1.20 (-1.77 to -0.63)	-0.25 (-0.37 to -0.13)	3.8×10 ⁻⁵	
PS	328	-0.70 (-1.27 to -0.13)	-0.15 (-0.26 to -0.03)	0.016	
Delayed treatment	328	—	—	—	
Trial 3					
BA + AT	328	-1.17 (-1.75 to -0.59)	-0.24 (-0.36 to -0.12)	8.6×10 ⁻⁵	0.019
BA	327	-1.18 (-1.77 to -0.60)	-0.24 (-0.37 to -0.12)	7.8×10 ⁻⁵	
AT	328	-0.98 (-1.56 to -0.40)	-0.20 (-0.32 to -0.08)	0.001	
Delayed treatment	328	—	—	—	
Trial 4					
BA + BI	311	-2.03 (-2.59 to -1.48)	-0.45 (-0.58 to -0.33)	1.2×10 ⁻¹²	4.2×10 ⁻⁴
BA	314	-1.15 (-1.71 to -0.59)	-0.26 (-0.38 to -0.13)	5.4×10 ⁻⁵	
BI	313	-2.30 (-2.86 to -1.75)	-0.51 (-0.64 to -0.39)	1.0×10 ⁻¹⁵	
Delayed treatment	313	—	—	—	

The least squares mean changes and their differences of the PHQ-9 scores were estimated with the mixed-effects models for repeated-measures (MMRM). The regression models included fixed effects of treatment, visit (as categorical), and treatment-by-visit interactions, adjusted for baseline PHQ-9 scores, employment status, age and gender. The interaction P-values are calculated from the MMRM models involving interaction terms of the corresponding treatments. P-values are two-sided and no adjustments are made for multiple comparisons.

† Minus value indicates insomnia reduction. () indicates 95% confidence intervals.

‡ Effect sizes (standardised mean differences) were calculated using the observed week 6 SD for group differences.

AT: Assertion training, BA: Behavioural activation, BI: Behaviour therapy for insomnia, CR: Cognitive restructuring, ISI: Insomnia Severity Scale, PS: Problem solving

Supplementary Data. Table 4. Analysis of main and interaction effects using the mixed-model repeated measures analysis (MMRM) for SWEMWBS score at week 6

Component	N	Estimated differences of change scores at Week 6 (95% CI)	Effect size for differences of change scores at Week 6 [‡] (95% CI)	P-value (Main effect)	P-value (Interaction of two components)
Trial 1					
BA + CR	329	0.68 (0.17 to 1.18)	0.16 (0.04 to 0.28)	0.009	1.0×10 ⁻³
BA	327	1.12 (0.62 to 1.63)	0.27 (0.15 to 0.39)	1.5×10 ⁻⁵	
CR	328	0.76 (0.26 to 1.27)	0.18 (0.06 to 0.30)	0.003	
Delayed treatment	328	—	—	—	
Trial 2					
BA + PS	329	0.77 (0.25 to 1.29)	0.18 (0.06 to 0.30)	0.004	0.076
BA	327	1.12 (0.60 to 1.64)	0.26 (0.14 to 0.38)	2.9×10 ⁻⁵	
PS	328	0.32 (-0.20 to 0.85)	0.08 (-0.05 to 0.20)	0.225	
Delayed treatment	328	—	—	—	
Trial 3					
BA + AT	328	0.99 (0.49 to 1.50)	0.23 (0.11 to 0.35)	1.1×10 ⁻⁴	0.020
BA	327	1.12 (0.62 to 1.63)	0.26 (0.14 to 0.38)	1.4×10 ⁻⁵	
AT	328	0.72 (0.22 to 1.22)	0.17 (0.05 to 0.28)	0.005	
Delayed treatment	328	—	—	—	
Trial 4					
BA + BI	311	0.78 (0.28 to 1.29)	0.18 (0.06 to 0.29)	0.002	1.1×10 ⁻⁵
BA	314	1.20 (0.69 to 1.70)	0.27 (0.16 to 0.39)	4.3×10 ⁻⁶	
BI	313	1.20 (0.70 to 1.71)	0.27 (0.16 to 0.39)	3.4×10 ⁻⁶	
Delayed treatment	313	—	—	—	

The least squares mean changes and their differences of the PHQ-9 scores were estimated with the mixed-effects models for repeated-measures (MMRM). The regression models included fixed effects of treatment, visit (as categorical), and treatment-by-visit interactions, adjusted for baseline PHQ-9 scores, employment status, age and gender. The interaction P-values are calculated from the MMRM models involving interaction terms of the corresponding treatments. P-values are two-sided and no adjustments are made for multiple comparisons.

† Plus value indicates well-being increase. () indicates 95% confidence intervals.

‡ Effect sizes (standardised mean differences) were calculated using the observed week 6 SD for group differences.

