Statistical Analysis Plan

Full/ Long title of the Trial: The Mental Imagery for Suicidality in Students Trial (MISST): A feasibility study.

Short Study title/ Acronym: MISST

Trial registration number(s): ISRCTN13621293

SAP version number with dates: V1; 22Jun2023

Protocol version number and date: V2; 13Feb2023

SAP revisions:

Protocol version	Updated SAP version number	Section number changed	Description of and reason for change	Date changed	Initials

Roles and responsibility: *Names, affiliations and roles of SAP contributors*

Dr. Chris Sutton, The University of Manchester, Lead Statistician

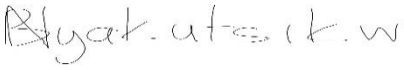
Blessing Nyakutsikwa, The University of Manchester, Trial Statistician

Dr Peter Taylor, The University of Manchester, Co-CI

Dr Jasper Palmier-Claus, Lancaster University, Co-CI

Roles and responsibility: Signatures

Trial Statistician:

Signature:  Date: 11/07/2023

Name: (please print): Blessing Nyakutsikwa

Lead Statistician:

Signature: ....  ..... Date: 12/07/2023

Name: (please print): .....Chris Sutton .....

Trial Steering Committee Chair:

Signature: .....  Date: 11/07/2023

Name: (please print):...Rory O'Connor

Chief Investigator:

Signature: .....  ..... Date: 11/07/2023

Name: (please print):.....Peter Taylor.....

TABLE OF CONTENTS

1 TABLE OF CONTENTS

GLOSSARY AND ABBREVIATIONS

ANCOVA	Analysis of covariance
BMAC	Broad Minded Affective Coping intervention
BHS	Beck Hopelessness Scale
BSS	Beck Scale for Suicide Ideation
SDES	(Short) Defeat and Entrapment Scale
GAD-7	Generalised Anxiety Disorder Assessment
MINI	Mini International Neuropsychiatric Interview
PANAS	Positive and Negative Affect Schedule
PCISS	Perceived Control of Internal States Scale
PHQ-9	Patient Health Questionnaire (Depression)
PSS	Perceived Stress Scale
SASII	Linehan Suicide Attempt Self-Injury Interview
SITBI	Self-Injurious Thoughts and Behaviours Interview (short form)
“8 week”	8-week post-randomisation appointment: in a 7 to 10 week window from randomisation

“16 week”	16-week post-randomisation appointment: in a 15 to 18 week window from randomisation
“24 week”	24-week post-randomisation appointment: in a 23 to 26 week window from randomisation

2 INTRODUCTION

2.1 Background and rationale

Going to university is an important milestone in many people’s lives but can also be a time of significant challenge and stress. There are growing concerns about mental health amongst student populations including suicide risk. Student mental health and counselling services have the potential to prevent suicide, but evidence-based therapies are required that fit these service contexts. The Broad-Minded Affecting Coping intervention (BMAC) is a brief, positive imagery-based intervention that aims to enhance students’ access to past positive experiences and associated emotions and cognitions. Pilot data provides preliminary support for the BMAC for students struggling with suicidal thoughts and behaviours, but this intervention has not yet been evaluated. MISST (the Mental Imagery for Suicidality in Students Trial) is a feasibility Randomised Controlled Trial (RCT), which aims to determine the acceptability and feasibility of evaluating the BMAC as an intervention for university students at risk of suicide within a larger efficacy trial. Key feasibility uncertainties have been identified relating to recruitment, retention, and missing data. Intervention acceptability and safety will also be evaluated.

2.2 Objectives

1. To determine whether University students are willing to be randomised to a trial targeting suicidal experiences.
2. To understand whether it is feasible to collect clinical outcome data in this population.
3. To explore whether patients engage with the BMAC intervention.
4. To determine the safety of the intervention and trial procedures.
5. To explore the initial promise of the intervention, in terms of impact upon clinical outcomes.
6. To investigate what aspects of suicidal experiences might be an appropriate primary clinical outcome for a full trial

7. To understand the potential factors affecting acceptability and delivery (e.g. facilitators and barriers).

   Assessed via the study's qualitative component only – this is therefore not subject to further coverage in this SAP.

