

A Trial Protocol for the Effectiveness of Digital Interventions for Preventing Depression in Adolescents: The Future Proofing Study

Aliza Werner-Seidler (✉ a.werner-seidler@blackdog.org.au)

University of New South Wales <https://orcid.org/0000-0002-9046-6159>

Kit Huckvale

Black Dog Institute

Mark E. Larsen

Black Dog Institute

Alison L. Calear

Australian National University

Kate Maston

Black Dog Institute

Lara Johnston

Black Dog Institute

Michelle Torok

Black Dog Institute

Bridianne O'Dea

Black Dog Institute

Philip J. Batterham

Australian National University

Susanne Schweizer

University of New South Wales

S. Rachel Skinner

University of Sydney

Katharine Steinbeck

University of Sydney

Julie Ratcliffe

Flinders University

Ju-Lee Oei

University of New South Wales

George Patton

Murdoch Children's Research Institute

Iana Wong

Black Dog Institute

Joanne Beames

Black Dog Institute

Quincy J. J. Wong

Western Sydney University

Raghu Lingam

University of New South Wales

Katherine Boydell

Black Dog Institute

Allison M. Salmon

Black Dog Institute

Nicole Cockayne

Black Dog Institute

Andrew Mackinnon

Black Dog Institute

Helen Christensen

Black Dog Institute

Keywords: depression, prevention, adolescents, school, mental health, eHealth, mHealth, digital health, smartphone application, CBT

Posted Date: November 13th, 2019

DOI: <https://doi.org/10.21203/rs.2.14135/v2>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published on January 2nd, 2020. See the published version at <https://doi.org/10.1186/s13063-019-3901-7>.

Abstract

Background Depression frequently first emerges during adolescence, and one in five young people will experience an episode of depression by the age of 18 years. Despite advances in treatment, there has been limited progress in addressing the burden at a population level. Accordingly, there has been growing interest in prevention approaches as an additional pathway to address depression. Depression can be prevented using evidence-based psychological programs. However, barriers to implementing and accessing these programs remain, typically reflecting a requirement for delivery by clinical experts and high associated delivery costs. Digital technologies, specifically smartphones, are now considered a key strategy to overcome the barriers inhibiting access to mental health programs. The Future Proofing Study is a large-scale school-based trial investigating whether Cognitive Behaviour Therapies (CBT) delivered by smartphone application can prevent depression.

Methods A randomized controlled trial targeting up to 20,000 Year 8 Australian secondary school students will be conducted. In Stage I, schools will be randomized at the cluster level to receive either the CBT intervention app (SPARX) or to a non-active control group comparator. The primary outcome will be symptoms of depression and secondary outcomes include psychological distress, anxiety and insomnia. At the 12-month follow-up, participants in the intervention arm with elevated depressive symptoms will participate in an individual-level RCT (Stage II) and be randomised to receive a second CBT app which targets sleep difficulties (Sleep Ninja) or a control condition. Assessments will occur at post-intervention (both trial stages), and at 6, 12, 24, 36, 48- and 60-months post-baseline. Primary analyses will use an intention-to-treat approach and compare changes in symptoms from baseline to follow-up relative to the control group using mixed effect models.

Discussion This is the first trial testing the effectiveness of smartphone apps delivered to school students to prevent depression at scale. Results from this trial will provide much needed insight into the feasibility of this approach. They stand to inform policy and commission decisions concerning if and how such programs should be deployed in school-based settings in Australia and beyond.

Background

According to the World Health Organization, depression is the leading cause of disability worldwide (1). For many individuals, depression first emerges during adolescence (2), between 10–24 years (3). Approximately one in five young people will experience a depressive episode by the age of 18 years (4), with earlier onset being associated with a range of adverse consequences including poor academic performance, low levels of school attendance, social dysfunction, substance abuse, anxiety, suicidality, poor sleep, diabetes and metabolic syndrome (5–7). As adolescents develop, the foundations are laid for the successful transition into the varied roles of adulthood including the time in life when employment productivity peaks. Experiencing depression as a young person can negatively impact psychological, social, emotional, educational and vocational pathways into adulthood, with profound implications both for the individual, their families and their communities, as well as the economy (5, 8). The global cost of mental illness is nearly US\$2.5 trillion, two thirds of which is incurred via indirect costs such as unemployment. This cost is expected to increase to US\$6 trillion by 2030 (9, 10). By way of benchmarking, total global health expenditure in 2009 was US\$5.1 trillion, of which less than 2% is spent on mental health (11). The immense cost of mental illness, together with the social and emotional consequences associated with earlier onset, highlights the crucial importance of prevention or intervention early in the course of illness, and consequently, at an early age.

There has been little progress in recent years addressing this challenge. Estimates suggest that only 13% of the disease burden due to depression can be alleviated through current treatments and service models (12). Even with optimal clinician competence, improved access and treatment adherence, only 36% of the burden is avoidable, leaving substantial unaddressed morbidity and disability (12). Currently, fewer than one third of young people seek help and gain access to first-line interventions for reasons that include accessibility, stigma and affordability (13). Of those who seek professional help, between 30–50% of young people do not respond to treatment (14), and among those who do, over 45% will experience a subsequent depressive episode (15). This has led to growing recognition that treatment alone will not successfully reduce the burden of depression at a population level (16).

Prevention represents a separate pathway through which the burden of depression may be addressed. There has been burgeoning interest in the use of preventive approaches, with several high quality but small-scale prevention studies demonstrating that the prevention of depression is possible in school aged young people. Several recent meta-analyses have examined the effect of psychological prevention programs specifically on adolescents and found consistent but small effect sizes in favor of delivering CBT for depression, at least in the short term (17–19). Of the prevention programs that are available, those based on Cognitive Behaviour Therapy (CBT) have been the most frequently evaluated interventions, with evidence to support these approaches (20). Other approaches may prove effective but have not yet been subject to rigorous empirical evaluation. CBT involves challenging unhelpful thoughts or beliefs, by gathering evidence which contradict these thoughts or beliefs from real world situations. This leads to changes in emotion and behaviour. CBT is structured, brief and time-limited, and aims to help individuals develop skills and strategies to manage their own mood and depression. Prevention approaches can be delivered either universally to all individuals within an identified population regardless of risk (universal prevention), or targeted to those with risk factors such as low socio-economic background (selective prevention), or to those who have sub-threshold disorder symptoms (indicated prevention). While there is evidence for all three prevention approaches (18, 19, 21), there have typically been larger effects reported for targeted approaches, at least in part due to limitations in the statistical power of universal prevention studies (22).

Despite consistent statistically and clinically meaningful effects in existing prevention studies, these studies have been limited by several factors. First, trials have generally involved small sample sizes, and the consequent limited statistical power means that potential effects go undetected and/or are measured imprecisely. Specifically, the median sample size identified in a recent meta-analysis of 47 studies examining the prevention of depression was 208, with an average effect size of $g = 0.23$ (19). Samples of this size are insufficient because preventive programs, particularly universally delivered programs, necessarily expose a large proportion of individuals not at risk to the intervention, and yield small to modest effects because of the low base rates of disorder in general population samples (21). Second, studies have rarely involved follow-up beyond 12 months, which limits what is known about long-term preventive effects.

Third, studies have relied solely on self-report or clinician-reported symptoms. While subjective reports reflect the gold-standard approach and is useful clinically, there is much to be gained by the inclusion of additional objectively quantifiable data sources (e.g., health care utilisation data, smartphone-collected sensor data). Fourth, most prevention approaches involve the delivery of a single intervention or program. There is rarely an option to identify who is most likely to respond to different interventions, or for a tiered approach whereby individuals for whom one approach has not been effective, may try a different approach afterwards. This is potentially a missed opportunity given that not everyone will respond to a single, or the first-line, approach. Finally, CBT has, until recently, been delivered to young people generally over 6–12 sessions using face-to-face approaches which are challenging to scale due to the expertise and cost required to deliver such programs. Digital approaches may be easier to scale up; however, in a review of 146 randomised controlled prevention trials, only two involved digitally-delivered interventions (18). In light of these limitations, there is scope to improve on and extend the empirical base to understand the effects of depression prevention programs.

The Current Study

The Future Proofing Study has been designed to overcome these challenges by conducting a large school-based RCT testing whether depression can be prevented using a CBT program delivered by smartphone application. Research evidence to date suggests that digital CBT programs for depression delivered in schools can be effective in reducing and preventing depression (23). Partnering in the delivery of prevention programs is a natural fit for schools, which are increasingly seen as settings in which to foster emotional skill development alongside with traditional academic learning. Mental ill health is associated with poorer academic performance (24), and depressed young people miss an average of 22 days of school per year (6), making mental wellbeing in line with the goals of educators and school administrators. Importantly, schools provide access to large student samples which are more representative of the general population than other settings, such as mental health services or medical practices. In addition, the potential to deliver a prevention program to all students universally may help reduce the stigma often associated with the delivery of mental health interventions only to those with symptoms (13).

This trial will involve up to 10,000 Australian school students who will be followed up for five years. A sample of this size provides appropriate statistical power to definitively and precisely evaluate effectiveness, as well as provide the opportunity to conduct robust moderation and mediation analyses. All study protocols will be delivered to participants digitally. The intervention/s will be delivered by smartphone application and study surveys can be completed on any Internet-enabled device (smartphone, tablet, laptop). This delivery method overcomes the aforementioned challenges associated with delivery by trained health professionals or teachers. It is simply too costly to train enough of the workforce to deliver CBT prevention programs either individually or in small groups to make an impact at a population level. Delivering interventions digitally is more cost effective than face-to-face therapies (25). Moreover, programs using smartphone app for delivery does not compromise effectiveness (26). Testing smartphone apps represents a major innovation of this project, with most evaluation of digital programs focusing on internet-based approaches (27). Digital technologies, and specifically smartphones, are now considered a key strategy to overcome the barriers inhibiting access to mental health programs. This mode of delivery is appropriate given young people show a preference for digital technologies when given the choice, due to the anonymity, convenience and privacy (12, 28).

This SPIRIT-compliant protocol describes the methodology for a multi-stage, two-arm, parallel group RCT. At Stage I, a cluster RCT will be undertaken to evaluate the effectiveness of SPARX, which is an app-based CBT gaming intervention (29). This intervention has been tested in the Australian adolescent population and was found to be effective for the prevention of depression (23). In the current trial, SPARX it will be delivered to Year 8 students (who are typically aged 12–14 years). The control group will consist of Year 8 students of an equivalent demographic who will not receive any intervention. Participants across both arms will be followed longitudinally over five years. The choice to use a non-active (treatment-as-usual) comparator was made to avoid placing unnecessary burden on participants and schools who would have been required to provide an attention-matched placebo intervention—an intervention that would have required the equivalent level of effort as that of the active condition—that would most likely provide no benefit. Other candidate programs that may have benefited other outcomes such as lifestyle or nutrition programs could not serve as a fair control condition because of their likely impact on mental health. At this Stage II (12 months following baseline), an individual-level RCT will be conducted within the intervention arm. Individuals who show elevated symptoms of depression at 12 months will be to be randomised to an intervention condition that receives a second CBT app which targets sleep difficulties (Sleep Ninja) or a control condition. Sleep Ninja focuses on sleep disturbances and delivers Cognitive Behaviour Therapy for Insomnia (CBT-I). An intervention targeting sleep difficulties was selected because sleep disturbance is both a risk factor and symptom of depression, with efficacy supporting its use in the prevention and treatment of depression in both adults and adolescents (28, 30–32).