AT: Assertion training, BA: Behavioural activation, BI: Behaviour therapy for insomnia, CR: Cognitive restructuring, PS: Problem solving, SWEMWBS: Short Warwick Edinburgh Mental Well-Being Scale

Supplementary Data. Table 5. The observed means and SDs of the PHQ-9 scores through the trial

Arm	Week 0			Week 1			Week 2			Week 3			Week 4			Week 5			Week 6		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
BA	327	7.90	2.49	313	5.88	2.87	302	5.47	3.03	313	5.14	3.00	304	4.92	3.06	304	4.70	3.08	313	4.48	3.12
CR	328	8.16	2.71	307	6.47	3.42	298	5.90	3.42	312	5.49	3.33	307	5.35	3.47	310	5.23	3.47	320	5.03	3.48
PS	328	8.00	2.81	296	6.48	3.59	303	5.84	3.47	297	5.79	3.77	298	5.67	3.91	296	5.27	3.60	311	4.94	3.75
AT	328	8.18	2.95	304	6.61	3.33	300	5.85	3.70	306	5.73	3.50	309	5.35	3.42	301	5.13	3.44	319	4.99	3.69
BI	328	8.06	2.73	323	5.76	3.11	319	5.43	3.24	320	5.26	3.21	319	4.93	3.45	319	4.79	3.46	321	4.64	3.62
BA + CR	329	8.13	2.93	310	6.00	3.36	308	5.49	3.21	313	5.30	3.32	308	5.02	3.32	302	5.08	3.62	313	4.58	3.25
BA + PS	329	8.09	2.78	316	6.33	3.26	313	5.79	3.43	319	5.20	3.22	309	5.35	3.27	309	4.89	3.25	321	4.50	2.93
BA + AT	328	8.17	2.73	301	6.47	3.39	307	5.82	3.51	309	5.47	3.37	305	5.28	3.43	304	5.11	3.81	316	4.74	3.63
BA + BI	327	7.95	2.59	312	6.13	3.37	309	5.66	3.43	311	5.60	3.77	307	5.08	3.47	314	5.04	3.70	318	4.74	3.60
Self-check	328	8.02	2.76	320	6.64	3.81	322	6.02	3.41	321	5.84	3.60	318	5.98	3.75	323	5.49	3.58	325	5.56	3.68
Health information	328	8.06	2.66	0	NA	NA	0	NA	NA	311	6.47	3.49	0	NA	NA	0	NA	NA	322	6.20	3.55
Delayed treatment	328	8.13	2.67	0	NA	NA	0	NA	NA	319	7.09	3.59	0	NA	NA	0	NA	NA	325	6.85	3.72
Follow-up proportion	100%			94.6%			93.9%			95.3%			94.0%			94.0%			97.2%		

Arm	Week 10			Week 14			Week 18			Week 22			Week 26		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
BA	294	5.17	3.62	293	5.03	3.45	285	5.07	3.78	287	4.93	3.75	302	5.09	3.94
CR	312	5.27	3.63	295	5.18	3.82	288	5.01	3.61	290	5.29	4.03	297	5.24	4.23
PS	295	5.24	3.88	286	5.25	3.97	278	5.33	3.76	273	5.26	4.07	289	5.20	4.03
AT	303	5.29	3.98	293	5.30	4.09	295	5.35	4.04	289	5.48	4.43	297	5.23	4.53
BI	303	5.04	3.38	299	4.64	3.77	289	4.90	3.96	274	4.85	3.72	291	4.81	3.79
BA + CR	290	4.81	3.25	284	4.88	3.44	284	4.80	3.31	272	4.65	3.54	290	4.69	3.68
BA + PS	301	5.04	3.77	296	5.05	3.52	284	5.12	3.53	282	5.14	3.85	293	4.77	3.48
BA + AT	299	5.20	3.84	284	5.00	3.90	276	4.92	3.95	269	4.89	3.98	281	4.95	3.78
BA + BI	301	4.91	3.58	287	4.89	4.16	278	4.96	4.11	273	4.80	3.85	286	5.05	4.38
Self-check	316	6.13	3.80	320	6.16	4.09	316	5.80	3.82	314	5.87	4.14	317	6.08	4.28
Health information	313	6.09	3.75	310	5.73	3.83	311	5.83	4.05	309	5.77	4.28	315	5.82	4.28
Delayed treatment	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Follow-up proportion	92.2%			90.0%			88.2%			86.8%			90.3%		

AT: Assertion training, BA: Behavioural activation, BI: Behaviour therapy for insomnia, CR: Cognitive restructuring, PS: Problem solving

Supplementary Data. Table 6. Sensitivity analysis of main and interaction effects using the mixed-model repeated measures analysis (MMRM) for PHQ-9 score (subgroup analysis by ISI score at baseline)

Arm	N	Estimated differences of change scores at Week 6 (95% CI)	Effect size for differences of change scores at Week 6 [‡] (95% CI)	P-value (Main effect)	P-value (Interaction of two components)	P-value (Interaction for the subgroups)
Trial 4 (ISI score ≥ 8)						
BA + BI	207	-2.15 (-2.78 to -1.51)	-0.59 (-0.76 to -0.41)	5.9×10 ⁻¹¹	1.2×10 ⁻⁹	0.388
BA	223	-2.60 (-3.22 to -1.97)	-0.71 (-0.88 to -0.54)	1.1×10 ⁻¹⁵		0.057
BI	218	-2.35 (-2.97 to -1.72)	-0.64 (-0.81 to -0.47)	4.0×10 ⁻¹³		0.216
Delayed treatment	220	—	—	—		
Trial 4 (ISI score ≤ 7)						
BA + BI	107	-1.71 (-2.48 to -0.93)	-0.56 (-0.81 to -0.30)	2.0×10 ⁻⁵	0.010	
BA	88	-1.55 (-2.37 to -0.73)	-0.50 (-0.77 to -0.24)	2.3×10 ⁻⁴		
BI	95	-1.65 (-2.45 to -0.85)	-0.54 (-0.8 to -0.28)	6.4×10 ⁻⁵		
Delayed treatment	93	—	—	—		

The least squares mean changes and their differences of the PHQ-9 scores were estimated with the mixed-effects models for repeated-measures (MMRM). The regression models included fixed effects of treatment, visit (as categorical), and treatment-by-visit interactions, adjusted for baseline PHQ-9 scores, employment status, age and gender. The interaction P-values are calculated from the MMRM models involving interaction terms of the corresponding treatments or subgroups. P-values are two-sided and no adjustments are made for multiple comparisons.