8. To gather participant and staff feedback to configure and optimise the intervention and full-scale trial design.

   Assessed via the study's qualitative component only – this is therefore not subject to further coverage in this SAP.

3 TRIAL METHODS

3.1 Trial design

A two-arm parallel-group 1:1 randomised feasibility trial to either: Risk assessment and signposting OR risk assessment and signposting plus the Broad Minded Affective Coping intervention (BMAC).

3.2 Randomisation

An unstratified permuted-block randomisation list was generated using the Sealed Envelope system with blocks of size 4 or 6 varying at random.

3.3 Sample size

There was no formal sample size calculation. A target combined sample size of 66 was chosen on the basis that it would enable investigation of the main research questions regarding feasibility and acceptability. For example, it will enable estimation of the retention rate to within approximately ±11% with 95% confidence, assuming the retention rate is no less than 70%, and sufficient data to estimate the *SD* of suicidal ideation.

3.4 Framework

Not applicable, although the underlying framework for an effectiveness trial would be superiority.

3.5 Statistical interim analysis and stopping guidance

There are no planned interim analyses.

Interim Analysis

Not applicable.

Guidelines for stopping a trial early

Not applicable.

3.6 Timing of final analysis

A single analysis is planned i.e. all outcomes will be analysed collectively.

3.7 Timing of outcome assessments

Outcomes are collected on four occasions:

- Baseline
- “Week 8” (7 to 10 week window from randomisation)
- “Week 16” (15 to 18 week window from randomisation)
- “Week 24” (23 to 26 week window from randomisation)

4 STATISTICAL PRINCIPLES

4.1 Confidence intervals (CI) and level of statistical significance

A 95% confidence level will be used throughout, where relevant, unless otherwise specified in this document. No adjustment for multiplicity will be used and no formal testing will be performed.

4.2 Adherence and protocol deviations

Adherence

Both trial arms include two “risk assessment and signposting” sessions with a trained clinician. The number of sessions attended by each participant will tabulated, both overall and by trial arm.

See Section 6.3 for details around adherence to the BMAC intervention as this is one of the key outcomes of this feasibility trial.

Non-compliances (Protocol deviations)

Non-compliance will include blind breaks, deviations in delivery of intervention, deviations in conducting assessments and deviations in data management. Deviations are logged and severity is judged by the research team.

Serious non-compliances will be line listed and their number and type will be reported by trial arm.

If there are any serious breaches these will also be line listed.

4.3 Analysis populations

The primary analysis population is “intention to treat” (as randomised). The trial safety population will be all participants who provided consent to take part.

5 TRIAL POPULATION

5.1 Screening data

The only screening data that will be reported will be the number that were eligible and the number that were ineligible along with reasons for ineligibility.

5.2 Eligibility

The inclusion criteria are:

- Aged ≥18 years.
- Accessing full or part time education through a HEI.
- Suicidal ideation and/or behaviours in the past three months, ascertained using the questions ‘have you had any thoughts about ending your life in the past three months?’ and ‘have you attempted to end your life in the past three months?’. Endorsement of either item will confirm eligibility for the trial and progression to full assessment. This approach is consistent with previous trials and is sensitive to detecting suicidal experiences amongst adults.

The exclusion criteria are:

- Active/historical full threshold first episode psychosis or bipolar disorder as identified by the patient or referring service and the MINI diagnostic interview.
- Known moderate to severe learning disability (IQ:<70).
- Organic cerebral disease/injury affecting receptive and expressive language comprehension.
- Non-English speaking to the degree that the participant is unable to answer questions and give written informed consent.

- Imminent and immediate risk to self or others, operationalised as the presence of active intent or planning to harm oneself or others in the near future (e.g. next month). Where individuals are excluded on this basis, with the person's consent, the researcher will aim to recontact them and the referrer in approximately one-month's time (or a time period agreed in collaboration with the individual) to determine if risk has subsided to a point where they are now eligible.