Aims

This trial will investigate the impact of two mental health apps on adolescent depression. The first objective of this trial is to determine whether delivering a CBT program via smartphone application can prevent high school students in Year 8 from developing depressive symptoms compared to students in the control arm. This study will assess the effect of the program on a range of secondary outcomes which include mental health symptoms (distress, anxiety, insomnia) at post-intervention, 6, 12, 24, 36, 48 and 60 months. The effect of a second intervention on depression in those who show elevated depressive symptoms one year after using the first SPARX app will also be evaluated. Individual differences associated with response to the interventions will also be investigated, including demographics, history of mental illness, personality factors, social support, social media use, and gender identification, among a range of other variables (see Table 1). In the control group, these factors will be examined as possible predictors for the development of depressive symptoms and other mental health problems.

The second objective of the trial is to identify the predictors of mental health outcomes from digital behavioural and cognitive data collected using participants' smartphones (see Table 2 for details). Sensor data will be collected from smartphones to provide information about activity levels (accelerometry), movement through space (GPS) and voice patterns to generate a 'digital phenotype' for each participant (33–35). Digital phenotyping can be understood as using sensor data collected by personal devices to quantify the individual level human phenotype. One of the functions of profiling individuals

using digital phenotypes in this way is to make predictions about changes in mental health. The sensors in smartphones provide a non-intrusive, minimally burdensome and scalable method to investigate the effect of activity and behaviour on mental health. Evidence of the validity of this approach is beginning to emerge, with location variability now showing an association with depression outcomes in small studies of adults (36, 37). However, the use of digital phenotyping approaches in young people at scale has not yet been undertaken. For participants in the intervention group, this information will be combined with intervention usage data and other study measures to explore predictors of response and non-response to the intervention apps. In the control group, this information will be used to prospectively examine predictors of psychopathology.

The third objective of this trial is to evaluate the implementation of the proposed CBT-based depression prevention program in the school environment. The data that will be collected to meet this objective will combine intervention usage with qualitative and quantitative methods. A protocol for the process evaluation of the program implementation will be detailed elsewhere.

Three different forms of data will be collected from this trial to address each of the objectives: self-report questionnaires (first objective), smartphone recorded sensor data and ecological momentary assessment data (second objective), and usage and implementation data (third objective). This study will also utilise health and education datasets as a fourth form of data that will be linked to the main study datasets one year after the trial commences, and then periodically until the trial is complete. Linked data will include academic outcomes, health service utilisation, infant development, alongside births and deaths records. Linkage will allow for a more complete picture of the education outcomes and the health of the sample and a cost-effectiveness analysis.

The first form of data (self-report questionnaire data) will be reported in the primary outcome paper and limited to primary and secondary outcomes. All other data types will be reported and published separately.

Hypotheses

The primary hypothesis is that participants allocated to the intervention arm will show a smaller increase in depressive symptoms relative to those in the control group at 12 months measured using the self-report Patient Health Questionnaire—Adolescent Version (PHQ-A; 38). Secondary hypotheses are that depressive symptoms will be lower (e.g., show a smaller increase) in the intervention arm relative to the control arm at other assessment points (post-intervention, 6, 24, 36, 48 and 60 months). Using the pre-defined clinical cut-offs on the PHQ-A will also allow a comparison between likely cases of depression across the two arms. It is hypothesised that the percentage of depression cases will be lower in the intervention arm relative to the control arm at post-intervention, 6, 12, 24, 36, 48 and 60 months. Secondary mental health outcomes are psychological distress, anxiety, and insomnia, and are expected to be lower in the intervention arm relative to the control, at post-intervention, 12, 24, 36, 48 and 60 months. Additional outcomes including suicidality, substance use and sleep quality will be reported separately.

For participants in the intervention condition who show clinically significant depressive symptoms at 12 months (PHQ-A score > 10) and who enter Stage II of the trial, it is hypothesised that those who receive the Sleep Ninja intervention will report greater decreases in symptoms of depression (primary outcome), anxiety and insomnia (secondary outcomes) from pre-to-post-intervention (Stage II) and from 12-to-24 months (after they receive the Sleep Ninja intervention), relative to the Stage II control arm.

Methods

Trial design

This study is a randomised-controlled, single-blind trial with two intervention stages, each consisting of two-parallel arms (intervention and inactive control) with 1:1 allocation at each stage. Cluster-randomisation will occur at the school-level for the first intervention stage, stratified by school size, school location, school student gender and socio-economic status (full details below in the randomization section).

The second intervention stage will occur directly after the 12-month follow-up and involve individual-level randomisation. Stage I of the trial will have eight measurement occasions: Baseline (pre-intervention); Stage I post-intervention (immediately after completing the Stage I intervention, 6 weeks after baseline); and 6-, 12- (primary endpoint), 24-, 36-, 48- and 60-month follow-up. The follow-up period is time since baseline. For the subset of participants involved in Stage II of the trial, an additional measurement occasion will take place 6 weeks after the Stage II intervention (Stage II post-intervention). The trial will take place in three recruitment waves. It is planned that Wave 1 will commence in Term 3 2019 (July 2019), Wave 2 will commence in Term 2 2020 (April 2020), and Wave 3 will commence in Term 3 2020 (July 2020). The Stage I intervention phase (SPARX) will last 6 weeks and have a follow-up period of five years. The Stage II intervention phase (Sleep Ninja) will last six weeks, occurs 12 months after Stage I (SPARX), and have a follow-up period of up to four years.

Data from each assessment point will be used to determine safety and whether any modifications to the trial protocol are required (see analysis section for details). A concurrent implementation process evaluation will be conducted, and study processes may be modified to better address the needs of the school context. Any change will be reflected in the study protocol and registration (ACTRN12619000855123).

Setting

This trial will be conducted in approximately 200 schools located predominantly in New South Wales, Australia. The school system in Australia is divided across into main school types: government schools, independent schools and Catholic schools. Schools from each of these groups will be recruited into the study. Schools across metropolitan, regional and rural locations will be invited to participate in order to generate a demographically representative sample. Similarly, schools with differing indices of community socio-educational advantage (ICSEA) will be targeted. If recruitment targets are not met, schools in

Australian states outside of New South Wales will be invited to participate. Recruitment is broadly aligned to the academic calendar, with delivery occurring during Term 2 (April) and Term 3 (July) such that assessment sessions take place during the school calendar year and risk issues can be addressed without delay. The first trial recruitment wave will be conducted primarily in independent schools in metropolitan Sydney and the Central Coast for reasons of convenience. The second and third waves will include both metropolitan and non-metropolitan sites, and include government, Catholic and independent schools. Trial management will take place at the Black Dog Institute, a translational research institute located in Sydney, Australia that is affiliated with the University of New South Wales (UNSW).

Participants

As this is a universal prevention study, there are no exclusion criteria. All adolescents enrolled in Year 8 at each participating school are eligible to participate in the trial if they have a smartphone with iOS or Android operating system and an active mobile phone number. The usual age range of students in Year 8 is 12–14 years, although Year 8 students outside of this age range will not be excluded.

Research Team Roles and Responsibilities

Several committees have been established to support the overall trial governance, each with their own charter and roles. The most senior committee is the Internal Steering Committee comprising senior investigators and staff located at the Black Dog Institute (HC, AWS, KH, KM, AS, NC). This group meets fortnightly and is responsible for guidance and decision making about overall trial design, intellectual contribution to the scientific quality and strategy, oversight of trial progress and compliance with good clinical practice. The day-to-day trial leadership team meets weekly (AWS, KH, KM), and lead the operational aspects of the trial including management of teams responsible for recruitment and consent (AWS, KM, LJ), data management data privacy and security (KH). A larger Investigator Committee comprises 20 investigators and collaborators on the project who meet annually to provide strategic guidance and input.

Youth and Public Involvement

Young people, parents and consumers have been consulted in the preparation phase for the current study. Recruitment processes and study materials were reviewed by the Lived Experience Advisory Panel at the Black Dog Institute, and a local school's Parents & Citizens (P&C) Association, with changes made based on suggestions provided. Young people both through the Black Dog Institute network, and a local mental health service youth reference group have been involved in reviewing all the questionnaires and informed the wording of questionnaires and explanation of concepts using 'explainers' to aid clarity. The two interventions have been subject to youth feedback in previous trials and iterations of the programs. For example, young people who have previously used the SPARX intervention provided acceptability feedback in a previous trial (23) while Sleep Ninja was developed in collaboration with young people, and a sample of 50 young people provided feedback about the intervention in acceptability surveys and interviews (32).

Recruitment

A flowchart outlining recruitment into the trial, randomisation at trial Stages I and II, the study timeline and participation is outlined in Figure 1. Approximately 200 schools will be recruited to participate in the trial via several recruitment pathways. The characteristics of the sample will be compared to school census data to determine representativeness. Schools will be recruited in three waves over two years: one wave for commencement of the trial in Term 3, 2019 (12 schools), then another two waves in Terms 2 and 3, 2020 respectively (approximately 188 schools).

School principals will be contacted in the first instance. They will be given brief information sheets containing trial details. Schools will have the opportunity to hear more information over the phone or via video conference. Those who wish to take part will be instructed to sign and return a letter of support. Upon receipt of the letter, schools will be accepted into the trial. In some instances, schools may approach the Black Dog Institute by submitting an expression of interest via a link on the study website (<http://www.futureproofing.org.au>), or by directly contacting staff at the Black Dog Institute.

After schools have consented, parents and students will be informed about the trial by their school via email or newsletter, depending on school preferences. A study invitation will be circulated to parents. This will include a link to an electronic information sheet and consent form. Here, parents and students can read information about the study, download a PDF copy of the study information sheet and provide electronic consent. Fourteen days following consent form distribution, schools with capacity will make follow-up phone calls to parents who have not submitted a response. The option to provide verbal consent will be given to these parents. Students whose parents have consented to their participation will be given the opportunity will provide consent electronically at the start of the first study session. On the consent form, participants indicate whether they consent to different components of the study, including the sensor data collection, the linkage component and whether they provide permission for their data to be included in a data repository for use by other researchers. Consent forms are available from the study website (www.futureproofing.org.au). No biological specimens will be collected. Students who do not provide active consent cannot participate in the trial. If a participant withdraws from the study, they are asked to indicate whether they give permission for their already provided data to be used by the study team. All recruitment strategies have ethical approval.