† Minus value indicates depression reduction. () indicates 95% confidence intervals.

‡ Effect sizes (standardised mean differences) were calculated using the observed week 6 SD for group differences.

BA: Behavioural activation, BI: Behaviour therapy for insomnia, ISI: Insomnia Severity Index

Supplementary Data. Table 7. Sensitivity analysis of main and interaction effects using the mixed-model repeated measures analysis (MMRM) for GAD-7 score (sensitivity analysis for GAD-7 at baseline ≥ 5)

Arm	N	Estimated differences of change scores at Week 6 (95% CI)	Effect size for differences of change scores at Week 6 [‡] (95% CI)	P-value (Main effect)	P-value (Interaction of two components)
Trial 1					
BA + CR	203	-1.85 (-2.47 to -1.23)	-0.54 (-0.73 to -0.36)	7.3×10^{-9}	6.9×10^{-5}
BA	195	-2.09 (-2.72 to -1.46)	-0.62 (-0.80 to -0.43)	1.2×10^{-10}	
CR	198	-1.56 (-2.18 to -0.94)	-0.46 (-0.64 to -0.28)	9.9×10^{-7}	
Delayed treatment	195	—	—	—	
Trial 2					
BA + PS	206	-1.76 (-2.39 to -1.13)	-0.51 (-0.69 to -0.33)	4.7×10^{-8}	0.001
BA	195	-2.10 (-2.74 to -1.46)	-0.61 (-0.79 to -0.42)	2.5×10^{-10}	
PS	192	-1.15 (-1.80 to -0.51)	-0.33 (-0.52 to -0.15)	4.6×10^{-4}	
Delayed treatment	195	—	—	—	
Trial 3					
BA + AT	200	-1.56 (-2.18 to -0.94)	-0.45 (-0.63 to -0.27)	1.1×10^{-6}	8.9×10^{-5}
BA	195	-2.05 (-2.68 to -1.42)	-0.59 (-0.77 to -0.41)	3.2×10^{-10}	
AT	198	-1.30 (-1.92 to -0.67)	-0.37 (-0.55 to -0.19)	5.3×10^{-5}	
Delayed treatment	195	—	—	—	
Trial 4					
BA + BI	191	-1.33 (-1.97 to -0.70)	-0.39 (-0.57 to -0.20)	4.2×10^{-5}	1.1×10^{-8}
BA	184	-2.08 (-2.72 to -1.43)	-0.60 (-0.79 to -0.41)	4.7×10^{-10}	
BI	196	-1.91 (-2.54 to -1.28)	-0.55 (-0.74 to -0.37)	3.9×10^{-9}	
Delayed treatment	186	—	—	—	

The least squares mean changes and their differences of the PHQ-9 scores were estimated with the mixed-effects models for repeated-measures (MMRM). The regression models included fixed effects of treatment, visit (as categorical), and treatment-by-visit interactions, adjusted for baseline PHQ-9 scores, employment status, age and gender. P-values are two-sided and no adjustments are made for multiple comparisons.

† Minus value indicates anxiety reduction. () indicates 95% confidence intervals.

‡ Effect sizes (standardised mean differences) were calculated using the observed week 6 SD for group differences.

AT: Assertion training, BA: Behavioural activation, BI: Behaviour therapy for insomnia, CR: Cognitive restructuring, GAD-7: General Anxiety Disorder-7, PS: Problem solving

Supplementary Data. Table 8. Sensitivity analysis of main and interaction effects using the mixed-model repeated measures analysis (MMRM) for ISI score (sensitivity analysis for ISI at baseline ≥ 8)

