Statistical Analysis Plan
DREAMS START (feasibility trial)

Julie Barber

Version 1.0 : 23rd August 2017

1. Introduction

This statistical analysis plan has been developed initially based on information from the study protocol (version 1) but will be combined with detail agreed through discussion with the trial team. Further drafts will be updated based on examination of a portion of the data when this is clean and available. The plan is focused on quantitative analysis of the follow-up data to 3 months, but not health economic outcomes.

Data is expected to be available for analysis by August 2017.

The statistical package STATA (version 14) will be used for all analyses.

2. Trial summary

Aims and Objectives (from protocol)

Our research question is, how feasible is a pragmatic randomised study to investigate the clinical and cost-effectiveness of a manualised intervention for carers (DREAMS START) to manage significant sleep disturbance in NHS patients with dementia living in their own homes.

Aims
To develop a manualised behavioural intervention for sleep disorder in dementia and examine feasibility of a full scale trial.

Objectives
1. To obtain estimates of acceptability and feasibility that will inform continuation to the main trial. Specifically to estimate (with 95% confidence intervals) the proportion of participants offered the intervention that adhere to it and the proportion of eligible participants who agree to participate in the trial.
2. To obtain estimates required for the main trial's sample size calculation in relation to potential primary outcomes [standard deviations; correlation between baseline and follow-up measurements and drop-out rate].
3. To use qualitative interviews to assess acceptability of the intervention and to detail any required refinements.

Design
Randomised controlled trial of DREAMS START manual compared to treatment as usual (TAU) for feasibility and acceptability.

Randomisation

Randomisation was conducted by a statistician independent of the analysis of the trial. Randomisation was 2:1 to the intervention and treatment as usual for at least 60 people with dementia and was stratified by three Sites (Site 1: Camden & Islington, Site 2: Barnet, Enfield & Haringey, Site 3: UCL/Join Dementia Research) using random block sizes.

Outcomes

Primary outcomes
Feasibility of recruitment-agreement to study/randomisation
Treatment Adherence [attending predetermined session numbers (≥4) - intervention group only]

Secondary outcomes

- Referral rates
- Follow-up rates
- All psychotropic medication prescription (to define rescue medication's role)
- Reported side effects: co-morbid physical illnesses and patient falls
- Choice of outcomes for main trial. We have several measures of sleep (objective and subjective). We will consider feasibility and acceptability by:
    - completion rates of instruments
    - acceptability of tools from qualitative interviews
    - estimates of statistical power and sample requirements based on detecting significant differences in outcomes.

Data collected (information mainly from CRF)
Data from patients and family carers is collected at screening (pre-randomisation), baseline (pre-randomisation), and 3 month follow-up.

General information collected:

1. number of patient-carer dyads screened
2. number eligible
3. number of eligible patients consenting
4. number of eligible carers consenting

Screening (recorded at screening, but provide baseline measures):

1. Socio-demographic details for included participants with dementia: sex, date of birth, type of dementia diagnosed, age when left education, last occupation, current marital status, ethnicity (*gender is available for all participants referred to the study – including those who did not provide consent to be screened, those who were not eligible, and those not randomised).*
2. Socio-demographic details for the included carers: sex, date of birth, current or last occupation, carer relationship to patient, co-resident carer (yes/no), average number of visits/month (non-resident carer), ethnicity (*gender & relationship to person with dementia are available for all referred to the study – including those who did not provide consent to be screened, those who were not eligible, and those not randomised).*
3. Sleep Disorders Inventory is validated for measuring sleep disorder in people with dementia. It has 8 items which describe sleep-disturbed behaviours with each rated according to frequency

(scale 0-4) and severity (rated 0-3) of sleep-disturbed behaviours. Mean frequency and mean severity of items 1-7 give subscores and their product provides the SDI global score. [Note item 8 is not used in scoring].

Baseline: Patient measures:

1. Dementia - type diagnosed (from referral – variable *P_DIA_B*; including diagnoses post-referral – variable *P_DIApostref_B*)
2. Severity of dementia (Clinical Dementia Rating; CDR). The CDR has six domains: Memory, Orientation, Judgment and Problem solving, Community Affairs, Home and Hobbies, and Personal Care. Each scored 0 (none), 0.5 (questionable), 1 (mild), 2 (moderate), 3 (severe). Expect for Personal Care, scored 0/0.5 (none/questionable), 1 (mild), 2 (moderate), 3 (severe). Ratings are used to create the Global Clinical Dementia Rating with 5 categories – healthy (0), very mild (0.5), mild (1), moderate (2), severe (3).
3. Actigraphy - monitors movement and estimates circadian phase. The watch is worn for a maximum of 14 days and average values calculated for:
   Sleep measures:
   - sleep efficiency (%) - capturing both initiation and maintenance of sleep, reflecting proportion of time in bed spent asleep
   - Sleep time (mins)
   - Wake time (mins)
   - time of lights out
   - time of falling asleep
   - time of waking up
   - time of getting up
   - time in bed (hours)
   - fragmentation index