The Black Dog Institute has existing relationships with 100 schools in New South Wales, both through previous research studies and via delivery of community mental health literacy programs to high school students. These relationships will be leveraged to support school recruitment. To maximise school consent rates, potential schools will be approached at least one school term in advance of study implementation. This means that schools can schedule study activities well in advance. Schools who participate in Wave 1 of the study will also be eligible to participate in Waves 2 or 3 (as these occur in the subsequent

year and therefore involve a different cohort). As a universal study, schools will be encouraged to organise data collection assessment sessions as whole-of-classroom activities. To maximise parental consent rates, an information video specifically designed for parents will be made available, and parents have the option to directly contact the research team for further study information.

Interventions

This trial will involve two interventions: a universal prevention intervention (Stage I; SPARX) and an indicated prevention intervention (Stage II; Sleep Ninja) for participants showing elevated depressive symptoms 12-months after receiving the universal intervention. Participants can begin or continue with any existing treatments during the study, including antidepressants and psychotherapy, and may receive additional support from health professionals if required.

SPARX

SPARX is a gaming intervention that was developed as a treatment program for mild-moderate depression (29). In its original form as a computerised game, research has shown SPARX to be effective in the treatment of adolescent depression, and equally effective as treatment as usual, which primarily involved face-to-face psychotherapy with a mental health professional (29). This trial will use SPARX-Future Proofing (SPARX-FP; referred to as SPARX in all other sections of the document), which is an adaptation of SPARX, to make it suitable for the *prevention* of depression, for delivery via mobile phone application. The content is the same, but instead of a focus on existing symptoms and depression, terminology has been updated to focus on times when the participants have felt 'down or stressed' and included skills and strategies reframed to focus on dealing with problems as they arise. The prevention version of SPARX has been trialled in an Australian sample of 540 final year secondary school students. This trial showed that, relative to a control group, those who used the program showed lower symptoms of depression in the lead up to final school exams (23). The move to a mobile phone application from a web-based program was informed by this previous study finding technological barriers and participant preferences for delivery via smartphone.

SPARX consists of seven modules (levels) which cover: finding hope, being active, dealing with strong emotions, problem solving, recognising unhelpful thoughts, challenging unhelpful thoughts and bringing it all together. Each module takes approximately 20 minutes to complete. Each module is designed to be completed on separate days, and users are encouraged to do one to two modules per week. Participants will have 6 weeks to complete the SPARX program on their personal smartphones, in their own time, or in class time if schools wish to offer it. Skills learnt through the SPARX program include emotion identification, emotion regulation, behavioural activation (being active), recognising and challenging unhelpful thoughts, and practical problem solving. This intervention is delivered in a game format, where participants begin by choosing a personalised avatar. The gaming component sees the participant undertaking skill-building challenges in the context of a fantasy world, where the aim is to restore balance in a world taken over by gloomy, negative, automatic thoughts. The gaming component is supplemented by direct instruction, education and activities (or homework) provided by a 'guide' avatar, who relates the content of the program to users' real life experiences.

Sleep Ninja

Sleep Ninja is a smartphone app based on CBT-I, which is the gold standard, evidence-based treatment for insomnia. Sleep Ninja was developed in collaboration with young people, and has been tested for feasibility, acceptability and preliminary effects on insomnia and depression (32, 39). The intervention involves psychoeducation, stimulus control, sleep-focused cognitive therapy, basic sleep hygiene and behavioural activation.

Sleep Ninja takes the form of a chat-bot. When the app is open, conversation messages appear on the phone screen from the Sleep Ninja character (who acts as a sleep coach). Users interact with the Sleep Ninja by selecting from pre-determined responses, to which the Sleep Ninja is programmed to respond. Users progress through six levels or "belts", starting at a white belt and working their way through to black belt status by the end of the program. To level up to the next belt, users must complete within a one-week period: (i) one training session (which takes approximately 5–10 minutes) and; (ii) three nights of sleep tracking using a sleep diary within the app. At the completion of each level, users are provided with a brief report card and graph summarising their self-reported sleep over that period. There are additional optional app features that users can access which include a meditation recording, extra sleep information, quick sleep tips and an easy way to send app-information to parents. If participants agree, the app sends a reminder each morning to enter sleep tracking from the night before, and a reminder an hour before bed to begin winding down in preparation for sleep, and a final reminder at bedtime for users to go to bed. Participants will have 6 weeks to complete the Sleep Ninja program in their own time on their personal smartphones.

Strategies to enhance and monitor adherence

To enhance adherence, digital notifications and reminders will be sent to participants reminding them to activate, visit and use the app/s as intended. Schools will provide class time to complete some of the activities contained in the apps (minimum of four class sessions), with the rest set for homework. School staff will provide verbal reminders and encouragement to students to use the Future Proofing App and for intervention arm schools, the first intervention app (Stage I; SPARX) during the intervention implementation period. There will also be a \$20 one-off incentive to cover any connectivity and data costs delivered at post-intervention to participants in both arms (6 weeks), which may serve to improve adherence. Adherence to the intervention app/s will be monitored by evaluating usage data that is collected automatically by the apps while they are active. Usage data is available at the individual level and consists of when and how many times users enter the apps, how long they spend in the app, how many modules they complete, and for Sleep Ninja only, any sleep tracking data that is inputted by the user.

Outcomes

The administration schedule for each of the assessment measures described below is provided in Table 1.

Primary outcome measure

Patient Health Questionnaire—Adolescent Version

The primary outcome measure is depression severity measured at 12-months post-baseline using the Patient Health Questionnaire for Adolescents (PHQ-A; 38). PHQ-A depression scores will also be collected at baseline, 6-weeks post-intervention (Stages I and II), and at 6, 24, 36, 48, and 60-months post-baseline. The PHQ-A is a validated modification of the PHQ-9 for adolescents, a 9-item self-administered depression severity screening and diagnostic tool based on DSM-IV criteria. The scale assesses the frequency of occurrence of nine depression symptom criteria during the previous two weeks, with items rated on a 4-point scale ranging from 0 (Not at all) to 3 (Nearly every day). Total scale scores on the PHQ-A depression scale can range from 0–27, with higher scores reflecting more severe depressive symptoms. The accepted clinical cut-offs are as follows: a score of 0–4 indicates nil to minimal symptoms, 5–9 indicates mild symptoms, 10–14 indicates moderate symptoms, 15–19 indicates moderately severe symptoms and 20–27 indicates severe symptoms. The psychometric quality of the PHQ-9 and PHQ-A is well established (38, 40).

Secondary outcome measures

Secondary outcomes will be monitored for the duration of the trial at baseline, 6-week post-intervention (Stages I and II) and the 12-, 24-, 36-, 48- and 60-month follow up assessments (see Table 1 for the specific measurement occasions corresponding to each outcome measure). The presence of mental health symptoms will be determined by established cut-offs relevant to each scale. Analyses will be comparisons of mean scores between the intervention and control arms at each time point. The following measures will be used to assess the secondary outcomes:

Distress Questionnaire–5

The Distress Questionnaire–5 (DQ–5; 41) is a 5-item brief screening tool for identifying general psychological distress. Respondents rate each item on a 5-point scale, ranging from 1 (Never) to 5 (Always). The total scores on the scale range from 5 to 25, with higher scores indicating greater psychological distress. The scale has demonstrated optimal to high internal consistency and convergent validity, and has been found to be more accurate in identifying psychological distress than some other commonly used screeners (41, 42).

Spence Children’s Anxiety Scale Short-Form (including Generalized Anxiety and Social Phobia Subscales)

The Spence Children’s Anxiety Scale Short-Form (SCAS-SF) is an 8-item brief measure of anxiety for children and adolescents, based on the SCAS (43). The SCAS was designed to measure the severity of children and adolescents’ anxiety symptoms based broadly on DSM-IV criteria for anxiety disorders (44). Respondents rate the degree to which they experience each symptom on a 4-point frequency scale, ranging from 0 (Never) to 3 (Always). Total scale scores on the 8-item SCAS-SF can range from 0–24, with higher scores reflecting greater anxiety. The SCAS Social Phobia (6 items; total scores = 0–18) and Generalised Anxiety (6 items; total scores = 0–18) subscales will also be administered in this trial to provide specific validated measures of these common anxiety disorders in adolescents. The SCAS has demonstrated high internal consistency and satisfactory test-retest reliability (43). The SCAS has also been reported to show both convergent and divergent validity (45).

Insomnia Severity Index

The Insomnia Severity Index (ISI) is a psychometrically sound, 7-item self-report measure of insomnia symptoms over the previous two weeks (46). Responses are reported on a Likert scale ranging from 0 (Not at all) to 4 (Very), producing total scores of 0 to 28. Cut-off scores are as follows: 0–7 reflects no clinically significant insomnia, 8–14 indicates subthreshold insomnia, 15–21 suggests moderate severity insomnia, and 22–28 indicates severe insomnia. The ISI was designed for use in adults but has been widely administered to, and validated in, adolescent samples (47, 48).

Additional outcome measures, potential mediators and risk factors

The following additional outcome measures will be examined and reported separately from the primary and secondary outcomes. Also included are variables which may mediate outcomes in the intervention arms. Potential risk factors for the development and maintenance of mental health disorders are also included (see Table 1 for the specific measurement occasions corresponding to each measure).

Suicidal Ideation Attributes Scale

The Suicidal Ideation Attributes Scale (SIDAS) is a recently developed brief measure of suicidal ideation severity in the past month. Using a general population sample, it has demonstrated high internal consistency and good convergent validity (49). The SIDAS consists of five questions pertaining to frequency of thoughts, controllability of thoughts, closeness to attempt, level of distress associated with the thoughts, and impact on daily functioning. Each item is assessed on a 10-point scale. Endorsement of item 3 (how close the individual has come to making an attempt in the last month) will trigger the trial's risk management procedure (see details below). A total scale score is calculated by summing item scores and can range from 0 to 50. Higher scores indicate higher levels of suicidal ideation severity.

Youth Risk Behaviour Survey

The Youth Risk Behaviour Survey (YRBS) was designed to assess health risk behaviours among secondary school students. Three items from the YRBS will be used in the current trial to assess suicide-related behaviours (thoughts, plans and attempts) in the past 12 months, for which participants indicate a 'Yes' or 'No' response. Studies have shown that the suicidality items demonstrate both substantial reliability (50) and good convergent and divergent validity in a secondary school sample (51).

Strengths and Difficulties Questionnaire

The Strengths and Difficulties Questionnaire (SDQ; 52) is a widely used behavioural screening questionnaire for 4 to 17 year-old children and adolescents. It consists of 25 items divided between 5 sub-scales: Emotional Symptoms, Conduct Problems, Hyperactivity/Inattention, Peer Relationship Problems and Prosocial Behaviour. Respondents indicate on a 3-point Likert scale the extent to which each item applies to them, using the options 0 (Not true), 1 (Somewhat true), or 2 (Certainly true). Total scale scores on each of the sub-scales can range from 0–10. A higher score is indicative of more problems for all subscales, except for the prosocial scale, where higher scores correspond to fewer difficulties in prosocial behaviour. The SDQ has demonstrated good internal consistency across studies (53, 54).