Arm	N	Estimated differences of change scores at Week 6 (95% CI)	Effect size for differences of change scores at Week 6 [‡] (95% CI)	P-value (Main effect)	P-value (Interaction of two components)
Trial 1					
BA + CR	231	-1.41 (-2.12 to -0.71)	-0.32 (-0.47 to -0.16)	8.0×10^{-5}	0.006
BA	235	-1.69 (-2.38 to -0.99)	-0.38 (-0.53 to -0.22)	2.4×10^{-6}	
CR	220	-1.13 (-1.83 to -0.42)	-0.25 (-0.41 to -0.09)	0.002	
Delayed treatment	231	—	—	—	
Trial 2					
BA + PS	238	-1.44 (-2.16 to -0.71)	-0.31 (-0.46 to -0.15)	1.0×10^{-4}	0.023
BA	235	-1.71 (-2.43 to -0.98)	-0.36 (-0.52 to -0.21)	4.6×10^{-6}	
PS	215	-0.94 (-1.68 to -0.19)	-0.20 (-0.36 to -0.04)	0.014	
Delayed treatment	231	—	—	—	
Trial 3					
BA + AT	238	-1.43 (-2.15 to -0.71)	-0.31 (-0.46 to -0.15)	1.0×10^{-4}	0.006
BA	235	-1.66 (-2.39 to -0.94)	-0.36 (-0.51 to -0.20)	7.0×10^{-6}	
AT	239	-1.21 (-1.93 to -0.49)	-0.26 (-0.41 to -0.11)	0.001	
Delayed treatment	231	—	—	—	
Trial 4					
BA + BI	207	-2.40 (-3.12 to -1.69)	-0.55 (-0.71 to -0.39)	7.8×10^{-11}	2.2×10^{-5}
BA	223	-1.69 (-2.39 to -0.99)	-0.39 (-0.55 to -0.23)	2.7×10^{-6}	
BI	218	-2.90 (-3.61 to -2.20)	-0.66 (-0.82 to -0.50)	2.1×10^{-15}	
Delayed treatment	220	—	—	—	

The least squares mean changes and their differences of the PHQ-9 scores were estimated with the mixed-effects models for repeated-measures (MMRM). The regression models included fixed effects of treatment, visit (as categorical), and treatment-by-visit interactions, adjusted for baseline PHQ-9 scores, employment status, age and gender. P-values are two-sided and no adjustments are made for multiple comparisons.

† Minus value indicates insomnia reduction. () indicates 95% confidence intervals.

‡ Effect sizes (standardised mean differences) were calculated using the observed week 6 SD for group differences.

AT: Assertion training, BA: Behavioural activation, BI: Behaviour therapy for insomnia, CR: Cognitive restructuring, ISI: Insomnia Severity Index, PS: Problem solving

Supplementary Data. Table 9. Sensitivity analysis of main and interaction effects using the mixed-model repeated measures analysis (MMRM) for SWEMWBS score (sensitivity analysis for SWEMWBS at baseline ≤ 20)

Arm	N	Estimated differences of change scores at Week 6 (95% CI)	Effect size for differences of change scores at Week 6 [‡] (95% CI)	P-value (Main effect)	P-value (Interaction of two components)
Trial 1					
BA + CR	104	1.25 (0.33 to 2.17)	0.35 (0.09 to 0.61)	0.008	0.650
BA	97	1.13 (0.19 to 2.07)	0.32 (0.05 to 0.58)	0.018	
CR	112	0.42 (-0.47 to 1.31)	0.12 (-0.13 to 0.37)	0.357	
Delayed treatment	110	—	—	—	
Trial 2					
BA + PS	122	0.87 (-0.02 to 1.76)	0.24 (-0.01 to 0.48)	0.055	0.531
BA	97	1.10 (0.16 to 2.05)	0.30 (0.04 to 0.56)	0.022	
PS	96	0.19 (-0.75 to 1.14)	0.05 (-0.20 to 0.31)	0.687	
Delayed treatment	110	—	—	—	
Trial 3					
BA + AT	113	0.74 (-0.14 to 1.63)	0.20 (-0.04 to 0.44)	0.100	0.074
BA	97	1.11 (0.19 to 2.03)	0.30 (0.05 to 0.55)	0.019	
AT	119	0.79 (-0.08 to 1.66)	0.22 (-0.02 to 0.45)	0.073	
Delayed treatment	110	—	—	—	
Trial 4					
BA + BI	106	0.39 (-0.53 to 1.32)	0.10 (-0.14 to 0.35)	0.402	0.002
BA	92	1.27 (0.31 to 2.23)	0.33 (0.08 to 0.59)	0.010	
BI	115	1.22 (0.32 to 2.13)	0.32 (0.08 to 0.56)	0.008	
Delayed treatment	107	—	—	—	

The least squares mean changes and their differences of the PHQ-9 scores were estimated with the mixed-effects models for repeated-measures (MMRM). The regression models included fixed effects of treatment, visit (as categorical), and treatment-by-visit interactions, adjusted for baseline PHQ-9 scores, employment status, age and gender. P-values are two-sided and no adjustments are made for multiple comparisons.

† Plus value indicates well-being increase. () indicates 95% confidence intervals.

‡ Effect sizes (standardised mean differences) were calculated using the observed week 6 SD for group differences.

AT: Assertion training, BA: Behavioural activation, BI: Behaviour therapy for insomnia, CR: Cognitive restructuring, PS: Problem solving, SWEMWBS: Short Warwick Edinburgh Mental Well-Being Scale

Supplementary Data. Table 10. Analysis of main and interaction effects using the mixed-model repeated measures analysis (MMRM) for PHQ-9 score at Week 26

