DREAMS TRIAL: STATISTICAL ANALYSIS PLAN (SAP)

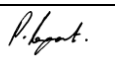
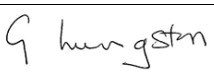
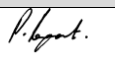
Full title of trial: A parallel multi-centre randomised controlled trial to determine the clinical and cost-effectiveness of DREAMS START (Dementia RElAted Manual for Sleep; STrAtegies for RelaTives) for people living with dementia and their carers

Version of SAP: 2

Date: 21st November 2023

Prepared by: Julie Barber & Mariam Adeleke

Based on protocol version 4 (13th Jan 2023)

Authors	Title	Signature	Date
Dr Julie Barber	Co-ap statistician		21st Nov 2023
Dr Mariam Adeleke	Trial statistician		21st Nov 2023
Reviewers	Title	Signature	
Dr Sarah Amador	Trial Manager		21st Nov 2023
Dr Penny Rapaport	Chief Investigator		21st Nov 2023
Prof Gill Livingston	Chief Investigator		21st Nov 2023
Approver	Title	Signature	Date
Dr Penny Rapaport	Chief-Investigator		21st Nov 2023

Version History Log

Version	Date	Details of change
0.1	8th June 2023	
0.2	25th July 2023	Updated missing data and adjustment factors following review by Penny and Gill.
0.3	16th Aug 2023	Updated detail on adjustment factors after review by Penny, Sarah and Gill
1.0	29th Aug 2023	Added Cross-classified model detail

1.1	16th Nov 2023	Updates following DMEC review and addition of estimand
2	21st Nov 2023	Updated after discussion with Penny and Gill

1. TRIAL SUMMARY

This summary of the DREAM START trial is based on the study protocol (version 4). For full detailed information see the study protocol.

1.1 OBJECTIVES

Primary objective: To determine whether the DREAMS START intervention improves sleep disturbance in people living with dementia at home at 8 months compared to usual NHS treatment.

Secondary objectives: to determine

1. Whether the DREAMS START intervention improves sleep disturbance in people living with dementia at home at 4 months.
2. Whether it reduces daytime sleepiness.
3. Whether it increases people with dementia's quality of life.
4. Whether it is cost effective.
5. The role of psychotropic medication and melatonin in any change.
6. Whether it increases family carers' quality of life
7. Whether it improves family carers' sleep and decreases their affective symptoms and burden.
8. What are the mechanisms of change?
9. If effective, how can we optimise the intervention for implementation at scale in the NHS?

1.2 TRIAL DESIGN

Randomised controlled trial with process evaluation in people living with dementia and family carers.

1.3 RANDOMISATION

Randomisation will occur at the level of the participant and will be blocked and stratified by site using a 1:1 intervention: treatment as usual ratio.

1.4 OUTCOMES

1.4.1 PRIMARY OUTCOME

The primary outcome is resident sleep at 8 months and will be measured using the SDI, a validated instrument. The SDI has seven sleep sub-questions: difficulty falling asleep; getting up during the night

(not scored as positive if someone gets up once or twice per night to pass urine and quickly falls back to sleep); wandering, pacing or getting involved in inappropriate activities at night; awakening the carer during the night; awakening at night, dressing, and planning to go out, thinking that it is morning and time to start the day; awakening too early in the morning (earlier than is his/her habit); and sleeping excessively during the day. Each item is rated according to frequency (scale 0 (Not present in the last two weeks) – 4 (once or more per day (every night)) and severity (scale 1 (mild) -3 (marked)) of sleep-disturbed behaviours and, when multiplied possible item scores range from 0-12. Data will be treated as continuous. This instrument will be used at baseline, 4, 8 month and 24 months follow up (with 8 months as the primary time-point).

1.4.2 SECONDARY OUTCOMES

All secondary outcome measures will be recorded at baseline, 4 and 8 months. A selection of outcomes will also be measured at the 24 month follow up (figure 1).