Screen for Disordered Eating

The Screen for Disordered Eating (SDE; 55) was recently developed to screen for eating disorders. The SDE comprises five items, to which respondents indicate whether they experience any disordered eating on a dichotomous scale (Yes or No). An individual is screened as positive if he/she endorses two or more items. In the primary care setting, this measure has demonstrated good discriminative accuracy (55).

Self-Harm Questionnaire

The Self-Harm Questionnaire (SHQ; 56) was designed to improve identification of self-harm in adolescents. The complete questionnaire consists of three screening questions enquiring about any past incidents of self-harming behaviour or thinking, followed by 12 additional questions that are only presented to adolescents reporting previous self-harm. Among a sample of psychiatric service in- and outpatients, the SHQ has demonstrated good concurrent and predictive validity (56). To assess self-harm prevalence and frequency in the current trial, only screening item 3 will be administered in order to assess past episodes of self-harm ('Have you ever actually harmed yourself on purpose? For example, have you ever cut yourself or taken an overdose and it was not an accident?'). Participants respond to this item on a 4-point scale of 'No', 'Yes, once', 'Yes, two, three or four times' and 'Yes, five or more times'. This item will allow the assessment of the prevalence of self-harm and its frequency.

Alcohol Use and Substance Use Questionnaires

The Alcohol Use questionnaire included in this trial was originally adapted from the School Health and Alcohol Harm Reduction Project (57) and used the Climate Schools Projects (58), which are Australian school-based trials aimed at reducing alcohol and cannabis use. The questionnaire includes a standard drink diagram contains nine items assessing age of first use, alongside frequency and quantity of alcohol use. An additional questionnaire assessing Other Substance Use was adapted from the Australian Institute of Health and Welfare 2007 National Drug Strategy Household Survey. The questionnaire contains five items assessing recency of substance use, with a specific focus on cannabis (additional items on age of first use and frequency of use), and also assesses tobacco, amphetamine, ecstasy, hallucinogens, sedatives and inhalant use.

Adolescent Psychotic-Like Symptom Screener

The Adolescent Psychotic-Like Symptom Screener (APSS; 59) is a 7-item measure designed to identify people who are at increased risk of future clinical psychotic disorder. In this study, only three items will be administered to assess paranoia, auditory and visual hallucinations. For each question, there are 3 possible responses: 'Yes, definitely', 'Maybe', 'No, never'. Responses are scored 1, 0.5, and 0 respectively. Higher total scale scores on this measure are indicative of greater psychotic-like symptoms. In the general population, this instrument has demonstrated good sensitivity and specificity in identifying young adolescents with psychotic-like experiences (59).

Pittsburgh Sleep Quality Index

The Pittsburgh Sleep Quality Index (PSQI; 60) was designed to assess sleep quality and disturbances over a 1-month interval. It consists of 19 items, which are used to compute seven component scores: sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleep medications and daytime dysfunction. Each item is weighted equally on a scale that ranges from 0 (no difficulty) to 3 (severe difficulty) scale. The seven component scores are then summed to yield a global score, ranging from 0 to 21; higher scores indicate worse sleep quality. The PSQI has demonstrated acceptable to good internal homogeneity, test-retest reliability and convergent validity across studies (60–62).

Child Health Utility 9D

The Child Health Utility 9D (CHU-9D; 63) is a 9 dimension generic preference-based measure designed to assess child and adolescent health related quality of life and suitable for application in economic evaluation. Current child/adolescent health related quality of life is assessed across the domains of worry, sadness, pain, tiredness, annoyance, school, sleep, daily routine and activities. Each dimension is rated on a 5-point response scale ranging from 'no' to 'severe impairment'. Responses are then converted to utilities on the 0–1 dead to full health quality adjusted life years scale (QALY) using a preference weighted scoring algorithm (64). Previous validation studies with adolescents from the community and mental health services have demonstrated that the self-complete instrument has acceptable internal consistency and convergent validity for children and adolescents age 7 to 17 years (65–67).

Short Warwick-Edinburgh Mental Wellbeing Scale

The Short Warwick-Edinburgh Mental Wellbeing Scale (SWEMWBS; 68) is a shortened 7-item version of the 14-item Warwick-Edinburgh Mental Well-being Scale (WEMWBS), which was developed to assess mental wellbeing in the general population. The SWEMWBS consists of seven statements about thoughts and feelings over the past two weeks. Ratings are made on a 5-point Likert scale (1 = None of the time, 2 = Rarely, 3 = Some of the time, 4 = Often, 5 = All of the time). Total scale scores are calculated by summing item scores and transforming the total score using a conversion table. Total scores can range from 7 to 35. A higher score indicates a higher level of mental well-being. The SWEMWBS has demonstrated adequate reliability and validity across studies (69, 70).

Demographic Information

At the baseline assessment, participants will be asked to provide their date of birth, postcode, country of birth, language spoken at home, who they live with at home, Aboriginal and Torres Strait Islander status, and socioeconomic status (71).

Height and Weight

Self-reported height (cm) and weight (kg) will be provided.

Big Five Personality Inventory

The Big Five Personality Inventory (BFI-10; 72) is a 10-item scale measuring the Big Five personality traits: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness. The scale was developed based on the 44-item Big Five Inventory (BFI-44; 73). Respondents indicate their level of agreement with each described trait on a 5-point scale, ranging from 1 = 'Disagree strongly' to 5 = 'Agree strongly'. Total sub-scale scores are calculated by summing item scores (two items per sub-scale). Subscale scores can range from 2 to 10. The BFI-10 has demonstrated acceptable test-retest reliability and good construct validity across studies (72, 74).

School Connectedness

School connectedness will be assessed using questionnaire items developed by the Organisation for Economic Co-operation and Development (OECD) Programme for International Student Assessment (75). Six items will be administered, which are rated on a four-point scale from 1 (Strongly agree) through to 4 (Strongly disagree). Total scale scores can range from 6 to 24, with higher scores reflecting greater school connectedness.

Schuster Social Support Scale

The Schuster Social Support Scale (SSSS; 76) is a 15-item measure of positive and negative interactions with family, friends and spouse. In the current study, 10 items will be administered to assess interactions with family and friends only. Each item is rated on a 4-point scale ranging from 0 (Never) through to 3 (Often). Scores are interpreted per category, for friends and family, with higher scores on the supportive interactions scales indicative of more supportive interactions, and higher scores on the negative interactions scales indicate more negative interactions.

Maladaptive Facebook Usage Scale (Adapted)

The Maladaptive Facebook Usage Scale (77) is a 7-item measure of maladaptive Facebook usage, which assesses an individual's tendency to undertake negative social evaluations and social comparisons when they use Facebook. This scale has demonstrated good test-retest reliability and convergent validity (77). In this trial, the scale has been adapted to incorporate social media more broadly. A screening item has been added asking respondents to nominate which social media platforms they use at least once per week. Options include: Facebook, Instagram, Snapchat, Tumblr, Twitter, YouTube, Reddit, Other, and "I

don't use social media". The seven items of the Maladaptive Facebook Usage Scale were adapted to apply to any social media platform, e.g. 'I tend to read the social media status updates of others to see if they are feeling the way I am,' and 'Reading the social media status updates of others tends to make me feel down on myself'. Items are rated on a 1 (Strongly Disagree) to 7 (Strongly Agree) response scale. Total scale scores can range from 7 to 49. Higher scores indicate greater tendencies to seek online social comparisons and negative evaluations.

Gender Identification and Sexual Identification/Preferences

Gender identification (two items: sex at birth and current gender identity), and sexual identification and preferences (one item for each) will be examined at baseline and annually between 12- and 60-month follow up.

Romantic Relationships

Two items will be presented at multiple time-points to assess romantic relationships. These items will ascertain the number of special or important romantic relationships in the past year, as well as the number of break-ups in the past 12 months.

Sexual Behaviour

A series of questions will be presented at the 24-month follow-up when participants are in Year 10 (aged 15–16 years) to assess sexual behaviour, taken from the National Survey of Australian Secondary Students and Sexual Health (78). These items will assess experience and age of first "making out", intimate genital touching and sex, as well as number of people they have had sex with during the past year and frequency of condom use. This section will include an initial gating item assessing history of intimate sexual contact. Those who do not endorse this item will not receive the other sexual behaviour items.

Trauma Behavioural Risk Factor Surveillance System—Adverse Childhood Experience Module

The Trauma Behavioural Risk Factor Surveillance System—Adverse Childhood Experience Module (BRFSS-ACE; 79) consists of 11 items that assess exposure to nine types of adverse childhood experiences in the first 18 years of life, including: verbal abuse, physical abuse, sexual abuse, household mental illness, household alcohol abuse, household drug abuse, domestic violence, parental separation/divorce, and incarcerated family members. The responses are dichotomised to indicate exposure and summed to create an ACE score ranging from 1 to 8 for each subdomain, with higher scores indicating greater exposure. Previous studies have demonstrated that this instrument has adequate internal consistency and validity (80, 81). In this trial, the complete BRFSS-ACE will be administered at 48-month follow up when participants are near 18 years of age. A modified version of the questionnaire consisting of 8 items will be administered at baseline when participants are approximately 13 years old. This modified scale excludes items on physical and sexual abuse, and includes additional items assessing out of home or foster care and feelings of endangerment or physical harm.

History of Mental Health Diagnosis

One item will assess lifetime history of diagnosed mental health problems (Major Depression, Social Anxiety Disorder, Generalised Anxiety Disorder, Obsessive Compulsive Disorder, Panic Disorder, Alcohol Use Disorder, Substance Use Disorder, Attention Deficit Hyperactivity Disorder, Post Traumatic Stress Disorder, Schizophrenia/Psychosis).

Hospitalisation

One item will assess hospitalisation in the previous 12 months and will ask participants to distinguish between hospitalisation for physical and mental health problems.

History of Disability Diagnosis

One item will assess lifetime history of diagnosed disability (Autism or Asperger's Syndrome, Intellectual Disability, Specific Learning Disability, Tourette Syndrome, Cerebral Palsy, Acquired Brain Injury, Other Neurological Disability, Hearing Impairment, Visual Impairment).

Pubertal Development

Open ended questions for females and males will be used to assess pubertal development (e.g. 'at what age did you get your first period/did your voice begin to break?'). Participants will also be provided with line drawings of stages of pubertal development (Tanner Stages) and asked to rate their current stage of physical development against these images.

Bullying

Bullying items were drawn from previous school-based trials. The three items examine whether participants have been bullied, cyber bullied, or bullied others in the past 12 months. Items are responded to on a 5-point frequency scale from 1 (Not at all) to 5 (Most days).

App Use and Feedback Surveys

App use and feedback surveys for both the Future Proofing App and SPARX app will be administered at the post-assessment time point. The Future Proofing App Survey was developed for this study and has 9 items which assesses app task preferences and reasons for discontinued use. The SPARX Feedback Survey is an 11-item survey that has been used previously to assess SPARX use (23), app acceptability and asks participants to select any skills they learnt from using the app.