10a. Using the Health information arm as the control

Arm	N	Change scores from baseline to Week 26 of the PHQ-9		Differences in change scores of the PHQ-9 at Week 26			
		Estimated least squares mean change scores [†]	Effect size [‡]	Estimated least squares mean differences	Effect size [‡]	<i>P</i> -value (Main effect)	<i>P</i> -value (Interaction of two components)
Trial 1							
BA + CR	329	-2.81 (-3.68 to -1.94)	-1.04 (-1.36 to -0.72)	-1.23 (-1.83 to -0.63)	-0.36 (-0.54 to -0.19)	4.1×10 ⁻⁴	0.587
BA	327	-2.41 (-3.27 to -1.56)	-0.89 (-1.21 to -0.57)	-0.83 (-1.43 to -0.24)	-0.24 (-0.42 to -0.07)	0.006	
CR	328	-2.37 (-3.23 to -1.51)	-0.88 (-1.19 to -0.56)	-0.79 (-1.39 to -0.19)	-0.23 (-0.41 to -0.06)	0.010	
Health information [reference]	328	-1.58 (-2.44 to -0.72)	-0.58 (-0.90 to -0.27)	—	—	—	
Trial 2							
BA + PS	329	-3.08 (-3.84 to -2.31)	-1.14 (-1.43 to -0.86)	-1.20 (-1.79 to -0.60)	-0.35 (-0.53 to -0.18)	4.8×10 ⁻⁴	0.827
BA	327	-2.71 (-3.47 to -1.94)	-1.01 (-1.29 to -0.72)	-0.83 (-1.42 to -0.23)	-0.24 (-0.42 to -0.07)	0.007	
PS	328	-2.52 (-3.29 to -1.74)	-0.93 (-1.22 to -0.65)	-0.63 (-1.24 to -0.03)	-0.19 (-0.36 to -0.01)	0.038	
Health information [reference]	328	-1.88 (-2.65 to -1.11)	-0.70 (-0.99 to -0.41)	—	—	—	
Trial 3							
BA + AT	328	-2.98 (-3.88 to -2.08)	-1.10 (-1.43 to -0.77)	-1.05 (-1.66 to -0.44)	-0.30 (-0.47 to -0.13)	0.001	0.485
BA	327	-2.76 (-3.65 to -1.87)	-1.02 (-1.34 to -0.69)	-0.83 (-1.43 to -0.23)	-0.23 (-0.40 to -0.06)	0.007	
AT	328	-2.62 (-3.52 to -1.71)	-0.96 (-1.29 to -0.63)	-0.69 (-1.29 to -0.08)	-0.19 (-0.36 to -0.02)	0.026	
Health information [reference]	328	-1.93 (-2.83 to -1.03)	-0.71 (-1.04 to -0.38)	—	—	—	
Trial 4							
BA + BI	311	-2.89 (-3.73 to -2.04)	-1.10 (-1.42 to -0.78)	-0.87 (-1.50 to -0.25)	-0.25 (-0.42 to -0.07)	0.006	0.072
BA	314	-2.84 (-3.68 to -2.01)	-1.08 (-1.40 to -0.77)	-0.83 (-1.45 to -0.21)	-0.24 (-0.41 to -0.06)	0.009	
BI	313	-3.05 (-3.88 to -2.22)	-1.16 (-1.48 to -0.85)	-1.03 (-1.66 to -0.41)	-0.29 (-0.47 to -0.12)	0.001	
Health information [reference]	313	-2.02 (-2.85 to -1.18)	-0.77 (-1.09 to -0.45)	—	—	—	

The least squares mean changes and their differences of the PHQ-9 scores were estimated with the mixed-effects models for repeated-measures (MMRM). The regression models included fixed effects of treatment, visit (as categorical), and treatment-by-visit interactions, adjusted for baseline PHQ-9 scores, employment status, age and gender. P-values are two-sided and no adjustments are made for multiple comparisons.

† Minus value indicates depression reduction. () indicates 95% confidence intervals.

‡ Effect sizes (standardised mean differences) were calculated using the observed baseline SD for within-group change scores and the observed week 6 SD for group differences.

AT: Assertion training, BA: Behavioural activation, BI: Behaviour therapy for insomnia, CR: Cognitive restructuring, PHQ-9: Patient Health Questionnaire-9, PS: Problem solving