   Nonparametric circadian rhythm analysis (NPCRA) measures:
   - relative amplitude
   - inter-daily stability
   - intra-daily variability
   - L5 - activity count for the five most restful hours
   - L5 start hour (of 5 most restful hours)
   - M10 - activity count for the 10 most active hours
   - M10 start hour (of 10 most active hours)

Core night-time analysis sleep measures:

- Core night-time (midnight to 6am) sleep efficiency
- Core night-time (midnight to 6am) sleep time (mins)
- Core night-time (midnight to 6am) wake time (mins)

[Note that if data is available for less than 7 days/nights, the values for the non-parametric circadian rhythm analysis measures will be set to missing; if fewer than 7 nights could be analysed, these values for the sleep analysis measures will be set to missing]

4. Neuropsychiatric symptoms (Neuropsychiatric Inventory - NPI)]: Assesses the neuropsychiatric symptoms in dementia patients. It measures the frequency and severity of 12 types of symptoms. Multiplying the frequency and severity ratings will give a score for that symptom, accumulating all symptom scores results in the total NPI score. Carers answer for the patient.

5. Epworth sleepiness scale (daytime sleepiness): assesses daytime somnolence. This is an eight item measure assessing tendency to sleep/doze in specific daily situations. Each item is scored 0-3, total score is sum of items (score range 0-24 (maximum sleepiness); a score of >10 indicating excessive sleepiness).
6. DEMQOL proxy: a 32 item interviewer-administered questionnaire answered by a carer. Items 1-31 are scored 1 (a lot) through to 4 (not at all). Item 32 is scored 1 (very good) through to 4 (poor). Total score is calculated as the sum across items 1-31 and ranges from 31 to 124, with higher scores indicating better QL. (Missing values will be replaced with the mean from all available items. Total scores will not be calculated if there are insufficient (<50% of items).)
7. Patient services use (CSRI) used for cost effectiveness.
8. Medication - as part of the CSRI –from this we identify psychotropic medications (present/absent, number, names)
9. Side effects measure for fall and comorbidities - Safety and Tolerability Assessment to record the occurrence of falls, gastrointestinal symptoms (diarrhoea, nausea, sore mouth, vomiting), neurological symptoms (headaches, visual/auditory disturbances, dizziness), infections and whether these were mild, moderate or severe. It also asks about any other side effects (yes/no) and what those were, as well as any other comments.

<u>Three months – Patient measures (details of measures given above):</u>

1. Sleep Disorders Inventory - SDI global score
2. Actigraphy - monitors movement and estimates circadian phase. The watch is worn for a maximum of 14 days at the 3 month follow-up and average values calculated for the range of measures listed above for baseline.
3. Neuropsychiatric symptoms (Neuropsychiatric Inventory - NPI)]: 12 symptom scores and total NPI score.
4. Epworth sleepiness scale (daytime sleepiness): total score
5. DEMQOL proxy: total score
6. Patient services use (CSRI) used for cost effectiveness.
7. Psychotropic medications from the CSRI (present/absent, number, names)
8. Side effects measure for fall and comorbidities – as for baseline
9. Number of intervention sessions attended (adherence to intervention)

<u>Carer measures - baseline and three months:</u>

1. Two questionnaires for sleep quality
    a. Pittsburgh Sleep Quality Index: (PSQI): measures the effect on carer sleep. It has 9 main items and 7 component scores. The global PSQI score is calculated as the sum of 7 component scores (scored 0-3 each) giving a range from 0 to 21. A total score of 5 or greater is indicative of poor sleep quality.
    b. Sleep Condition Indicator (SCI): An eight item scale characterising sleep 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. Hospital Anxiety and Depression Scale (HADS): Detects 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. Three scores can be calculated: HADS-Depression & HADS-Anxiety 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. Caregiver Health Status Questionnaire (HSQ-12): Measures the impact of health on social, emotional, and physical functioning. Responses to twelve items produce a score over the range on 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).