Person living with dementia (all proxy measures):

1. Neuropsychiatric symptoms (Neuropsychiatric Inventory, NPI) scale assesses dementia-related behavioural symptoms. It has 12 domains measuring the frequency and severity of 12 types of symptoms. Multiplying the frequency and severity ratings gives a score for each symptom (0-12). Accumulating all symptom scores results in the total NPI score (0-44). FOR the total score and all domain scales, higher scores indicate worse severity.
2. Epworth sleepiness scale (ESS) is an eight item measure assessing tendency to sleep/doze in specific daytime situations. Each item is scored 0- 3 so that total scores range from 0-24; a score of >10 indicating excessive sleepiness.
3. DEMQOL-Proxy is a 31 item interviewer-administered questionnaire measuring quality of life in people with dementia. Items are scores as A lot = 1; Quite a bit = 2; A little = 3; Not at all = 4. In calculating a total score, positive items are reversed so that for all items a higher score means better quality of life. (Items: 1, 4, 6, 8, 11 and 32 need reversing). Total score is the sum across 31 items, giving a range for the total score from 31 to 124. Higher scores indicate better QL. (As recommended, where less than 50% of the items are missing, these items will be imputed using the within person mean from all available items. Total scores will not be calculated if there are insufficient (<50%) items).
4. Modified Client Service Receipt Inventory (CSRI) (Beecham and Knapp, 1992) a proxy questionnaire asking about health and social care service use information in the past 4 months for the patient (including care home admission, extra patient care during therapy). The analysis of this outcome measure will be described in the health economic plan.
5. EQ-5D 5 level (EQ-5D-5L)[49] proxy is a generic measure of health related quality of life. Carer proxy responses will be used to calculate QALYs and incremental cost per QALY gained. The analysis of this outcome measure will be described in the health economic plan.
6. Medication- psychotropic medication to delineate the role of rescue medication and any effect of intervention on prescribing. This data will be collected as part of the CSRI.
7. Side effects measure for fall and comorbidities at baseline. Using a Safety, and Tolerability Assessment to record the occurrence of falls, dizziness, headaches and gastrointestinal symptoms (appetite or bowel symptoms) and other side effects and whether these were mild, moderate or severe. This will allow us to assess potential harms.

8. One week actigraphy for person with dementia (Axivity AX3[42] at baseline and before 4 and 8 month follow-up). [This data will be used as part of the process evaluation]

Family carer:

1. Sleep Condition Indicator (SCI) is an eight item scale to assess sleep disturbance. It characterises sleep both dimensionally and against insomnia disorder criteria. Each item scored 0 to 4. Total SCI score is calculated as the sum of all scores giving a range 0 to 32 (a higher score meaning better sleep).
2. The hospital anxiety and depression scale (HADS) is a 14 item measured used to detect the states of depression and anxiety using two separate subscales. Carers are asked to rate, on a 4-point scale (0-3), different aspects of their mood. 7 items related to anxiety and 7 to depression. By summing the items, 3 scores can be calculated: HADS-Depression (sum of the depression items) & HADS-Anxiety (sum of anxiety items) which range from 0 (low severity) – 21 (high severity) and HADS-Total score (sum of HADS-D and HADS-A) ranging from 0 to 42.
3. Zarit Burden Interview (ZBI) measures the impact that care giving has on the carer. A 22-item self-report questionnaire asks different aspects of how people feel taking care of another person on a scale of 0-4. This results in an accumulated score ranging from 0 (no burden) – 88 (severe burden).
4. Health Status Questionnaire (HSQ-12) Measures the impact of health on social, emotional, and physical functioning. Responses to twelve items produce a score for 8 domains. Summary scores within the eight domains range from 0 (negative attribute) – 100 (positive attribute). Items have different number of responses and must be recoded before being summarised (Items 1, 5-7: 1-5; Items 2-4: 1-3; Items 8-12: 1-6). Sub-score of primary interest for analysis is the Mental Health score. The Physical Functioning score will also be reported.
9. Modified Client Service Receipt Inventory (CSRI) a questionnaire asking about health and social care service use information in the past 4 months. This will incorporate the Valuation of Informal Care Questionnaire (iVICQ) a measure of carer time and activity and the Brief Work Productivity and Activity Impairment (WPAI) a measure of productivity loss. The analysis of this outcome measure will be described in the health economic plan.
10. EQ-5D 5 level (EQ-5D-5L) is a generic measure of health related quality of life. The analysis of this outcome measure will be described in the health economic plan.