Smartphone collected data

Additional data will be collected directly via smartphones through a study app developed for this study (the Future Proofing App), as this data cannot be collected from the self-report questionnaires outlined above. Three forms of data will be collected, which include actively collected data, passively collected data, and app usage data. Actively collected data will assess self-reported mood ratings, voice samples and cognitive tasks. This form of data will overcome biases in retrospective reporting of mood and allow for an investigation of whether changes in voice are related to changes in mental health states. Cognitive tasks include two measures of executive function presented in affective and neutral contexts. Specifically, a measure of affective shifting (82) and the Affective Digit Span (83) and a typing speed task, which will allow for a mechanistic assessment of the way cognitive function affects psychological outcomes in adolescents.

Passively collected smartphone data will involve the collection of location data (GPS) and movement data (accelerometry), which will be used to investigate whether mental health changes can be predicted from location and activity data. Smartphone sensor data will be collected during the first three months of the study, and at each annual assessment point for three month periods. App usage data will allow for assessment of app use and completion, and time spent using the study app/s.

Linked data

Data relating to participants' academic outcomes, physical health, utilisation of health services, infant development, births and deaths will be linked from extant Australian government administrative datasets by an authorised agency of the New South Wales Ministry of Health, the Centre for Health Record Linkage (CheReL, www.cherel.org.au). Informed consent from each participant will be obtained prior to linkage and databases will be linked using personal identifiers such as names, dates of births, addresses and hospital identification numbers with probabilistic methods. Linked data will be provided to the researchers in a de-identified manner and will be used to determine related health and other outcomes for the participant, including birth and perinatal data, educational outcomes in standardized curriculum-based tests and hospitalization and mortality outcomes. Linkage will occur within the first 12 to 24 months of the study (at the end of 2020) and will be updated periodically every two years until the trial concludes.

Procedure and participant timeline

Figure 1 shows participant timeline and flow. An initial preparatory phase and first school visit will solicit parental and student consent, ensure that electronic study questionnaires are accessible by consented students and that the appropriate study apps are installed on each student's device. Each subsequent school visit will coincide with an assessment occasion. During these visits, students will access study questionnaires via a secure online portal, accessible using their mobile number and a one-time password sent via SMS. Study personnel will attend these school visits to assist with technical issues and participants' questions. Following completion of baseline questionnaires, participants in the intervention condition will be instructed to commence the intervention both during class time and in their own time, and will receive notifications and reminders over the following six weeks to do so. Schools are required to schedule four 20 minute in-class sessions for intervention completion but may choose to hold additional sessions (up to seven). Participants in the control condition will also be provided equivalent class time to complete activities on the Future Proofing app. After each assessment occasion, schools will be encouraged to hold 'catch up' sessions for students who do not complete their assessments (e.g., who are absent from school that day). Students who do not complete assessments in school will also have the option to do this in their own time. Students will have two weeks following the scheduled assessment date to complete questionnaires at each time point, after which the assessment surveys will be locked and no longer accessible. During this two-week window, automated electronic reminders will be sent to students to prompt completion, if required.

Sample size

Separate sample size estimates were derived to determine the numbers of participants needed to meet the aims of each trial stage. Power was set at 0.80, $\alpha = .05$ (two tailed), and a correlation of 0.5 assumed between baseline and endpoint symptom scores. To account for possible clustering effects at Stage I (i.e., participants from the same school having prognostically relevant characteristics and outcomes more alike than between schools), a design effect (84) was calculated assuming an intraclass correlation coefficient (ICC) of .03 (based on previous school-based studies; 85), and a cluster size of 50 students, yielding a design effect of 2.47. In order to detect a mean standardized difference of 0.3 between conditions at Stage I, a sample of 870 participants would be needed (435 per arm). Allowing for up to 30% attrition, recruitment of 1244 participants would be required (622 per arm).

Using similar parameters for Stage II (without a design effect) would require recruitment of 368 participants in total to detect the same effect size. However, participants in this stage must meet eligibility requirements regarding elevated symptoms and failure to respond to the first stage intervention. They must also be willing to participate. Accordingly, a scenario involving recruitment of 10,000 participants at Stage I was considered, assuming 7,000 participants would have endpoint data available at 12 months (allowing for 30% attrition). Conservatively estimating 20% having PHQ-A scores >10, 25% showing a reduction

due to the Stage I intervention, and that 50% of those eligible to enrol at Stage II do so, 525 participants would enter this stage II, with endpoint data available for at least 368 (allowing for up to 30% attrition). This sample would enable the detection of differences in changes in symptomatology scores of comparable size to those in Stage I of the trial ($d = 0.29$).

Given that a target sample of 10,000 is substantial, if the numbers of participants are not sufficient to reach the Stage II target, supplementing this sample with additional young people specifically targeted based on PHQ-A scores that fall in the top 20% of population scores to ensure that this stage of the trial has comparable power as Stage I to detect an intervention effect will be considered. This will ensure that sufficient numbers are included at both stages of the trial. Should the target be met, Stage I of the trial will have much higher power to detect changes in symptom scores and would be adequately powered to detect differences in incidence of based on PHQ-A diagnosis criteria.

Exploratory analyses for moderating and mediating effects will be feasible at both Stages. Stage II will have power to detect small to medium mediation effects (86). Power to detect medium size effects (0.5 SD) in moderation analyses will be maintained for subgroups down to approximately 120 participants. Power will be enhanced in longitudinal modelling (growth curves) which use multiple occasions of measurement to estimate rates of change (86).

Randomization

For Stage I, cluster randomisation (at the school level) will be employed for administrative convenience, to avoid control condition contamination, and for the ecological validity of providing the intervention at the cluster level. Schools will be randomised after they are recruited into the trial. The trial statistician who is not involved in the day-to-day running of the trial will perform the randomisation, and the identity of the school will be concealed from the statistician.

Schools will be randomised with a 1:1 allocation, as per a computer-generated randomisation schedule. Balance between the trial arms will be achieved by stratifying based on school size, school location (metropolitan vs regional), school type (co-ed or gender-selective) and Index of Community Socio-Educational Advantage (ICSEA) level. For Stage II, individual-level randomisation with a 1:1 allocation using a computer-generated randomisation schedule stratified by gender and depression severity scores will be employed. This procedure is automated through the Black Dog Institute research platform. Permuted block randomisation will be used at both stages but will not be disclosed to ensure concealment. Allocation to arm is not directly communicated to schools. However, because the study has no control intervention, schools, participants and study operational staff will be aware of school allocation to the intervention arm. With the exception of the trial data manager, other individuals not involved in the day to day running of the trial will remain blind to allocation. All outcome assessments are conducted electronically and not subject to assessor bias. Unblinding at the conclusion of analysis will be performed by the trial data manager with authorisation from the Trial Steering Committee.

Participant risk management protocol

An independent Data and Safety Monitoring Committee (DSMC) has been established to monitor the quality of trial data and the safety of research participants. The DSMC will be responsible for safeguarding the interests of participants through regular monitoring, including participant safety and adverse events. This group will be responsible for monitoring the efficacy of the interventions being tested on primary outcomes, as well as the overall conduct of the study, including recruitment, protocol compliance, accuracy and completeness of data collection. This group will also provide recommendations with respect to continuing, modifying or terminating the trial, on the basis of feasibility or safety concerns, and will have access to unblinded data. This group can recommend the trial team be unblinded if there are serious safety concerns and meets every six months to review trial conduct. There is no anticipated harm from this trial, as it involves a population-based sample and evidence-based interventions. However, should there be any unanticipated harms, this will be monitored by the DSMC who will provide recommendations about care for affected participants.

All students involved in the trial will be assessed for suicidal thinking and behaviour at each assessment point, using the SIDAS and YRBS. If participants indicate serious suicidal thinking, plans or behaviour on study surveys, an alert will be triggered whereby the research team and school counsellor are notified immediately using a purpose-built study portal. School counsellors will then follow-up with students within 48 hours to offer immediate support, or refer on when necessary, which is recorded in the portal. The research team will monitor this portal and directly contact any counsellors who have not indicated follow-up with students. If students have changed or left schools, parents will be notified. At the 48-month assessment point when students are in their final year of school, they will also be asked about a history of sexual abuse. If students indicate an experience of sexual abuse, this same process of notifying the school counsellor will be followed, and schools will assume the duty of care for mandatory reporting requirements. This information will be communicated to relevant university and school ethics committees, as well as the DSMC following each assessment point.

Data collection and management

All research data collected in this trial will be stored using a unique participant Id code. A list of identifiable participant information associated with each Id code will be stored separately to the research data. The privacy, security and ownership of the research data will be maintained and re-identifiable data will not be stored or accessible by another organisation. Access, storage and transmission logs will be recorded and regularly reviewed for anomalies. Annual audits will be conducted by the trial data manager to ensure compliance with data security processes outlined by UNSW and Black Dog Institute.

The data that will be collected includes questionnaire data, mobile phone data, and linked data. Coded questionnaire outcome data will be stored securely on the Black Dog Institute online research platform until ready for export. The research platform is stored on UNSW servers and supported by enterprise access controls and 256-bit encryption or higher. Data will be exported from the research platform into Microsoft Excel following assessments so that it can be checked by the data manager for data quality and accuracy. After checking, data will be exported into appropriate statistical software for analysis. Data

collected by mobile phone apps will be encrypted and transmitted to a secure database hosted by Google Cloud Services in Australia. Data will be securely removed and transferred on a scheduled basis to UNSW servers by the trial data manager. Access to Google Cloud Services and the UNSW server will require authentication and will be restricted to the data manager and named members of the research team. The data manager will be responsible for extracting and securely transferring data to the research team. Only researchers whose analyses require access to the specific dataset collected from each questionnaire, app and linkage data source will be able to access those data. Linked datasets will be subject to the requirements set out by linking agencies for storage on UNSW servers using the above outlined procedures.

Analysis

Primary outcome

Analysis of the primary outcome will be undertaken using an intention-to-treat approach including all participants randomised regardless of intervention received, controlling for baseline differences when appropriate. The primary analysis will be conducted using planned contrasts comparing a change in depression scores on the PHQ-A from baseline to 12-months between the trial arms (SPARX intervention vs. control), using a mixed-effects model repeated measure analysis (MMRM). MMRM is preferred due to the ability of this approach to include participants with missing data without using discredited techniques such as last observation carried forward (87, 88). School will be included as a random effect to evaluate and accommodate clustering effects. Variables used in determining allocation balance will be evaluated and retained in analyses where they are significant or quasi-significant. An unconstrained variance-covariance matrix will model within-individual dependencies. Transformation of scores, including categorisation, may be undertaken to meet distributional assumptions and accommodate outliers.

Secondary and additional outcomes

Secondary and additional outcome analyses will involve contrasts comparing change on secondary (DQ-5, SCAS, ISI), and other (e.g., SIDAS) outcomes from baseline to other occasions of measurement, using an MMRM approach, as described above.