Power calculation (as in protocol)

With 40 intervention participants (larger to allow a more precise estimate of proportion adhering to intervention) and 20 controls we will achieve the following 95% confidence intervals (CI) for our expected adherence and participation estimates:

1 Proportion of participants adhering to intervention- expected value 75%, 95% CI: 59-87%

2 Proportion of appropriate referrals consenting to the trial-expected value 50%, 95% CI:41-59%

We judge that these confidence intervals provide acceptable ranges to inform continuation to the main trial. Overall we expect that our "stop-go" measures would be related to the proportion adhering-

1. >=70% - go to main trial
2. 60-69 –consider a modified trial design to increase adherence
3. <60 do not progress to main trial using this model.

This sample size will also be sufficient to estimate the standard deviation required for the sample size calculation in the main trial[33;34].

We estimate that recruitment referral rate will be approximately six potential participants per week; two participants will be suitable and will agree to participate (one per research assistant); and follow-up will be approximately 80%.

Statistical analysis (as in protocol)

Our co-applicant statistician will lead and supervise the analysis. The statistical analyses will be described in a predefined statistical analysis plan, and planned and conducted according to ICH E9 and following the standard operating procedures of the PRIMENT clinical trials unit. A summary of the main analyses are given here.

Baseline data will be summarised by treatment group using means (with standard deviations), medians (with interquartile ranges), counts and proportions, as appropriate, to gauge the balance in characteristics between the randomised groups. A consort diagram will describe the flow of patients through the trial.

1. The proportion of participants in the intervention group that adhere to the treatment (attend at least 4 out of the 6 sessions) and the proportion of appropriate referrals consenting to the trial will be calculated with 95% confidence intervals.
2. Actigraph measurements of sleep efficiency (%), total sleep time (minutes) and wakefulness during the sleep period (minutes), the Epworth sleep inventory scores and the Sleep Disorders Inventory scores will be summarised to provide distributional information and to

obtain estimates of standard deviation and correlation needed for the sample size calculation of the main trial.

3. Referral rates, drop out and loss to follow-up rates will be calculated with 95% confidence intervals.
4. In addition, prescription of psychotropic medication (to define rescue medication's role) and co-morbid physical illnesses and patient falls will be summarised by randomised group.
5. Other patient and carer outcome data will also be summarised by randomised group using appropriate estimates with 95% confidence intervals.

3. Detailed analysis plan

Comparison of characteristics between those consenting and not consenting
The following characteristics will be summarised and compared between eligible patients and carers who consented vs. those who were screened and did not consent:

a. Patients: site, gender, relationship with carer (ethnicity may be summarised if available for sufficient numbers of not consented patients)
b. Carers: Gender (non resident/resident & ethnicity may be summarised if available for sufficient numbers of not consented carers)

These characteristics will be summarised in each group as counts and proportions and compared using appropriate two sample methods.

Summary of recruitment and follow-up
Flow of patients through the trial will be described using a consort diagram. The following will also be calculated with 95% confidence intervals:

a. Proportion of screened patients who were eligible for the trial;
b. Proportion of eligible referrals consenting to the trial;
c. Proportion of participants in each randomised group who dropped out or were lost to follow-up by 3 months. (Where available, reasons for losses will be summarised);
d. Proportion of participants in the intervention group who adhered to the intervention (attended at least 4 out of the 6 sessions);
e. Mean/Median number of sessions attended by those in the intervention group.

Summary of baseline data
Baseline data (sociodemographic characteristics, actigraph measures and other scores) for patients and carers will be summarised by treatment group using means (with standard deviations), medians (with interquartile ranges (IQR)), counts and proportions, as appropriate, to gauge the balance in characteristics between the randomised groups. Distributional features of continuous scores will be examined graphically.

3 month follow-up
Follow-up scores, actigraph measurements and other scores at 3 months will be summarised using means (with SDs), medians (IQR), counts (%) as appropriate. For continuous measures correlations between baseline and follow-up measurements will be calculated. Percentage of participants with missing values for each outcome, along with reasons for missingness (where available) will be summarised.

Measurements will be compared between randomised groups using appropriate regression models to provide estimates of the effect of the intervention with 95% confidence intervals (e.g. difference in means) adjusted for baseline score and site. For actigraph measurements analyses, we will focus on sleep efficiency, relative amplitude, sleep fragmentation index, start time of most restful hours, activity count for most restful hours, start time of most active hours, and activity count for most active hours.

Use of psychotropic medication

The frequency (%) of participants in each randomised group who had taken each type of medication (anxiolytics and hypnotics, antipsychotics, antidepressants, adjuvant psychotropics, and melatonin) during the 3 months prior to the baseline and follow-up will be calculated. The difference in proportions and odds ratio will be reported with 95% confidence interval. Estimates adjusted for site and baseline use will be obtained using logistic regression.

Side effects

Frequency (%) of co-morbid physical illnesses and patient falls will be summarised by randomised group.