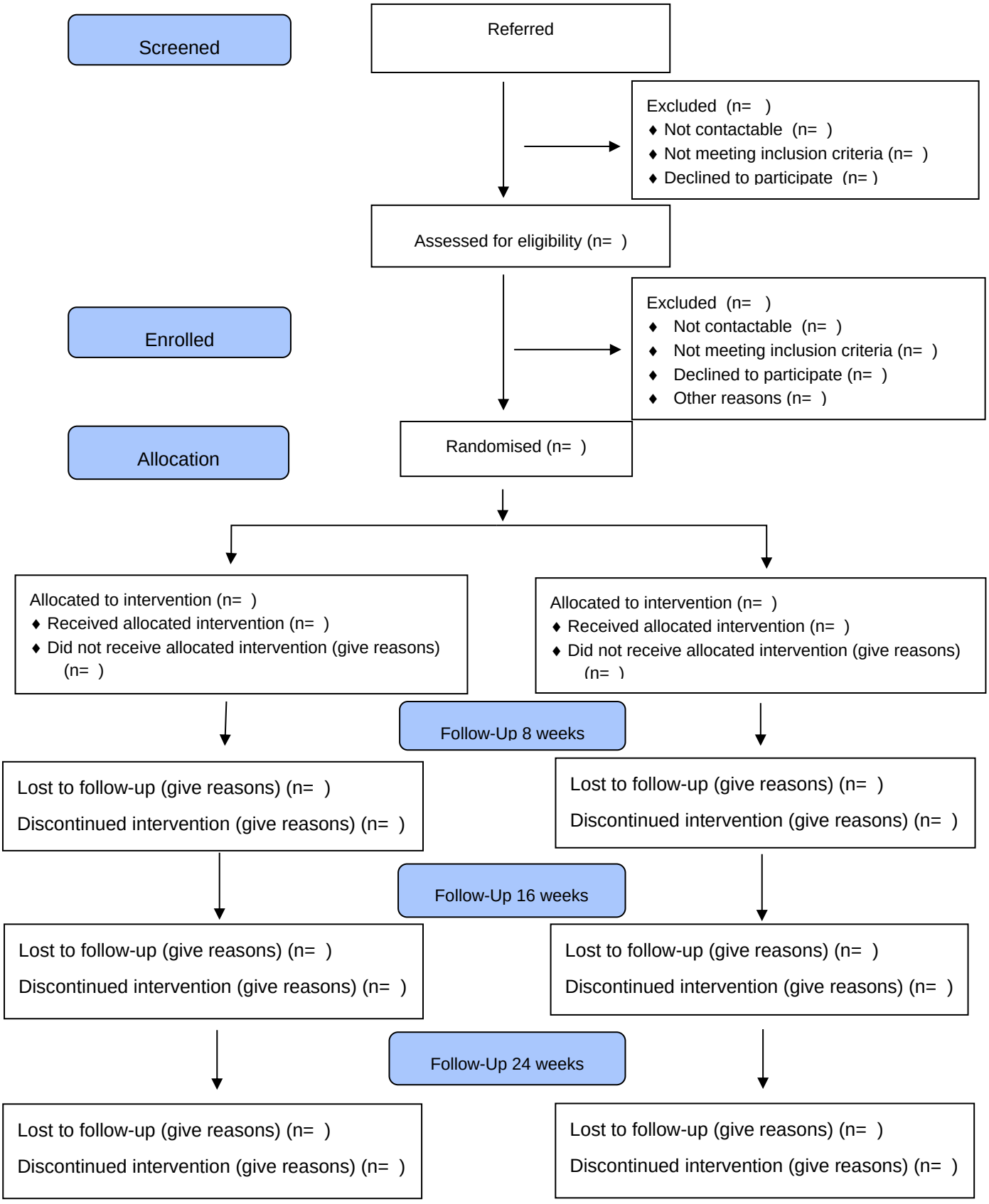
Eligibility information will be summarised as part of the CONSORT diagram. This will include number of referrals screened and found to be eligible, and reasons for non-eligibility.

5.3 Recruitment

This will be presented in a CONSORT diagram (see Figure 1) with numbers and reasons for ineligibility and for non-consent.