Additional analyses

Subsidiary complier analyses will be undertaken to compare individuals who complete the intervention relative to those who do not, in both trial stages. Regression models will be used to examine risk factors for symptoms of psychopathology across the study measurement points in the control arm. Mediation analyses will be explored using structural equation modelling. Machine learning approaches will be used to link multiple data types with outcomes to discover novel factors which predict outcomes in terms of specific disorders and their progression as well as investigate individual patterns of data that predict individual mental state or symptom trajectory. Cost-effectiveness analyses will be conducted at the primary endpoint and the final measurement point. Data will be linked to existing records and education and health outcomes will be reported.

Interim analyses

The percentage of participants meeting PHQ-A depression caseness, scores and reporting of suicidality at each assessment within both groups will be measured and reviewed by the DSMC in order to monitor relative levels of deterioration (accounting for baseline levels), for safety purposes.

Ethics and dissemination

This study has ethical approval from the University of New South Wales Human Research Ethics Committee (HC180836) and NSW Government State Education Research Applications Process Approval (SERAP 2019201) and has applied for approval from Catholic dioceses and the Tasmanian Education of Education. The trial is subject to annual progress reviews with these Ethics bodies. The trial is registered with the Australian New Zealand Clinical Trials Registry (ANZCTR12619000855123). All protocol amendments will be subject to approval by relevant ethics committees and listed on the ANZCTR registry. All trial findings will be presented in aggregate format so that no individual level data will be presented.

Public access to the full protocol can be granted from the authors on reasonable request. Access to participant-level dataset will be subject to governance processes set up around a data repository which will contain data from this study. This will be open for researchers to apply for access to.

Trial findings will be communicated using lay language and will be made available to participating schools for publication in school newsletters and/or school websites. Participants and parents will also be provided with these findings via email. Regular trial progress updates will be provided by the research team to schools for distribution through the school community, at their discretion. These findings will also be provided to other stakeholders in the wider community, including to government in policy documents, to school counsellor bodies, teacher groups and mental health groups. All findings will be provided in aggregate level. The results of the trial will be disseminated via peer-reviewed publications in scientific journals and conferences. No restrictions have been imposed on the dissemination of information by funders.

Discussion

The prevention of depression is critical if the burden of disease is to be reduced. This study will be the first to examine the prevention of depression in young people at scale, using easy to access, convenient and private mobile-phone applications. Efficacy of depression prevention has been established (17, 19). What the Future Proofing Study provides is a large definitive trial investigating the impact of delivering evidence-based prevention programs like CBT at scale, in real world settings. A large-scale trial such as this can provide insight into scaling and implementation processes, which can form the basis of an ongoing delivery and dissemination framework. The current study meets this need by examining a depression-prevention program, delivered by mobile phone applications, with access facilitated in the school environment which is where young people spend most of their time. This means that if successful, schools could become the setting in which access to these programs are rolled out continually, future proofing and inoculating all young people who go through the school system against mental illness.

There are several other features of this trial that make it novel. The use of a two-staged design, such that those who show symptoms of depression at the primary endpoint are randomised to receive a second intervention targeting insomnia, is innovative. This second intervention offers an alternative way to reduce risk for depression onset. The use of smartphone apps as the primary intervention delivery mechanism is new in this area (27, 89) and promises reduced costs and access barriers compared to traditional face-to-face and health professional-led programs. The inclusion of smartphone-collected objective data adds another dimension to the trial, allowing for machine learning approaches to explore whether activity and location data are associated with, or predictive of, changes in depression and other mental health symptoms. Moreover, linking this rich dataset to government education and health records will permit direct inferences to be made about the impact of mental illness on health and academic outcomes. Finally, the continuation of follow-up assessments for five years means that the long-term effects of the interventions on depression, anxiety, distress, insomnia and additional outcomes can be assessed. The results of the Future Proofing Study stand to provide a major contribution to the field of digital mental health prevention programs, and knowledge about the population-level impact of preventing depression using these methods.

Declarations

Ethics approval and consent to participate

This study has ethical approval from the University of New South Wales Human Research Ethics Committee (HC180836) and NSW Government State Education Research Applications Process Approval (SERAP 2019201). All schools, parents and participants must provide active consent to participate.

Consent for publication

Not applicable.

Availability of data and materials

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

Funding for this project came from a NHMRC Project Grant Awarded to Helen Christensen APP1120646, an Early-Mid Career Fellowship awarded to Aliza Werner-Seidler, an NHMRC Fellowship awarded to Philip Batterham APP1158707, and a Sir Henry Wellcome fellowship awarded to Susanne Schweizer. The funding bodies had no role in any aspect of the study design or this manuscript.

Author contributions

HC conceived of the study and secured funding, with assistance from AWS, ML, AC, MT, PB, NC, JO, GP, JR, RS, KS, and AJM. AWS and KH provided leadership in the overall management of the trial, including the trial design, technological requirements, and drafted the manuscript. ML provided leadership for the technical components of the study and AC provided leadership for the survey questionnaires. KM, LJ, AS, NC and IW provided research and operational support and contributed to study design processes. AJM and PB provided statistical and methodological leadership, and AJM is the trial statistician. BOD, JO, GP, JR, SS, RS, KS and KM provided expert input and guidance on aspects of the study. All authors contributed to, reviewed and refined the protocol and approved the final manuscript. All named authors adhere to the authorship guidelines of *Trials* and have agreed to publication. No professional writers have been involved.

Acknowledgements

The authors would like to acknowledge the many individuals who have been involved in the setting up of this study. Thank you to the schools, parents and young people who have provided their perspectives and guidance on the project. Thanks to members of the Black Dog Institute Education and Marketing team including Cameron Banks, Hiroko Fujimoto, Sarah Connor, Zoe Jenkins, Vivienne Zeaiter, Dean Winder and Fiona Treweweke. Sincere thanks to the IT team

including Jacinto Santamaria, Chaithra Madappa, Praveena Cheni Cheri, Toby Reid, Sam Townsend, Cesar Anonuevo, Jennifer Biggin, the Deakin team, and Victoria Carr. Thanks also to the many people who contributed their ideas and expertise to this project including Kathryn Woodcock, Drs Yael Perry, Simon Baker and Jane Hwang, and Professors Jennie Hudson, Sveta Venkatesh, and Maree Teesson.

Trial status

Study protocol version 2 (16th July 2019). Recruitment commenced on the 24th July 2019 and is expected to be completed by 25th September 2020.

Abbreviations

CBT—Cognitive Behaviour Therapy

CBT-I—Cognitive Behaviour Therapy for Insomnia

NNT—Number needed to treat

RCT - Randomised controlled trial

cRCT—Cluster randomised controlled trial

References

1. Friedrich MJ. Depression is the leading cause of disability around the world. *JAMA*. 2017;317(15):1517.
2. Merikangas KR, Cui L, Kattan G, Carlson GA, Youngstrom EA, Angst J. Mania with and without depression in a community sample of US adolescents. *Arch Gen Psychiatry*. 2012;69(9):943–51.
3. Sawyer SM, Azzopardi PS, Wickremarathne D, Patton GC. The age of adolescence. *Lancet Child Adolesc Health*. 2018;2(3):223–8.
4. Lewinsohn PM, Hops H, Roberts RE, Seeley JR, Andrews JA. Adolescent psychopathology: I. Prevalence and incidence of depression and other DSM-III-R disorders in high school students. *J Abnorm Psychol*. 1993;102(1):133–44.
5. Johnson D, Dupuis G, Piche J, Clayborne Z, Colman I. Adult mental health outcomes of adolescent depression: A systematic review. *Depress Anxiety*. 2018;35(8):700–16.
6. Lawrence D, Johnson S, Hafekost J, de Hann KB, Sawyer M, Ainley J, et al. The mental health of children and adolescents: Report on the second Australian child and adolescent survey of mental health and wellbeing. Canberra: Department of Health; 2015.
7. McLeod GF, Horwood LJ, Fergusson DM. Adolescent depression, adult mental health and psychosocial outcomes at 30 and 35 years. *Psychol Med*. 2016;46(7):1401–12.
8. Lund C, Cois A. Simultaneous social causation and social drift: Longitudinal analysis of depression and poverty in South Africa. *J Affect Disorders*. 2018;229:396–402.
9. Insel TR. The global cost of mental illness [Internet]. Bethesda, USA: National Institute of Mental Health; 2011. <https://www.nimh.nih.gov/about/directors/thomas-insel/blog/2011/the-global-cost-of-mental-illness.shtml>. Accessed 9 March 2019.
10. World Health Organization. Global status report on non-communicable diseases 2010. Geneva: World Health Organization; 2011.
11. World Health Organization. Mental Health Atlas 2017. Geneva: World Health Organisations; 2018.
12. Andrews G, Issakidis C, Sanderson K, Corry J, Lapsley H. Utilising survey data to inform public policy: Comparison of the cost-effectiveness of treatment of ten mental disorders. *Br J Psychiatry*. 2004;184(6):526–33.
13. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: A systematic review. *BMC Psychiatry*. 2010;10:113.
14. Treatment for Adolescents With Depression Study Team. Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression: Treatment for Adolescents With Depression Study (TADS) randomized controlled trial. *JAMA*. 2004;292(7):807–20.
15. Curry J, Silva S, Rohde P, Ginsburg G, Kratochvil C, Simons A, et al. Recovery and recurrence following treatment for adolescent major depression adolescent MDD treatment and recovery/recurrence. *JAMA Psychiatry*. 2011;68(3):263–9.
16. Ebert DD, Cuijpers P. It is time to invest in the prevention of depression. *JAMA Network Open*. 2018;1(2):e180335-e.
17. Merry SN, Hetrick SE, Cox GR, Brudevold-Iversen T, Bir JJ, McDowell H. Psychological and educational interventions for preventing depression in children and adolescents. *Cochrane Database Syst Rev*. 2011(12):Cd003380.
18. Stockings E, Degenhardt L, Dobbins T, Lee Y, Erskine H, Whiteford H, et al. Preventing depression and anxiety in young people: A review of the joint efficacy of universal, selective and indicated prevention. *Psychol Med*. 2015;46(1):11–26.
19. Werner-Seidler A, Perry Y, Calear AL, Newby JM, Christensen H. School-based depression and anxiety prevention programs for young people: A systematic review and meta-analysis. *Clin Psychol Rev*. 2017;51:30–47.
20. Hetrick SE, Cox GR, Witt KG, Bir JJ, Merry SN. Cognitive behavioural therapy (CBT), third-wave CBT and interpersonal therapy (IPT) based interventions for preventing depression in children and adolescents. *Cochrane Database Syst Rev*. 2016(8):CD003380.
21. Muñoz RF, Cuijpers P, Smit F, Barrera AZ, Leykin Y. Prevention of major depression. *Annu Rev Clin Psychol*. 2010;6:181–212.