Figure 1: Schedule of Assessments

	Screening (Pre-treatment)	**Randomisation**	**Treatment Phase**						**Follow up**	**Final visit**	**2 year Follow up**
Visit No:	1		2	3	4	5	6	7	8	9	12
	Baseline		Session 1	Session 2	Session 3	Session 4	Session 5	Session 6	4 month	8 month	+2year
Window of flexibility for timing of visits:		+/- 2 days							+/- 4 weeks	+/- 4 weeks	+ 4 weeks
Screening-Informed Consent	X										X
Screening-SDI (person with dementia)	X								X	X	X
Screening -Dementia type (person with dementia)	X										
Dementia severity (CDR) (person with dementia)	X										
Screening – AUDIT C (person with dementia)	X										
Screening -Eligibility confirmation	X										
Demographics (person with dementia)	X										
NPI	X								X	X	
Epworth sleepiness scale	X								X	X	

DEMQOL-Proxy	X								X	X	X
CSRI/medication (proxy, for person with dementia)	X								X	X	X
EQ-5D-5L (proxy, for person with dementia)	X								X	X	X
Side effects	X								X	X	X
Actigraphy	X								X	X	
Demographics (carer)	X										
Sleep condition indicator (carer)	X										X
HADS (carer)	X								X	X	X
Zarit (carer)	X								X	X	X
Health Status Questionnaire (12) (carer)	X								X	X	
CSRI/medication (carer)	X								X	X	X
EQ-5D-5L (carer)	X								X	X	X
Intervention acceptability									x	x	
Adverse Events review	X					X			X	X	
Withdrawal	X					X			X	X	

1.5 SAMPLE SIZE

We used the standard deviation (SD) of baseline SDI scores (15.74) and the correlation between baseline and follow-up measurements (0.57) observed in our feasibility trial. There is no published SDI minimum clinically important difference. We aim to detect a difference of ≥5.5 points, consistent with important differences identified through our survey of experts. This corresponds to a small-medium effect size of 0.35, and is realistic (an average difference of 5.6 was observed in feasibility work). To account for potential facilitator clustering in the intervention arm, we assumed an inter cluster correlation coefficient (ICC) as observed in the START and MARQUE studies (0.03). A full study with 1:1 randomisation to detect a difference of 5.5 on SDI (effect size 0.35) between intervention and TAU with power 90% and 5% significance requires 370 participants; 185 in each arm (assuming an average of 15 people per facilitator, up to 15% drop out and with inflation for non-normality). This calculation should provide a conservative estimate of the sample size needed in the case where analyses are based on transformed data (e.g. log or square root transformation may be appropriate)[54]. Calculation of sample size was carried out using STATA version 14.

2. DETAILED ANALYSIS PLAN

This analysis plan gives detail of the main quantitative statistical analyses of effectiveness and safety outcomes for the DREAMS trial up to 8 months. Plans for analysis of health economic outcomes, qualitative analyses and analyses of 24 month outcomes will be covered separately. This plan has been prepared following the brief outline provided in the protocol and in advance of any formal comparisons between the randomised groups. The plans for analysis given in this document do not preclude the undertaking of further ad-hoc analyses, although the results of any such further analyses would be interpreted carefully. Furthermore the SAP does not prevent the adaption of any part of the trial analysis, should situations arise in which such adaptation is necessary. Any such adaptation will be fully justified and transparent.

2.1 TIMING

The final trial analysis will take place once the SAP is formally signed off and the completed database including data for all participants up to 8 months has been locked.

2.2 DATA CHECKING

Before analysis and database lock, basic checks will be performed on the quality of the data, focusing on identifying:

- Missing data
- Data outside expected range
- Other inconsistencies between variables e.g. in the dates the questionnaires were completed

If any inconsistencies are found, the corresponding values will be double checked with the researchers and corrected if necessary in the source data. This checking process and subsequent changes will be documented.

2.3 STATISTICAL PRINCIPLES

Analyses will be planned and conducted according to the principles of GCP, the research governance framework, and ICH topic E9 'Statistical Principles for Clinical Trials' and following the SOPs of the PRIMENT clinical trials unit. Results will be reported following Consort guidance.

2.3.1 ANALYSIS POPULATION & PRIMARY ESTIMAND

The primary trial analyses will aim to estimate the effect of the intervention compared with standard care, regardless of the post-randomisation (intercurrent) events listed below for the case where the person living with dementia is alive at 8 months. The analysis will exclude those randomised in error as previously agreed by the trial steering committee/ DMEC.