Additionally, overall recruitment per month will be presented diagrammatically. Recruitment per month by type of service (NHS; University) will be presented (although numbers recruited by individual services will not be presented). The mean, SD and range of numbers recruited by each counselling service across the trial period will also be presented.

Figure 1: CONSORT Diagram



5.4 Withdrawal/follow-up

This will be presented in a CONSORT diagram (see Figure 1), with numbers and reasons for withdrawal and/or exclusion from analysis given at each stage: 8-week follow-up; 16-week follow-up; 24-week follow-up.

5.5 Baseline patient characteristics

A number of the following items (indicated by an *) permit free-text entries (in some cases to describe an 'other' response) and the research team will undertake an exercise at the end of data collection to group the terms that were used into "types". This mapping will be provided to the statistician and used to recode the items when summarizing. Likewise, for various diagnoses, therapies and medications, the research team will undertake a categorisation exercise to enable statistical analysis to be performed of commonly-occurring categories; any category including three or more responses will be used to form a new indicator variable which will be described as frequency (percentage).

<u>All the variables in the baseline CRF will be summarized by trial arm including</u>:

Gender (female, male, non-binary, gender fluid, F-M trans, M-F trans, other*)

Sexuality (heterosexual, gay/lesbian, bisexual, pansexual, asexual, other*)

Ethnicity (Arab, Asian British, Indian, Pakistani, Bangladeshi, Chinese, other Asian background, Black African, Black Caribbean, Black British, Irish Gypsy or Traveller, White British, White and Black Caribbean, White and Black African, White and Asian, White other, Other Mixed/Multiple Ethnic background)

Marital status (single, partnered, married, open relationship, polyamorous, other*)

Student status (undergraduate, PGR, PGT, other*)

Full or part time (FT, PT)

International student (yes, no)

On break from studies (yes, no)

In employment (yes, no)

Mental health diagnosis (yes, no) – if yes up to 3 may be specified as free text

Physical health diagnosis (yes, no) – if yes up to 3 may be specified as free text

In receipt of psychological therapy (yes, no) – if yes may be specified as free text

Past psychological therapies (yes, no) – if yes up to 3 may be specified as free text

Taking any current medication (yes, no) – if yes up to 3 may be specified as free text

Any past medication – if yes up to 3 may be specified as free text

Admitted to hospital for mental health reasons (yes, no) – if yes, (a) how many times and (b) time since most recent admission

Assessed disorders from the MINI:

Major depressive disorder (yes, no)

Panic disorder (yes, no)

Agoraphobia (yes, no)

Social anxiety (yes, no)

OCD (yes, no)

PTSD (yes, no)

Alcohol use disorder (yes, no)

Substance use disorder (yes, no)

The scales and subscales for the following questionnaires (PROMs) will be calculated and summarised by trial arm (mean, SD, median, minimum, lower quartile, median, upper quartile, maximum).

- BSS
- BHS
- SDES
- GAD-7
- PANAS
- PCISS
- PHQ9
- PSS
- SASII
- SITBI

6 ANALYSIS

The analysis will be focused on the **six objectives** which are addressed using the quantitative data, and the related 'success' criteria for progression to an evaluative trial.

6.1 Objective 1 - To determine whether University students are willing to be randomised to a trial targeting suicidal experiences.

Feasibility outcome

Recruitment rate: the number of participants recruited and randomised (both overall across the eleven-month recruitment window starting February 2022, and per month).

Analysis:

The number recruited and randomised (n_R) will be expressed descriptively:

- as a percentage (x_R) of a fixed denominator 66 i.e. $x_R=100*(n_R/66)$;
- as a percentage of those deemed eligible.

Progression traffic lights criteria:

Green: $x_R \geq 80$; Amber: $60 \leq x_R < 80$; Red: $x_R < 60$.

6.2 Objective 2 - To understand whether it is feasible to collect clinical outcome data in this population

Feasibility outcomes

The key outcome measure is attendance at the "24-week" visit and (at least part) completion of the outcome case-report form (CRF; i.e. set of questionnaires);

Attendance at the "8-week" visit and (at least part-) completion of the outcome CRF;

Attendance at the "16-week" visit and (at least part-) completion of outcome CRF.