10b. Using the Self-check arm as the control

Arm	N	Change scores from baseline to Week 26 of the PHQ-9		Differences in change scores of the PHQ-9 at Week 26			
		Estimated least squares mean change scores [†]	Effect size [‡]	Estimated least squares mean differences	Effect size [‡]	<i>P</i> -value (Main effect)	<i>P</i> -value (Interaction of two components)
Trial 1							
BA + CR	329	-3.11 (-3.82 to -2.40)	-1.14 (-1.40 to -0.88)	-1.37 (-1.97 to -0.77)	-0.40 (-0.58 to -0.23)	1.8×10 ⁻³	0.247
BA	327	-2.69 (-3.39 to -1.99)	-0.99 (-1.24 to -0.73)	-0.95 (-1.55 to -0.36)	-0.28 (-0.46 to -0.10)	0.002	
CR	328	-2.66 (-3.36 to -1.96)	-0.97 (-1.23 to -0.72)	-0.92 (-1.52 to -0.32)	-0.27 (-0.45 to -0.09)	0.003	
Self-check [reference]	328	-1.74 (-2.42 to -1.06)	-0.64 (-0.89 to -0.39)	—	—	—	
Trial 2							
BA + PS	329	-3.14 (-3.81 to -2.46)	-1.16 (-1.41 to -0.91)	-1.34 (-1.94 to -0.75)	-0.40 (-0.57 to -0.22)	1.1×10 ⁻⁵	0.376
BA	327	-2.75 (-3.43 to -2.08)	-1.01 (-1.26 to -0.77)	-0.96 (-1.55 to -0.36)	-0.28 (-0.46 to -0.11)	0.002	
PS	328	-2.57 (-3.25 to -1.89)	-0.95 (-1.20 to -0.70)	-0.77 (-1.37 to -0.17)	-0.23 (-0.41 to -0.05)	0.012	
Self-check [reference]	328	-1.80 (-2.46 to -1.14)	-0.66 (-0.90 to -0.42)	—	—	—	
Trial 3							
BA + AT	328	-3.08 (-3.81 to -2.35)	-1.12 (-1.39 to -0.86)	-1.20 (-1.81 to -0.59)	-0.34 (-0.51 to -0.17)	1.1×10 ⁻⁴	0.193
BA	327	-2.83 (-3.55 to -2.11)	-1.03 (-1.30 to -0.77)	-0.96 (-1.56 to -0.35)	-0.27 (-0.44 to -0.10)	0.002	
AT	328	-2.69 (-3.42 to -1.96)	-0.98 (-1.25 to -0.72)	-0.82 (-1.42 to -0.22)	-0.23 (-0.4 to -0.06)	0.008	
Self-check [reference]	328	-1.87 (-2.58 to -1.17)	-0.68 (-0.94 to -0.43)	—	—	—	
Trial 4							
BA + BI	311	-2.96 (-3.71 to -2.22)	-1.12 (-1.40 to -0.84)	-0.99 (-1.62 to -0.36)	-0.28 (-0.46 to -0.10)	0.002	0.017
BA	314	-2.91 (-3.65 to -2.17)	-1.10 (-1.38 to -0.82)	-0.94 (-1.57 to -0.31)	-0.26 (-0.44 to -0.09)	0.004	
BI	313	-3.12 (-3.86 to -2.39)	-1.18 (-1.46 to -0.90)	-1.15 (-1.78 to -0.52)	-0.32 (-0.50 to -0.15)	3.6×10 ⁻³	
Self-check [reference]	313	-1.98 (-2.70 to -1.25)	-0.75 (-1.02 to -0.47)	—	—	—	

The least squares mean changes and their differences of the PHQ-9 scores were estimated with the mixed-effects models for repeated-measures (MMRM). The regression models included fixed effects of treatment, visit (as categorical), and treatment-by-visit interactions, adjusted for baseline PHQ-9 scores, employment status, age and gender. P-values are two-sided and no adjustments are made for multiple comparisons.

† Minus value indicates depression reduction. () indicates 95% confidence intervals.

‡ Effect sizes (standardised mean differences) were calculated using the observed baseline SD for within-group change scores and the observed week 6 SD for group differences.

AT: Assertion training, BA: Behavioural activation, BI: Behaviour therapy for insomnia, CR: Cognitive restructuring, PHQ-9: Patient Health Questionnaire-9, PS: Problem solving

Supplementary Data. Table 11. Component analysis of specific effects of individual iCBT skills on the PHQ-9 scores at week 6, using the mixed-model repeated measures analysis (MMRM) with or without interactions

Component	N	Differences of change scores of the PHQ-9			Component	N	Differences of change scores of the PHQ-9		
		Estimated least squares mean difference	Effect size [†]	P-value			Estimated least squares mean difference	Effect size [†]	P-value
<i>ns</i>	3280	-0.93 (-1.37 to -0.49)	-0.27 (-0.39 to -0.14)	3.6×10 ⁻⁵	<i>ns</i>	3280	-0.63 (-1.12 to -0.13)	-0.18 (-0.32 to -0.04)	0.013
<i>ba</i>	1640	-0.38 (-0.61 to -0.16)	-0.11 (-0.17 to -0.05)	0.001	<i>ba</i>	1640	-1.00 (-1.49 to -0.50)	-0.28 (-0.43 to -0.14)	7.7×10 ⁻⁵
<i>cr</i>	657	-0.30 (-0.65 to 0.05)	-0.09 (-0.19 to 0.01)	0.091	<i>cr</i>	657	-0.62 (-1.12 to -0.13)	-0.18 (-0.32 to -0.04)	0.013
<i>ps</i>	657	-0.31 (-0.66 to 0.04)	-0.09 (-0.19 to 0.01)	0.081	<i>ps</i>	657	-0.55 (-1.05 to -0.06)	-0.16 (-0.3 to -0.02)	0.029
<i>at</i>	656	-0.26 (-0.61 to 0.09)	-0.07 (-0.17 to 0.03)	0.142	<i>at</i>	656	-0.64 (-1.13 to -0.15)	-0.18 (-0.32 to -0.04)	0.011
<i>bi</i>	655	-0.36 (-0.71 to -0.01)	-0.10 (-0.2 to 0.00)	0.044	<i>bi</i>	655	-0.94 (-1.44 to -0.45)	-0.27 (-0.41 to -0.13)	1.7×10 ⁻⁴
<i>wl</i>	328	0.58 (0.09 to 1.08)	0.17 (0.03 to 0.31)	0.020	<i>ba*cr</i>	—	0.65 (-0.05 to 1.35)	0.18 (-0.02 to 0.38)	0.070
					<i>ba*ps</i>	—	0.49 (-0.21 to 1.19)	0.14 (-0.06 to 0.34)	0.172
					<i>ba*at</i>	—	0.76 (0.06 to 1.46)	0.22 (0.02 to 0.42)	0.033
					<i>ba*bi</i>	—	1.18 (0.48 to 1.88)	0.34 (0.14 to 0.54)	0.001
					<i>wl</i>	—	0.58 (0.09 to 1.08)	0.17 (0.03 to 0.31)	0.020

The least squares mean changes and their differences of the PHQ-9 scores were estimated with the mixed-effects models for repeated-measures (MMRM). The regression models included fixed effects of components, their interactions with *ba*, visit (as categorical), and component-by-visit interactions, adjusted for baseline PHQ-9 scores, employment status, age and gender. P-values are two-sided and no adjustments are made for multiple comparisons.