22. Cuijpers P. Examining the effects of prevention programs on the incidence of new cases of mental disorders: The lack of statistical power. *Am J Psychiatry*. 2003;160(8):1385–91.
23. Perry Y, Werner-Seidler A, Calear AL, Mackinnon A, King C, Scott J, et al. Preventing depression in final year secondary students: School-based randomized controlled trial. *J Med Internet Res*. 2017;19(11):e369.
24. Fröjd SA, Nissinen ES, Pelkonen MUI, Marttunen MJ, Koivisto A-M, Kaltiala-Heino R. Depression and school performance in middle adolescent boys and girls. *J Adolesc*. 2008;31(4):485–98.
25. Donker T, Blankers M, Hedman E, Ljotsson B, Petrie K, Christensen H. Economic evaluations of Internet interventions for mental health: A systematic review. *Psychol Med*. 2015;45(16):3357–76.
26. Firth J, Torous J, Nicholas J, Carney R, Prata P, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: A meta-analysis of randomized controlled trials. *World Psychiatry*. 2017;16(3):287–98.
27. Grist R, Croker A, Denne M, Stallard P. Technology delivered interventions for depression and anxiety in children and adolescents: A systematic review and meta-analysis. *Clin Child Fam Psychol Rev*. 2019;22(2):147–71.
28. de Bruin EJ, Bogels SM, Oort FJ, Meijer AM. Efficacy of cognitive behavioral therapy for insomnia in adolescents: A randomized controlled trial with internet therapy, group therapy and a waiting list condition. *Sleep*. 2015;38(12):1913–26.
29. Merry SN, Stasiak K, Shepherd M, Frampton C, Fleming T, Lucassen MF. The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: Randomised controlled non-inferiority trial. *BMJ*. 2012;344:e2598.
30. Christensen H, Batterham PJ, Gosling JA, Ritterband LM, Griffiths KM, Thorndike FP, et al. Effectiveness of an online insomnia program (SHUTi) for prevention of depressive episodes (the GoodNight Study): A randomised controlled trial. *Lancet Psychiatry*. 2016;3(4):333–41.
31. de Bruin EJ, Bogels SM, Oort FJ, Meijer AM. Improvements of adolescent psychopathology after insomnia treatment: Results from a randomized controlled trial over 1 year. *J Child Psychol Psychiatry*. 2018;59(5):509–22.
32. Werner-Seidler A, Wong Q, Johnston L, O’Dea B, Torok M, Christensen H. A pilot evaluation of the Sleep Ninja—a smartphone-application for adolescent insomnia symptoms. *BMJ Open*. 2019;9:e026502.
33. Bickman L, Lyon AR, Wolpert M. Achieving precision mental health through effective assessment, monitoring, and feedback processes. *Adm Policy Ment Health*. 2016; 43–271.
34. Insel TR. Digital phenotyping: Technology for a new science of behavior. *JAMA*. 2017;318(13):1215–6.
35. Onnela J-P, Rauch SL. Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. *Neuropsychopharmacology*. 2016;41(7):1691–96.
36. Farhan AA, Yue C, Morillo R, Ware S, Lu J, Bi J, et al.. Behavior vs. introspection: Refining prediction of clinical depression via smartphone sensing data. 2016 OEEE Wireless Health, 2016 Oct 25–27; Bethesda, MD: USA.
37. Place S, Blanch-Hartigan D, Rubin C, Gorrostieta C, Mead C, Kane J, et al. Behavioral indicators on a mobile sensing platform predict clinically validated psychiatric symptoms of mood and anxiety disorders. *J Med Internet Res*. 2017;19(3):e75.
38. Johnson JG, Harris ES, Spitzer RL, Williams JB. The patient health questionnaire for adolescents: Validation of an instrument for the assessment of mental disorders among adolescent primary care patients. *J Adolesc Health*. 2002;30(3):196–204.
39. Werner-Seidler A, O’Dea B, Shand F, Johnston L, Frayne A, Fogarty A, et al. Development of the Sleep Ninja: A smartphone application for adolescents with sleep disturbance. *JMIR Mental Health*. 2017;4(3):e28.
40. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: Validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606–13.
41. Batterham PJ, Sunderland M, Carragher N, Calear AL, Mackinnon AJ, Slade T. The Distress Questionnaire–5: Population screener for psychological distress was more accurate than the K6/K10. *J Clin Epidemiol*. 2016;71:35–42.
42. Batterham PJ, Sunderland M, Slade T, Calear AL, Carragher N. Assessing distress in the community: Psychometric properties and crosswalk comparison of eight measures of psychological distress. *Psychol Med*. 2018;48(8):1316–24.
43. Spence SH, Barrett PM, Turner CM. Psychometric properties of the Spence Children’s Anxiety Scale with young adolescents. *J Anxiety Disord*. 2003;17(6):605–25.
44. Spence SH. A measure of anxiety symptoms among children. *Behav Res Ther*. 1998;36(5):545–66.
45. Essau CA, Muris P, Ederer EM. Reliability and validity of the Spence Children’s Anxiety Scale and the Screen for Child Anxiety Related emotional disorders in German children. *J Behav Ther Exp Psychiatry*. 2002;33(1):1–18.
46. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Med*. 2001;2(4):297–307.
47. Alvaro PK, Roberts RM, Harris JK. The independent relationships between insomnia, depression, subtypes of anxiety, and chronotype during adolescence. *Sleep Med*. 2014;15(8):934–41.
48. Chung KF, Kan KK-K, Yeung W-F. Assessing insomnia in adolescents: Comparison of insomnia severity index, Athens insomnia scale and sleep quality index. *Sleep Med*. 2011;12(5):463–70.
49. Van Spijker BA, Batterham PJ, Calear AL, Farrer L, Christensen H, Reynolds J, et al. The suicidal ideation attributes scale (SIDAS): Community-based validation study of a new scale for the measurement of suicidal ideation. *Suicide Life Threat Behav*. 2014;44(4):408–19.
50. Brener ND, Kann L, McManus T, Kinchen SA, Sundberg EC, Ross JG. Reliability of the 1999 youth risk behavior survey questionnaire. *J Adolesc Health*. 2002;31(4):336–42.

51. May A, Klonsky ED. Validity of suicidality items from the Youth Risk Behavior Survey in a high school sample. *Assessment*. 2011;18(3):379–81.
52. Goodman R. The Strengths and Difficulties Questionnaire: A research note. *J Child Psychol Psychiatry*. 1997;38(5):581–6.
53. Goodman R. Psychometric properties of the Strengths and Difficulties Questionnaire. *J Am Acad Child Adolesc Psychiatry*. 2001;40(11):1337–45.
54. Muris P, Meesters C, Eijkelenboom A, Vincken M. The self-report version of the Strengths and Difficulties Questionnaire: Its psychometric properties in 8- to 13-year-old non-clinical children. *Br J Clin Psychol*. 2004;43(Pt 4):437–48.
55. Maguen S, Hebenstreit C, Li Y, Dinh JV, Donalson R, Dalton S, et al. Screen for Disordered Eating: Improving the accuracy of eating disorder screening in primary care. *Gen Hosp Psychiatry*. 2018;50:20–5.
56. Ougrin D, Boege I. Brief report: The Self Harm questionnaire: A new tool designed to improve identification of self harm in adolescents. *J Adolesc*. 2013;36(1):221–5.
57. McBride N, Farrington F, Midford R, Muleners L. School Health and Alcohol Harm Reduction Project: Details of intervention development and research procedures. Perth, WA: National Drug Research Institute, Curtin University of Technology; 2006.
58. Newton NC, Teesson M, Vogl LE, Andrews G. Internet-based prevention for alcohol and cannabis use: Final results of the Climate Schools course. *Addiction*. 2010;105(4):749–59.
59. Kelleher I, Harley M, Murtagh A, Cannon M. Are screening instruments valid for psychotic-like experiences? A validation study of screening questions for psychotic-like experiences using in-depth clinical interview. *Schizophr Bull*. 2011;37(2):362–9.
60. Buysse DJ, Reynolds III CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Res*. 1989;28(2):193–213.
61. Hinz A, Glaesmer H, Brähler E, Löffler M, Engel C, Enzenbach C, et al. Sleep quality in the general population: Psychometric properties of the Pittsburgh Sleep Quality Index, derived from a German community sample of 9284 people. *Sleep Med*. 2017;30:57–63.
62. Raniti MB, Waloszek JM, Schwartz O, Allen NB, Trinder J. Factor structure and psychometric properties of the Pittsburgh Sleep Quality Index in community-based adolescents. *Sleep*. 2018;41(6): zsu066.
63. Stevens K. Developing a descriptive system for a new preference-based measure of health-related quality of life for children. *Qual Life Res*. 2009;18(8):1105–13.
64. Ratcliffe J, Huynh E, Chen G, Stevens K, Swait J, Brazier J, et al. Valuing the Child Health Utility 9D: Using profile case best worst scaling methods to develop a new adolescent specific scoring algorithm. *Soc Sci Med*. 2016;157:48–59.
65. Furber G, Segal L. The validity of the Child Health Utility instrument (CHU9D) as a routine outcome measure for use in child and adolescent mental health services. *Health Qual Life Outcomes*. 2015;13(1):22.
66. Ratcliffe J, Flynn T, Terlich F, Stevens K, Brazier J, Sawyer M. Developing adolescent-specific health state values for economic evaluation: An application of profile case best-worst scaling to the Child Health Utility 9D. *Pharmacoeconomics*. 2012;30(8):713–27.
67. Stevens K, Ratcliffe J. Measuring and valuing health benefits for economic evaluation in adolescence: An assessment of the practicality and validity of the Child Health Utility 9D in the Australian adolescent population. *Value Health*. 2012;15(8):1092–9.
68. Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): Development and UK validation. *Health Qual Life Outcomes*. 2007;5(1):63.
69. Haver A, Akerjordet K, Caputi P, Furunes T, Magee C. Measuring mental well-being: A validation of the Short Warwick–Edinburgh Mental Well-Being Scale in Norwegian and Swedish. *Scand J Public Health*. 2015;43(7):721–7.
70. Vaingankar JA, Abidin E, Chong SA, Sambasivam R, Seow E, Jeyagurunathan A, et al. Psychometric properties of the short Warwick Edinburgh mental well-being scale (SWEMWBS) in service users with schizophrenia, depression and anxiety spectrum disorders. *Health Qual Life Outcomes*. 2017;15(1):153.
71. Svedberg P, Nygren JM, Staland-Nyman C, Nyholm M. The validity of socioeconomic status measures among adolescents based on self-reported information about parents occupations, FAS and perceived SES; implication for health related quality of life studies. *BMC Med Res Methodol*. 2016;16:48.
72. Rammstedt B, John OP. Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *J Res Pers*. 2007;41(1):203–12.
73. John OP, Donahue EM, Kentle RL. The Big Five Inventory - Versions 4a and 5a. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research; 1991.
74. Balgiu B. The psychometric properties of the Big Five inventory–10 (BFI–10) including correlations with subjective and psychological well-being. *Glob J of Psychol Res: New Trends and Issues*. 2018;8(2):61–9.
75. OECD. Learning for tomorrow's world: First results from PISA 2003 Paris: PISA, OECD Publishing; 2004.
76. Schuster TL, Kessler RC, Aseltine RJ. Supportive interactions, negative interactions, and depressed mood. *Am J Commun Psychol*. 1990;18(3):423–38.
77. Smith AR, Hames JL, Joiner TE. Status update: Maladaptive Facebook usage predicts increases in body dissatisfaction and bulimic symptoms. *J Affect Disorders*. 2013;149(1–3):235–40.
78. Fisher CM, Waling A, Kerr L, Bellamy R, Ezer P, Mikolajczak G, et al. 6th National survey of Australian secondary students and sexual health 2018. Bundoora: Australian Research Centre in Sex, Health & Society, La Trobe University; 2019.
79. Centers for Disease Control and Prevention (CDC). Behavioral Risk Factor Surveillance System Survey Questionnaire [Internet]. Atlanta, Georgia: US: Department of Health and Human Services, Centers for Disease Control and Prevention; 2009, 2010 <https://www.cdc.gov/brfss/questionnaires/index.htm>. Accessed 1 September 2018.