Analysis

Retention rate, computed as the percentage of randomised participants who complete the outcome CRF at each time-point (i.e. $x_{c8,}$ x_{c16}, and x_{c24}, respectively). These percentages will be presented descriptively along with 95% confidence intervals, both

overall and by trial arm. The confidence intervals will be computed using the Wilson method. [3]

Progression traffic lights criteria:

Green: $x_{c24} \geq 80\%$, Amber: $60\% \leq x_{c24} < 80\%$, Red: $x_{c24} < 60\%$.

6.3 Objective 3 - To explore whether patients engage with the BMAC intervention

Feasibility Outcome

The BMAC intervention comprises six sessions within an eight-week treatment window measured from the baseline assessment, plus an optional 'booster' session.

The number of BMC sessions received during the 8-week treatment window is one BMAC 'engagement' outcome.

Another engagement outcome is receipt of a "booster" session within the 8 weeks following the main 8-week BMAC treatment period.

Of primary interest, however, is *adherence*. This is defined as receipt of ≥ 2 BMAC sessions within the eight-week treatment window.

Analysis

The *adherence rate*, computed as the percentage (x_A) of the participants allocated to the BMAC intervention who receive ≥ 2 sessions within the 8-eight week treatment window

Additionally, the number of BMAC sessions received by each participant will be tabulated as frequency (%) and will also be presented as mean (SD).

Percentage of BMAC arm participants who opted to attend the 'booster' session.

Progression traffic lights criteria:

Green: $x_A \geq 80\%$; Amber: $60\% \leq x_A < 80\%$; Red: $x_A < 60\%$.

6.4 Objective 4 - To determine the safety of the intervention and trial procedures

Safety Outcomes - **Adverse event categories for MISST**

1. Self-harm

This includes acts of intentional harm directed towards the self regardless of suicidal intent, and so can include suicide attempts as well as acts of non-suicidal self-harm and self-harm where the intent is unclear.

2. Mental health

Events other than self-harm involving a marked deterioration of mental health symptoms, the onset or development of new mental health problems, or occurrence of events linked to mental health deterioration (e.g. hospitalisation).

3. Unintended injury

Occurrence of unintentional injury (e.g. a fall, or hitting head). This would include injuries sustained through events such as road traffic accidents of assault, where the occurrence of injury was not intended by the participant.

4. Substance/alcohol use

Adverse events related to use of substances or alcohol. If these events could also be classed as self-harm then that category should be used.

5. Other physical health

Occurrence or onset of physical health condition, or marked worsening of existing physical health condition, excluding events that could be categories as self-harm or unintended injury (above). Physical health deterioration that is directly linked to substance or alcohol use (e.g. accidental overdose) should be classed as substance/alcohol use.

6. Other

Any event not captured by the above categories.

AEs are categorised as above and coded for severity (3-point numerical scale, from 1 to 3, with 3 indicating greater severity). They are also categorised by type.

All SAEs have been coded for causality (5-point scale ranging from unrelated to "Definite" related), severity (4-point scale ranging from "mild" to "life threatening") and expectedness (binary, expected or unexpected).

AEs and SAEs are recorded separately: the total number of adverse events reported will therefore be the sum of the number of AEs and the number of SAEs etc.

Analysis

Analysis of safety data will be descriptive in nature:

- The number of AEs will be tabulated by trial arm and severity overall and for each category.
- The number of participants with at least one AE will be tabulated by trial arm overall and greatest severity for each category.
- SAEs will be line listed separately for each trial arm including at a minimum category, causality, expectedness, and severity.
- The number of SAEs will be tabulated by trial arm and severity overall and for each category.
- The number of participants with at least one SAE will be tabulated by trial arm overall and greatest severity for each category.

Progression criterion:

Discontinuation of the feasibility trial or non-progression on safety grounds will be considered if the intervention or procedures are deemed to elevate risk. The Trial Steering Committee will oversee SAEs across treatment arms and will be involved in any decision-making around discontinuation or non-progression to an evaluation trial.