† Minus value indicates depression reduction. () indicates 95% confidence intervals.

‡ Effect sizes (standardised mean differences) were calculated using the observed week 6 SD for group differences.

ns (non-specific treatment effect including weekly encouragement emails and self-checks with the PHQ-9), *ba* (specific effect of behavioural activation), *cr* (specific effect of cognitive restructuring), *ps* (specific effect of problem solving), *at* (specific effect of assertion training), *bi* (specific effect of behaviour therapy for insomnia), *wl* (waiting list).

AT: Assertion training, BA: Behavioural activation, BI: Behaviour therapy for insomnia, CR: Cognitive restructuring, PHQ-9: Patient Health Questionnaire-9, PS: Problem solving

Supplementary Data. Table 12. Component analysis of specific effects of individual iCBT skills on the PHQ-9 scores at week 26, using the mixed-model repeated measures analysis (MMRM) with or without interactions

Component	N	N at week 26	Differences of change scores of the PHQ-9		
			Estimated least squares mean difference	Effect size [†]	P-value
<i>ns</i>	3280	2943	0.03 (-0.5 to 0.56)	0.01 (-0.14 to 0.16)	0.916
<i>ba</i>	1640	1452	-0.43 (-0.7 to -0.16)	-0.12 (-0.2 to -0.05)	0.002
<i>cr</i>	657	587	-0.68 (-1.1 to -0.26)	-0.19 (-0.32 to -0.07)	0.002
<i>ps</i>	657	582	-0.59 (-1.01 to -0.16)	-0.17 (-0.29 to -0.05)	0.007
<i>at</i>	656	578	-0.54 (-0.96 to -0.11)	-0.15 (-0.28 to -0.03)	0.013
<i>bi</i>	655	577	-0.59 (-1.02 to -0.17)	-0.17 (-0.29 to -0.05)	0.006

Component	N	Differences of change scores of the PHQ-9		
		Estimated least squares mean difference	Effect size [†]	P-value
<i>ns</i>	3280	0.29 (-0.3 to 0.88)	0.08 (-0.09 to 0.25)	0.330
<i>ba</i>	1640	-0.96 (-1.56 to -0.37)	-0.28 (-0.45 to -0.11)	0.002
<i>cr</i>	657	-0.93 (-1.53 to -0.34)	-0.27 (-0.44 to -0.1)	0.002
<i>ps</i>	657	-0.78 (-1.38 to -0.18)	-0.22 (-0.4 to -0.05)	0.011
<i>at</i>	656	-0.82 (-1.42 to -0.23)	-0.24 (-0.41 to -0.06)	0.007
<i>bi</i>	655	-1.19 (-1.79 to -0.59)	-0.34 (-0.51 to -0.17)	9.6×10 ⁻⁵
<i>ba*cr</i>	—	0.51 (-0.34 to 1.36)	0.15 (-0.1 to 0.39)	0.237
<i>ba*ps</i>	—	0.39 (-0.46 to 1.24)	0.11 (-0.13 to 0.36)	0.366
<i>ba*at</i>	—	0.57 (-0.28 to 1.42)	0.16 (-0.08 to 0.41)	0.188
<i>ba*bi</i>	—	1.21 (0.36 to 2.05)	0.35 (0.10 to 0.59)	0.005

The least squares mean changes and their differences of the PHQ-9 scores were estimated with the mixed-effects models for repeated-measures (MMRM). The regression models included fixed effects of components, their interactions with *ba*, visit (as categorical), and component-by-visit interactions, adjusted for baseline PHQ-9 scores, employment status, age and gender. P-values are two-sided and no adjustments are made for multiple comparisons.

† Minus value indicates depression reduction. () indicates 95% confidence intervals.

‡ Effect sizes (standardised mean differences) were calculated using the observed week 6 SD for group differences.

ns (non-specific treatment effect including weekly encouragement emails and self-checks with the PHQ-9), *ba* (specific effect of behavioural activation), *cr* (specific effect of cognitive restructuring), *ps* (specific effect of problem solving), *at* (specific effect of assertion training), *bi* (specific effect of behaviour therapy for insomnia), *wl* (waiting list).