80. Ford DC, Merrick MT, Parks SE, Breiding MJ, Gilbert LK, Edwards VJ, et al. Examination of the factorial structure of adverse childhood experiences and recommendations for three subscale scores. *Psychol Violence*. 2014;4(4):432–44.
81. Meinck F, Cosma AP, Mikton C, Baban A. Psychometric properties of the Adverse Childhood Experiences Abuse Short Form (ACE-ASF) among Romanian high school students. *Child Abuse Negl*. 2017;72:326–37.
82. Schweizer S, Parker J, Leung JT, Griffin C, Blakemore S-J. Age-related differences in affective control and its association with mental health difficulties. *Dev Psychopathol*. 2019: doi:10.1017/S0954579419000099.
83. Schweizer S, Leung J, Kievit R, Speekenbrink M, Trender W, Hampshire A, et al. Protocol for an app-based affective control training for adolescents: Proof-of-principle double-blind randomized controlled trial [version 2; peer review: 3 approved, 1 approved with reservations]. *Wellcome Open Res*. 2019;4(91); doi:10.12688/wellcomeopenres.15229.2
84. Campbell MK, Elbourne DR, Altman DG. CONSORT statement: Extension to cluster randomised trials. *BMJ*. 2004;328(7441):702–8.
85. Calear AL, Christensen H, Mackinnon A, Griffiths KM, O’Kearney R. The YouthMood Project: A cluster randomized controlled trial of an online cognitive behavioral program with adolescents. *J Consult Clin Psychol*. 2009;77(6):1021–32.
86. Thoemmes F, MacKinnon DP, Reiser MR. Power analysis for complex mediational designs using Monte Carlo methods. *Struct Equ Modeling*. 2010;17(3):510–34.
87. Carpenter JR, Kenward MG. *Missing data in randomised controlled trials: A practical guide*. Birmingham: Health Technology Assessment Methodology Programme; 2007.
88. Hamer RM, Simpson PM. Last observation carried forward versus mixed models in the analysis of psychiatric clinical trials. *Am J Psychiatry*. 2009;166(6):639–41.
89. Grist R, Porter J, Stallard P. Mental health mobile apps for preadolescents and adolescents: A systematic review. *J Med Internet Res*. 2017;19(5):e176.

Tables

Table 1. Summary of primary, secondary and tertiary outcome measures, potential mediating and risk factors, and data collection timepoints.

Outcome measures		Baseline	Post (Stage I)	6 mo	12mo	Post (Stage II) *	24 mo	36 mo	48 mo	60 mo
Primary outcome measure										
Depression	Patient Health Questionnaire: Adolescent Version (PHQ-A)	ü	ü	ü	ü	ü	ü	ü	ü	ü
Secondary outcome measures										
Psychological distress	Distress Questionnaire 5 (DQ5)	ü	ü	-	ü	-	ü	ü	ü	ü
Anxiety	Spence Children's Anxiety Scale (SCAS) including Generalized Anxiety and Social Phobia subscales	ü	ü	-	ü	ü	ü	ü	ü	ü
Insomnia	Insomnia Severity Index (ISI)	ü	ü	-	ü	ü	ü	ü	ü	ü
Tertiary outcome measures, potential mediating variable risk factors										
Suicidal ideation	Suicidal Ideation Attributes Scale (SIDAS)	ü	ü	-	ü	-	ü	ü	ü	ü
Suicide behaviour	Youth Risk Behaviour Survey (YRBS)	ü	ü	-	ü	-	ü	ü	ü	ü
Externalising behaviours	Strength and Difficulties Questionnaire (SDQ)	ü	ü	-	ü	-	ü	ü	ü	ü
Eating disorders	Screen for Disordered Eating (SDE)	ü	ü	-	ü	-	ü	ü	ü	ü
Self-harm	Single question from Self-Harm Questionnaire	ü	ü	-	ü	-	ü	ü	ü	ü
Alcohol use	Questions from previous National Drug and Alcohol Research Centre, UNSW studies	ü	ü	-	ü	-	ü	ü	ü	ü
Other substance use	Questions from Australian Institute of Health and Welfare 2007 National Drug Strategy Household Survey	ü	ü	-	ü	-	ü	ü	ü	ü
Psychosis	Adolescent Psychotic-like Symptom Screener (APSS)	ü	ü	-	ü	-	ü	ü	ü	ü
Sleep quality	Pittsburgh Sleep Quality Index (PSQI)	ü	ü	-	ü	ü	ü	ü	ü	ü
Quality of life	Child Health Utility 9D (CHU-9D)	ü	ü	-	ü	-	ü	ü	ü	ü
Well-being	Short-form Warwick-Edinburgh Mental Well-being Scale (SWEMWBS)	ü	ü	-	ü	-	ü	ü	ü	ü
Demographics	Age, ethnicity, location, language, perceived socio-economic status	ü	-	-	-	-	-	-	-	-
Height and weight	Height in cm, weight in kg	ü	-	-	ü	-	ü	ü	ü	ü
Outcome measures (continued)		Baseline	Post (Stage I)	6 mo	12mo	Post (Stage II) *	24 mo	36 mo	48 mo	60 mo
Personality	Big Five Inventory-10 (BFI-10)	ü	-	-	-	-	-	-	-	-
School connectedness	Items from OECD Programme for International Student Assessment (PISA)	ü	ü	-	ü	-	ü	ü	ü	-
Social support	Schuster Social Support Scale (SSSS)	ü	ü	-	ü	-	ü	ü	ü	ü
Social media use	Maladaptive Facebook Use Survey (MFUS), adapted to apply to social media use more broadly	ü	ü	-	ü	-	ü	ü	ü	ü
Gender identification	Sex at birth and current gender identification	ü	-	-	ü	-	ü	ü	ü	ü
Sexual identification and preferences	Single item for each	ü	-	-	ü	-	ü	ü	ü	ü
Sexual behaviour	History of sexual behaviour	-	-	-	-	-	ü	ü	ü	
Romantic relationships	History of romantic relationships	ü	-	-	ü	-	ü	ü	ü	ü
Trauma	Adapted Behavioural Risk Factor Surveillance System - Adverse Childhood Experience (BRFSS-ACE)	ü	-	-	-	-	-	-	ü	-
History of diagnosed mental health condition	Diagnosed mental health conditions	ü	ü	-	ü	-	-	-	ü	ü
Hospitalisation	History of hospitalisation	ü	ü	-	ü	-	-	-	ü	ü
History of diagnosed disability	Diagnosed disability	ü	-	-	-	-	-	-	ü	-
Pubertal development	Menarche, voice breaking & Tanner stages	ü	-	-	ü	-	ü	ü	ü	ü
Bullying	From previous school trials.	ü	ü	-	ü	-	ü	ü	ü	ü
App feedback/acceptability		Baseline	Post (Stage I)	6 mo	12mo	Post (Stage II)*	24 mo	36 mo	48 mo	60 mo
App Feedback	Future Proofing App Feedback Survey	-	ü	-	-	-	-	-	-	-
App Feedback	Post-intervention SPARX-FP Feedback Survey	-	ü	-	-	-	-	-	-	-

* Participants selected to take part in Stage II of the trial only.
 ile 2. Future Proofing app tasks, data, and collection timepoints.

Smartphone collected data	Approximate time	Baseline	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	6 months	12 months	24 months	36 months	48 months
Self-rated mood visual analogue scale (VAS)	1'	=	=	=	=	=	=	=	=	=	=	=	=
Affective card sorting task (CST) cognitive game	3'	=	=	=	=	=	=	=	=	=	=	=	=
Affective backward span task (BST) cognitive game	3'	=	=	=	=	=	=	=	=	=	=	=	=
Speed of typing task	2'	=	=	=	=	=	=	=	=	=	=	=	=
Voice sample task	2'	=	=	=	=	=	=	=	=	=	=	=	=
Accelerometry (movement) device sensor data	<i>3-months of continuous data collection</i>												
Gyroscope (movement) device sensor data	<i>3-months of continuous data collection</i>												
GPS (location) device sensor data	<i>3-months of continuous data collection</i>												

Figures

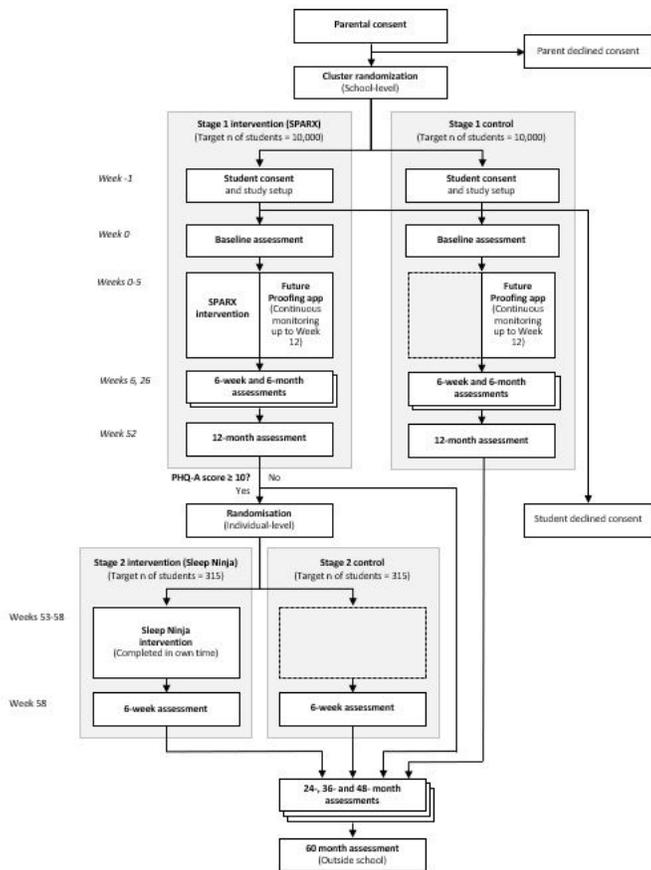


Figure 1

Study Flow

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Appendix1WHODataset.docx](#)