The primary estimand attributes are defined below:

Estimand Aspect	
Population	Dyads meeting the eligibility criteria
Treatment condition	Up to 6 sessions of intervention + standard care compared with standard care alone, regardless of compliance with treatment, treatment discontinuation or use of rescue medications
Endpoint	SDI at 8 months
Summary measure	Mean difference
Handling of intercurrent events	
Study treatment discontinuation/non-compliance	Treatment policy
Use of rescue / other medications or interventions	Treatment policy
Death of PLWD	Hypothetical strategy (treatment effect estimated assuming no deaths within 8 months)
Death of Carer	Treatment policy
Care home admission	Treatment policy
Hospitalisation of carer or PLWD	Treatment policy

Where possible, outcome data will be collected after each of the listed intercurrent events. Where data is missing/not available, we will assume that the participant's outcome after the event will be similar to that of all other participants. We expect death will be a reason for outcome data being unavailable however our feasibility study indicated there should be few deaths (1 out of 62 people

randomised in the feasibility trial) and we do not expect numbers of deaths to differ by treatment arm. Given this, these unavailable data will be treated as missing data in the primary analysis, therefore estimating the treatment effect under the hypothetical scenario that people do not die within 8 months. Extent, reasons and characteristics of those without data will be examined. Sensitivity analyses will be used to consider the impact of the primary approach for handling unavailable/missing data on the results.

2.3.2 CONFIDENCE INTERVALS AND P-VALUES

Confidence intervals will be presented at the 95% level and, along with P-values, will be 2 sided for all analyses.

2.3.3 STATISTICAL SOFTWARE

Analyses will be conducted by Mariam Adeleke and Julie Barber using Stata version 18 (or above) (StataCorp 2023).

2.3.4 BLINDING

Blinding in this trial is limited to those collecting the outcome data. All efforts will be made to protect this blinding when preparing for analyses. No summaries of outcome data will be reported to the trial team by randomised group until the final analysis report is ready.

2.4 DESCRIPTION OF STUDY SAMPLE

2.4.1 CONSORT DIAGRAM

A consort diagram will be constructed to describe the flow of subjects through the trial (http://www.consort-statement.org/). The diagram will detail the number of subjects: approached and assessed for eligibility; eligible; agreeing to enter the study (with reasons for refusal); receiving the intervention (with reason for not receiving this); followed up and withdrawn (with reasons).

2.4.2 EXTERNAL VALIDITY

To examine external validity, available screening data (site, sex and relationship with person with dementia) will be summarised and compared between those eligible and consenting to the trial and those eligible but who did not consent.

2.4.3 SUMMARY OF INTERVENTION DELIVERY AND ADHERENCE

Session attendance: The number of sessions attended by those in the intervention group will be tabulated, including the proportion of participants who completed ≥ 4 intervention sessions (out of a maximum of 6) which is defined as adherent to the intervention. Reasons for non-completion of the intervention sessions will be summarised (where available).

Facilitators: We will state the number of facilitators who delivered the intervention during the trial, the number of participants per facilitator and the number of facilitators per site. For each participant in the intervention group, we will examine whether all sessions were delivered by the same facilitator and summarise numbers of participants who had contact with more than one facilitator. For analysis purposes, in cases where more than one facilitator delivers the intervention, we will assign that participant to the facilitator from whom they received the majority of their sessions. In the case where they receive equal numbers of sessions from more than one facilitator, we will assign them to the first facilitator they had contact with. We expect the majority of intervention participants to have interaction with only one facilitator.

Mode of delivery: For the intervention group participants, mode of delivery of the sessions will be summarised, giving numbers who received their sessions by video alone, telephone alone, mix of remote (telephone and video) delivery, face to face alone or mixed remote and face to face. Data about mode of delivery is not in the main trial database but will be provided separately by the research team.

2.4.4 PROTOCOL DEVIATIONS

Timing of 4 and 8 month data collection relative to baseline will be reported by randomised group, including the proportion with assessments outside of the pre-specified window (+/- 4 weeks around the 4 or 8 month date relative to baseline). Any other protocol deviations will be summarised.

2.4.5 ATTRITION AND MISSING DATA

Some loss to follow-up is expected over 8 months, particularly due to death. Reasons for outcome data being unavailable will be described and frequency (%) of subjects with missing data, by reason, will be provided for each randomised group (and for each outcome).

Baseline characteristics of those with unavailable outcome data will be examined.