6.5 Objective 5 - To explore the initial promise of the intervention, in terms of impact upon clinical outcomes

(Clinical) Outcomes

Handling of item missing: See section 6.7

The following 'clinical' outcomes will be used to explore the promise of the intervention:

The putative **primary outcome** for a subsequent evaluation trial is the BSS which is a 19-item questionnaire assessing a person's suicidal intent (with an additional two questions asking about any previous suicide attempts). Only Q4 to Q19 are used in the scoring. Scoring is performed as follows:

BSS

- If both Q4 and Q5 are scored 0 then Q6 to Q19 should be scored 0. If either Q4 or Q5 is scored as 1 or more, then Q6 to Q19 remain as originally scored.
- Q1 to Q19 are then summed to generate a total score.
- The total score is an integer between 0 and 38 with higher scores being more adverse.

Potential **secondary outcomes** for a subsequent evaluation trial are:

BHS

- This is a 20-item questionnaire.
- All items use binary (0, 1) response options. A response of "true" = 1 and a response of "false" = 0.
- First, the following items should be reversed (before summing them), so that a score of 1 ("true") becomes a score of 0, and a score of 0 ("false") becomes a score of 1: Q1, Q3, Q5, Q6, Q8, Q10, Q13, Q15, Q19.
- Then all items are summed to create an overall score, from 0 to 20.

BHS subscales

In addition to the overall score the questionnaire also has three subscales (Boduszek & Dhingra, 2015).

Items are summed (following reversal of positively worded items as noted above) to create totals for each subscale.

- Subscale 1: Feelings about future: Sum item scores for Q1, Q5, Q6, Q13, Q15, Q19 (subscale scores range from 0 – 6, with higher scores indicating more hopelessness)
- Subscale 2: Loss of motivation: Sum item scores for Q2, Q3, Q9, Q11, Q12, Q16, Q17, Q20 (subscale scores range from 0 – 8, with higher scores indicating more hopelessness).
- Subscale 3: Future expectations: Sum item scores for Q4, Q7, Q8, Q10, Q14, Q18 (subscale scores range from 0 – 6, with higher scores indicating more hopelessness).

SDES (Defeat and Entrapment Scale – Short-Form)

- This is an 8-item questionnaire
- Each item is a statement (e.g. “I feel defeated by life”) which the respondent scores from 0 (Not at all) to 4 (Extremely like me)
- The item scores for Q1 to Q8 are summed to create a total score.
- The total score is an integer between 0 and 32 with higher scores being more adverse.

GAD-7

- This is a 7-item questionnaire used to assess the respondent’s anxiety.
- Each item is scored from 0 (Not at all) to 3 (Nearly every day)
- The item scores for Q1 to Q7 are summed to create a total score.
- The total score is an integer between 0 and 21 with higher scores being more adverse.

PANAS-X - Positive affect and negative affect subscales

- The PANAS-X comprises 60 words or phrases which the respondent indicates feelings and emotions. Each word is ‘scored’ from 1 to 5 to reflect the extent to which the respondent has felt that way over the past week.
- From the 60 words or phrases, 10 are used to create the positive affect subscale and 10 are used to create the negative affect subscale.
- These are putative mediators for the BSS.
- For the positive affect subscale, scores for the following 10 items are summed: Interested, Excited, Strong, Enthusiastic, Proud, Alert, Inspired, Determined, Attentive, Active
- For the negative affect subscale, scores for the following 10 items are summed: Distressed, Upset, Guilty, Scared, Hostile, Irritable, Ashamed, Nervous, Jittery, Afraid
- Each subscale ranges from 10 to 50 with higher scores denoting “more” of the scale attribute.

PCISS

- The PCISS (Perceived Control of Internal States Scale) is an 18-item questionnaire that measures perceived control over thoughts, emotions, and bodily sensations.
- The scores (reversals already coded on the form) for the 18 questions are summed to create a total score.

- The total score is an integer between 18 and 90 with lower scores being more adverse.
- This is a putative mediator for the BSS.

PHQ-9

- The PHQ-9 is a 9-item questionnaire that assesses depressive symptoms.
- Each item is scored from 0 (Not at all) to 3 (Nearly every day).
- The scores for the 9 items are summed to create a total score.
- The total score is an integer between 0 and 27 with higher scores being more adverse.