AT: Assertion training, BA: Behavioural activation, BI: Behaviour therapy for insomnia, CR: Cognitive restructuring, PHQ-9: Patient Health Questionnaire-9, PS: Problem solving

4. CONSORT 2010 Checklist for nonpharmacological and factorial trials



CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item from CONSORT 2010	Items added for nonpharmacological / factorial trials	Reported on page No
Title and abstract				
	1a	Identification as a randomised trial in the title	Identification as a factorial randomized trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)		3
Introduction				
Background and objectives	2a	Scientific background and explanation of rationale	Scientific background and rationale for using a factorial design, including whether an interaction is hypothesized	4, 22
	2b	Specific objectives or hypotheses	Specific objectives or hypotheses and a statement of which treatment groups form the main comparisons ^b	4
Methods				
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	When applicable, how care providers were allocated to each trial group Description of the type of factorial trial (such as full or partial, number of factors, levels within each factor ^b) and allocation ratio	22
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons		28
Participants	4a	Eligibility criteria for participants	When applicable, eligibility criteria for centers and for care providers Eligibility criteria for each factor, noting any differences, if applicable	22
	4b	Settings and locations where the data were collected		22
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	Precise details of both the experimental treatment and comparator	23-24
	5a		Description of the different components of the	23

			interventions and, when applicable, description of the procedure for tailoring the interventions to individual participants.	
	5b		Details of whether and how the interventions were standardized.	NA because this is an app
	5c		Details of whether and how adherence of care providers to the protocol was assessed or enhanced	NA
	5d		Details of whether and how adherence of participants to interventions was assessed or enhanced	27
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed		25-26
	6b	Any changes to trial outcomes after the trial commenced, with reasons		NA
Sample size	7a	How sample size was determined	When applicable, details of whether and how the clustering by care providers or centers was addressed How sample size was determined for each main comparison, including whether an interaction was assumed in the calculation	26
	7b	When applicable, explanation of any interim analyses and stopping guidelines	When applicable, explanation of any interim analyses and stopping guidelines, noting any differences across main comparisons and reasons for differences	NA
Randomisation: Sequence generation	8a	Method used to generate the random allocation sequence	Type of randomization, details of any restriction (such as blocking and block size), and, if applicable, whether participants were randomized to factors at different time points	22
	8b	Type of randomisation; details of any restriction (such as blocking and block size)		22
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned		22

Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions		22
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	If done, who was blinded after assignment to interventions (e.g., participants, care providers, those administering co-interventions, those assessing outcomes) and how	22-23
	11b	If relevant, description of the similarity of interventions		23
	11c		If blinding was not possible, description of any attempts to limit bias	10
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	When applicable, details of whether and how the clustering by care providers or centers was addressed Statistical methods used for each main comparison for primary and secondary outcomes, including: ● Whether the target treatment effect for each main comparison pertains to the effect in the presence or absence of other factors ● Approach to analysis, such as factorial or multiarm ● How the approach was chosen, such as prespecified or based on estimated interaction ● If factorial approach was used, whether factors were adjusted for each other ● If applicable, how nonconcurrent recruitment to factors was handled ● Method(s) used to evaluate statistical interaction(s)	26-27
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses		27
Results				
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	The number of care providers or centers performing the intervention in each group and the number of patients treated by each care provider or in each	Figure 1, Supplementary Data Table

			center	S5
			For each main comparison, the number of participants who were randomly assigned, received intended treatment, and were analyzed for the primary outcome	
	13b	For each group, losses and exclusions after randomisation, together with reasons	For each main comparison, losses and exclusions after randomization, together with reasons	Figure 1
	13c		For each group, the delay between randomization and the initiation of the intervention	22
	New		Details of the experimental treatment and comparator as they were implemented	23-24
Recruitment	14a	Dates defining the periods of recruitment and follow-up	Dates defining the periods of recruitment and follow-up for each factor, noting any differences, with reasons	5
	14b	Why the trial ended or was stopped		NA
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	When applicable, a description of care providers (case volume, qualification, expertise, etc.) and centers (volume) in each group	Supplementary Data Table S1
			A table showing baseline demographic and clinical characteristics for each main comparison	
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	For each main comparison, the number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	Figure 1
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	For each primary and secondary outcome, results for each main comparison, the estimated effect size, and its precision (such as 95% CI)	5-6, Tables 2-4
			For each primary outcome, the estimated interaction effect and its precision	Extended Data Tables 1-5
			If done, the estimated interaction effects and precision for secondary outcomes	
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended		NA
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses,		7

	18b	distinguishing pre-specified from exploratory	Participant flow, losses and exclusions, baseline data, and outcome data (including primary and secondary outcomes, harms, and adherence) presented by treatment groups ^b	Figure 1, Supplementary Data Table S1
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	All important harms or unintended effects for each main comparison	7
Discussion				
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	In addition, take into account the choice of the comparator, lack of or partial blinding, and unequal expertise of care providers or centers in each group	10
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	Generalizability (external validity) of the trial findings according to the intervention, comparators, patients, and care providers and centers involved in the trial	10
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		10-11
Other information				
Registration	23	Registration number and name of trial registry		22
Protocol	24	Where the full trial protocol can be accessed, if available		22, Supplement 1
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders		12

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.

^b Each overall intervention group to be compared is a factor (eg, in a 2 × 2 trial with factors A and B, active A and control A together comprise one factor and active B and control B together comprise another factor). The specific interventions within a factor are the levels (eg, active A and control A are the 2 levels of factor A). The unique combinations of factors and levels are treatment groups (eg, in a 2 × 2 trial with factors A and B there will be 4 treatment groups: active A plus control B, active A plus active B, etc). What treatment groups will be compared against each other to draw main conclusions about the effectiveness of each intervention is the main comparison.