2.4.6 BASELINE TABLE

We will summarise family carer and person with dementia's baseline characteristics by randomised group using means (with standard deviations), medians (with interquartile ranges), counts and proportions, as appropriate. We will use this summary to gauge the balance in characteristics achieved between randomised groups. We will examine the distribution of continuous variables using histograms. No significance testing will be used.

2.5 ANALYSIS OF PRIMARY OUTCOME

2.5.1 SUMMARY INFORMATION

For each randomised group we will summarise the primary outcome (SDI at 8 months) using means with standard deviations and medians with interquartile ranges. We will also graphically examine the

distribution of the scores. We will examine repeated measurements of the SDI outcome at 4 and 8 months by treatment group using summary statistics and profile plots.

2.5.2 MAIN ANALYSIS

The effect of the intervention will be described using the difference in mean 8-month SDI scores (between intervention and control groups) calculated with a 95% confidence interval and P-value. This estimate will be obtained from a 3 level, linear mixed effects multiple regression model which allows analysis of repeated outcome measurements at 4-month and 8-month and for clustering by facilitator in the intervention arm The model will include a treatment group indicator, time indicator, an interaction between treatment and time indicators, baseline SDI score and indicators for site as fixed effects.

We will fit the following heteroscedastic mixed effects model (Candlish *et. al*, 2018) :

$$Y_{ijk} = \beta_0 + \beta_T T_{ij} + \beta_M M_{ijk} + \beta_{TM} T_{ij} M_{ijk} + \beta_1 Y_{0ij} + \beta_2 S_{1ij} \ldots + \beta_{12} S_{10ij} + u_i T_{ij} + \gamma_{ij} + \epsilon_{ijk} T_{ij} + \xi_{ijk}(1 - T_{ij})$$

Where

i is the therapist subscript

j is the participant subscript

k is the time subscript

Y_{ijk} is the outcome, SDI score at 8 months

T_{ij} is the indicator variable for the intervention arm (=0, 1)

M_{ijk} is the indicator variable for the time (0 = 4-month , 1 = 8-month)

β_T is adjusted difference between trial arms at 4 months

$\beta_T + \beta_{TM}$ is the parameter of interest, adjusted difference between trial arms at 8 months

$S_{1ij} \ldots S_{10ij}$ are the indicator variable representing the 11 study sites

Y_{0ij} is the baseline SDI score

$u_i \sim N(0, \sigma_u^2)$ is a random effect representing variation between facilitators in the intervention arm

$\gamma_{ij} \sim N(0, \sigma_\gamma^2)$ is the random effect at the participant level

$\epsilon_{ijk} \sim N(0, \sigma_1^2)$ is the within participant residual in the intervention arm

$\xi_{ijk} \sim N(0, \sigma_0^2)$is the within participant residual in the control arm

We will use adjusted degrees of freedom (kenward-roger, Stata option *dfmethod(kroger)*) and restricted maximum likelihood (Stata option *REML*) for estimation, as recommended (Candish *et al* 2018).

Example STATA code for the primary model:

mixed SDI SDI$_0$ i.randgrp##i.time i.site, || clusterid: randgrp, nocons ||participantid:, residual(independent ,by(randgrp)) dfmethod(kroger) reml

The intra-cluster correlation coefficient (with 95% confidence interval) will be calculated to describe facilitator clustering.

Analysis will include all those with available data.

In the event of the model not fitting
If the model with heteroscedastic residuals does not converge, we will first fit a model with homoscedastic residuals. If convergence issues remain, we will first fit the model without site and if necessary, also exclude facilitator clustering.

In the event of non-normal residuals
Model assumptions will be checked. SDI scores are expected to be slightly positively skewed, however with adjustment for baseline SDI included, the model residuals are likely to be approximately normally distributed. If, however, residuals are found to be severely non normal, the primary model will be refitted after suitable transformation of the SDI score (e.g. a log transformation).

2.5.3 SENSITIVITY ANALYSES FOR MISSING OUTCOME

Under the assumption that data are MAR, 2 approaches will be taken for the primary outcome:

1) We will refit models to obtain estimates adjusted for variables associated with missingness. To identify predictors of missing data, characteristics of participants with and without missing outcome data will be compared using logistic regression models (with missing yes/no as the outcome). The main analysis model will be refitted to adjust for any characteristics found to be associated with missingness and the outcome of interest.
2) We will use multiple imputation methods. The imputation model will include repeated measurements of the outcome of interest, socio-demographic baseline data and any other variables possibly related to missingness and the outcome. The imputations will be performed by study arm. The primary analysis model will then be re-fitted using the imputed data. We will use the number of imputations that is around the proportion of missingness (e.g. 20 imputed sets for 20% missing data) and combine the results using Rubin's rules.