PSS

- The PSS is a 10-item questionnaire.
- Each item is a question, about the person's feelings and thoughts during the previous month, for which the respondent chooses one of five possible answers, ranging from 'Never' to 'Very Often'.
- Scores for Q4, Q5, Q7, and Q8 are reverse scored (e.g. 0 = 4, 1 = 3, 2 = 2, 3 = 1, 4 = 0.).
- The resulting scores for the ten questions are then summed to create a total score (0 to 40) with higher scores being more adverse.

Interview measures

SASII

- At baseline and each follow-up time-point, two interview-based questions are asked of the participant. These questions are denoted by S1 and S2. S1 is a dichotomous question asking about self-harm events or suicide attempts whereas S2 is a question, completed only if the answer to S1 is 'Yes, reflecting the number of self-harm events including suicide attempts.
- Baseline items for the SASII describe past history of self-harm. S2 should be treated as a continuous frequency scale. If participants had a score of 0 for S1 (a "no" response) then S2 will be scored as zero, indicating no lifetime self-harm.
- For the follow-up assessment points, the SASII is used to record how much self-harm occurred since the last assessment. S2 should be treated as a continuous frequency scale. If participants had a score of 0 for S1 (a "no" response) then S2 should be scored as zero, indicating no self-harm since last assessment point. The key question for analysis is S2 ("If yes, how many times have you deliberately harmed or

injured yourself or attempted suicide since the last assessment?"). This question captures self-harm behaviour occurring during the follow-up period.

SITBI

- The SITBI is used as an assessment of self-harm but provides this information separately for suicide attempts and non-suicidal self-injury, in contrast to the SASII which only records acts of self-harm (which may be suicidal or non-suicidal in nature) without making this distinction.
- The baseline SITBI comprises 15 questions can be used to provide further information on lifetime self-harm, split into suicide attempts (8 questions) and non-suicidal self-injury (7 questions). For suicide attempts, if the first question, Q2, is scored 0 (no) then Q6 to Q9 are all be scored as zero, indicating no suicide attempts. For the non-suicidal self-injury section, if the first question, Q10, is scored 0 (no) then Q13 to Q16 can all be scored zero, indicating no self-injury.
- The follow-up SITBI can be used to provide further information on the presence and frequency of self-harm since the last assessment, split into suicide attempts and non-suicidal self-injury. For suicide attempts, if the first question, Q2, is scored 0 (no) then Q4 to Q6 can all be scored zero (no suicide attempts). If Q7 is scored 0 (no) then Q8 to Q10 can all be scored zero (no self-injury). The key questions for the analysis are Q4 ("How many suicide attempts have you made in the last 8 weeks?"), and Q8 ("How many times in the past 8 weeks have you purposively hurt yourself without wanting to die?"). These capture self-harm behaviour occurring during the follow-up period.
- SITBI questions are not combined: each is analysed separately (except for the coding detailed above).

Analysis

All interval ('continuous') outcomes will be summarised by trial arm and overall at each time-point (baseline, 8 weeks, 16 weeks, 24 weeks) as mean (SD). Any categorical outcomes will be summarised by trial arm and overall at each time-point (baseline, 8 weeks, 16 weeks, 24 weeks) as number (%).

BSS

A longitudinal (baseline, “8 weeks”, “16 weeks”, “24 weeks”) plot will be created displaying both individual and group mean BSS profiles). Some random jitter may be applied to help distinguish points.

A constrained longitudinal data analysis (cLDA) model [1] will be fitted by maximum likelihood with a common mean at baseline and separate means for each trial arm at each of the three follow-up timepoints (7 fixed-effect parameters in total, e.g.using xtreg with random intercept in Stata). Bootstrapping (10,000 replications) will be used for the calculation of bias-corrected percentile intervals with the focus being on the BMAC to control contrast at “24 weeks”. No adjustment for other factors will be included in the model as no “stratification” factors were used in the randomisation algorithm and there was a single therapist.