We will use pattern mixture models for sensitivity analysis under MNAR. These will involve modifying the MAR imputed data (as created for the sensitivity analysis 2) above) to reflect agreed MNAR scenarios (adding a specified factor d to imputed values), fitting the primary model and combining estimates using Rubin's rules (Cro *et al,* 2020).

This method will be used to impute 8-month missing SDI values with the following conditions:

a) If the participant has missing outcome data because they were admitted to a care home, then we will add *d1* to the MAR imputed SDI scores.
b) If the participant has missing outcome data due to end of life or death we will add *d2* to MAR imputed values.
c) If the participant has missing outcome data for any other reason, data will remain as previously imputed.

Given disturbed sleep is often a predictor of entry into care, we might expect worse sleep for those that move to a care home. For *d1* we will consider increases of 25%, 50% and 75% of the absolute change of the SDI score observed over 8 months (from baseline value) for all participants.

It is difficult to predict how death/end of life will impact on sleep so *d2* will take a broader range of values considering increases and decreases of 25%, 50% and 75% of the absolute change of the SDI score over 8 months based on all participants.

2.5.4 SUPPORTIVE ANALYSES/OTHER ANALYSES

1) A supportive analysis will be conducted that involves refitting of the main analysis model including fixed effect adjustment for any notable baseline factors with concerning imbalances between randomised groups (if any).
2) A Complier Average Causal Effect (CACE) analysis will be conducted to estimate the treatment effect relevant to the subgroup of participants who would always adhere to the intervention regardless of the assignment (principle treatment strategy). Participants who have been offered and could participate in at least 4 sessions in the intervention group will be deemed as being compliers. A two stage least squares instrumental variables regression model will be used (using STATA command *ivregres 2sls*).
3) The primary analysis will be repeated excluding cases where the 8 month outcome lies outside the prespecified window for assessment (+/- 4 weeks around 8 month date calculated relative to baseline)
4) We will estimate a secondary estimand with similar specification to the primary estimand except that death of the person living with dementia will be handled using a 'while alive' rather than hypothetical strategy. This analysis will use the imputed 8 month values for those with missing data not due to death (as in analysis 2) in 2.5.3), but will exclude all participants who died before 8 months. The analysis will compare 8 month outcomes between randomised group using a 2 level, linear (heteroscedastic) mixed effects multiple regression model which allows for clustering by facilitator in the intervention arm. The model will include a treatment group indicator, baseline SDI score and indicators for site as fixed effects.

2.6 ANALYSIS OF SECONDARY OUTCOMES (EXCLUDING THOSE FOR HEALTH ECONOMIC ANALYSES)

2.6.1 CONTINUOUS SECONDARY OUTCOMES

The following secondary outcomes produce continuous scores:

Person with dementia: NPI**,** ESS**,** DEMQOL-Proxy

Family carer: SCI, HADS - anxiety**,** HADS - depression, ZBI, HSQ-12 - mental

Analyses carried out for these scores will be similar to those described for the primary outcome (section 2.5.1, 2.5.2 & 2.5.3). Main analyses will focus on scores measured at 8 months.

Sensitivity analyses for missing outcome

Similar sensitivity analyses as planned for the primary outcome will be considered for secondary outcomes with concerning amounts of missing data. (section 2.5.4)

Other summaries

HSQ-12 Physical health scores will be summarised by randomised group but not formally compared.

2.6.2 BINARY SECONDARY OUTCOME: PSYCHOTROPIC MEDICATION

The frequency (%) of participants in each randomised group who have taken at least one type of medication during the 4-month and 8-month follow-up period will be calculated. The frequency (%) of each type of medication (anxiolytics and hypnotics, antipsychotics, antidepressants, adjuvant psychotropics, and melatonin) will also be summarised.

Randomised groups will be compared in terms of the proportions who have taken at least one type of medication (within those categories listed above), obtaining an estimate of the difference in proportions and odds ratio with 95% confidence intervals at 4 and 8 months.