As per Lee et al [6], multiple confidence interval levels (75%, 80%, 85%, 90% & 95%) will be calculated and plotted together for the BMAC to control contrast at “24 weeks”. Ideally the plot would also display the minimal clinically important difference (MID) if available. Currently there is no robust MID in the literature for the BSS and so the research team will discuss and also seek feedback from the PPI group to guide a decision on a tentative value to be used in the plot.

Progression criteria

Progression will be recommended if the confidence intervals for the putative primary outcome measure show that there is promise that the BMAC is effective and may meet or exceed the MID. Progression will also be considered if there remains the potential for BMAC to the effective on an outcome measure that could be deemed an appropriate candidate for an evaluation trial.

Similar analyses will be undertaken for each of the other “non-interview” secondary outcomes.

The quantitative outcomes for the interview measures, however, will be counts and exploratory analysis will be essential before being able to formulate appropriate longitudinal models for these outcomes.

6.6 Objective 6 - To investigate what aspects of suicidal experiences might be an appropriate primary clinical outcome for a full trial

Feasibility Outcomes

Completion of BSS (putative primary outcome) at each of baseline, “8 week”, “16 week” outcomes “24 week” appointments.

Analysis

BSS completion rates (x_b) will be computed as the percentage of participants who provide a valid BSS outcome at all four time-points (Baseline, “8 week”, “16 week”, “24 week”).

Progression criteria

Green: $x_b \geq 80\%$; Amber: $60\% \leq x_b < 80\%$; Red: $x_b < 60\%$.

Sensitivity analyses

There are no planned sensitivity analyses.

Subgroup analyses

There are no planned subgroup analyses.

6.7 Missing data

Out of window assessments will be treated as missing.

Handling of item missing for PROMs: No official scoring instructions for scales and/or subscales provide explicit details for handling item missing then the provided instructions will be followed. In the absence of such instructions we will use within case mean imputation for situations where scale/subscale items are scored identically and no more than 20% of constituent items are missing.

6.8 Additional analyses

Analysis of putative mediators:

Three putative mediators of any effect of the intervention on the BSS at 24 weeks are proposed; the PCISS and the positive and negative affect scales from the PANAS. The sample size is too small for a rigorous mediation analysis but exploratory analyses will be undertaken to help inform any future trial proposal.

An indication of any effect of the trial intervention on each putative mediator will be available from the previously described longitudinal analyses of secondary outcomes. The investigators anticipate that effects will be apparent by 8 weeks.

If there is any indication of an effect of the intervention on both the BSS at 24 weeks (BSS.24w) and a putative mediator at 8 weeks (pmed.8w) then a simple ANCOVA model with BSS.24w as response and trial arm and baseline BSS will be fitted first without and then with the addition of pmed.8w to see if there is a reduction in the treatment effect which would tentatively support

the mediation hypothesis. This analysis will be repeated with pmed.16w in place of pmed.8w to better reflect the temporal aspect.

6.9 Harms

An assessment of safety is an objective for the trial, please see section 6.4 for details.

6.10 Statistical software

Analysis is likely to be performed using Stata V14 (or later), although graphs may be created using other software (e.g. R).

7 REFERENCES

1. Coffman CJ, Edelman D and Woolson RF (2016). To condition or not condition? Analysing 'change' in longitudinal randomised controlled trials. *BMJ Open*. 6:e013096.
2. Sim J and Lewis M (2012). The size of a pilot study for a clinical trial should be calculated in relation to considerations of precision and efficiency. Journal of clinical epidemiology. 65(3):301-308
3. Wilson EB (1927). Probable inference, the law of succession, and statistical inference. J. Amer. Statist. Assoc. 22, 209-212.
4. Brown, LD, Cai T, & Dasgupta, A (2001). Interval Estimation for a Binomial Proportion. Statistical Science, 16(2), 101-133.
5. Boduszek D and Dhingra K (2016). Construct validity of the Beck Hopelessness Scale (BHS) among university students: A multitrait-multimethod approach. Psychol Assess. (10), 1325-1330
6. Lee EC, Whitehead AL, Jacques RM and Julious SA (2014). The statistical interpretation of pilot trials: should significance thresholds be reconsidered? BMC Medical Research Methodology 14:41.