The difference in proportions will be estimated using a mixed effects binomial generalised linear model with identity link (Stata command *gllamm* with options *link(id) family(binomial)*). The model will include indicators for treatment group, time indicator, an interaction between treatment and time indicator, site and baseline psychotropic medication as fixed effects and facilitator and participant identifier as a random effect.

Odds ratios will be estimated using a similar mixed effects logistic regression model.

Estimates will be obtained for 4 month and 8 month follow up points.

2.6.3 SIDE EFFECTS

Side effects- which may not be side effects but part of illnesses- (occurrence of falls, dizziness, headaches and gastrointestinal symptoms (appetite or bowel symptoms) and other side effects) and whether these were mild, moderate or severe will be summarised by randomised group using frequency (%).

2.7 SUBGROUP ANALYSES (EXPLORATORY)

Analyses will be conducted for the primary outcome to explore whether the treatment effect differs according to prespecified baseline factors. These analyses will involve adding interaction terms between treatment group and the baseline factor to the primary outcome analysis model (section 2.5.2) and assessing the statistical significance of the interaction terms.

Baseline factors of interest are

1) severity of dementia defined by CDR in 3 groups : 0.5/1 = mild, 2= moderate, 3=severe.
2) HSQ physical health score (continuous score)

2.8 FURTHER EXPLORATORY ANALYSIS

The intervention was delivered to participants remotely, in-person or using a mix of the two. The frequency (%) and mean (SD) of the SDI at 8 months within each group will be summarised. Treatment effects for each mode of delivery will be obtained from a regression model similar to that used for the primary analysis (section 2.5.2) but with 2 treatment indicators to represent the 3 different modes of delivery. This model will additionally be adjusted for age of carer at baseline, ethnicity, relationship of carer to recipient (spouse/partner, child, other), type of dementia (DLB+parkinson's+Vascular / other types) and baseline HSQ physical health score.

2.9 PROCESS EVALUATION (QUANTITATIVE)

In considering change mechanisms, analysis will focus on movement recorded as the rest-activity amplitude from the 7-day actigraph. This measure reflects the relative difference between the least active five hours (L5) and the most active ten hours (M10) in the day. Rest-activity amplitude will be summarised at baseline, 4 and 8 months using means with SD and medians with IQR, by randomised group.

The mediating effect of this measure on the relationship between the randomised group and SCI, ESS and DemQol outcomes will be examined using causal mediation analysis (Stata 18 command *mediate*). Models will adjust for the baseline measurement of SDI/rest-activity amplitude as appropriate. The model for SDI will additionally adjust for age of carer (baseline), relationship of carer and recipient (spouse/partner, child, other), baseline HSQ physical health score and baseline falls (yes/no in previous 4 months)- Interaction terms will be included that allow the effect of rest-activity amplitude to vary across treatment groups. Robust standard errors will be used to account for intra-site correlation (option *vce(clustvar)* in Stata). Results will partition the total effect of the intervention on SDI into a direct effect and indirect effect. These will be reported with 95% confidence intervals. We will also report the proportion of the total effect that is due to mediation (post estimation Stata command *estat proportion*) with a 95% confidence interval.

These analyses will be carried out separately for 8 month and 4 month outcomes.

(Note - We may also consider use of the user written Stata command *ml-mediation* which allows mediation analysis with 2 levels and therefore may better allow for clustering by site)

2.10 FIDELITY ANALYSIS (DATA PROVIDED SEPARATELY BY RESEARCH TEAM)

To analyse fidelity of delivery of DREAMS START, we will assess the number of appointments delivered across all intervention participants. Checklists will be applied independently of the facilitator to a random selection of one recorded intervention session for each participant and researcher. A mean fidelity score will be produced by dividing the number of items on the checklist identified as being delivered in the appointment, by the number of items on the checklist that should have been delivered per appointment, per researcher and across all appointments. We will adopt thresholds used in other intervention fidelity work: where 81–100% constitutes high fidelity, 51-80 is moderate fidelity and

50% or lower constitutes low fidelity. We will also report the median (interquartile range) score for each of 4 process factors.

2.11 REFERENCES

StataCorp. 2023. *Stata Statistical Software: Release 18*. College Station, TX: StataCorp LLC.

Candlish *et al.* BMC Medical Research Methodology (2018). 18:105 https://doi.org/10.1186/s12874-018-0559-x

Cro *et al,* Statistics in Medicine (2020). 39:21 https://doi.org/10.1002/sim.8569