Oxford Respiratory Trials Unit

PREDICT

Positive Airway Pressure in Older People: A Randomised Controlled Trial

STATISTICAL ANALYSIS PLAN

Version 1.1

Version	Date	Comments
0.1	06/04/2011	First draft
0.2	09/05/2011	Updated following SAP meeting
0.21	14/06/2011	Minor updates prior to TSC meeting
0.3	08/03/2012	Multiple Imputation section added
0.4	02/04/2012	Updates following SAP meeting
0.5	13/04/2012	Minor changes before finalising
1.0	19/04/2012	Final draft of plan
1.1	10/06/2013	See appendix for details of changes

Contents

1. INTRODUCTION

This document details the planned statistical analyses for an investigator-blind, randomised controlled trial that compares continuous positive airway pressure plus best supportive care (CPAP) against best supportive care only (BSC) for treatment of obstructive sleep apnoea hypopnoea syndrome (OSAHS) in patients aged 65 and over.

Full details of the background to the trial and its design are presented in the trial protocol.

The analyses described in this document will be performed by the designated statistician at the MRC Clinical Trials Unit. All data will be analysed using STATA version 12.

The trial statisticians responsible for writing this document in discussion with the co-chief investigators and other principal investigators and conducting the final analyses are:
Daniel Bratton, MRC CTU
Andrew Nunn, MRC CTU.

2. DESIGN

2.1 Summary

This study is a two-arm, investigator-blind, parallel group, multi-centre randomised controlled trial. A total of 270 participants will be recruited to the study and followed-up for 12 months. At baseline participants will be randomised to receive CPAP plus BSC (active) or BSC only (control).

The co-primary endpoints are:

1. the therapeutic outcome of change in Epworth Sleepiness Scale score between the mean of the scores at months 3 and 4 and the baseline score
2. the cost efficiency of CPAP therapy calculated through the impact of CPAP on health-related quality of life and health service utilisation over 12 months of follow-up.

2.2 Inclusion Criteria

- Age $\geq$65 years
- A clinical diagnosis of OSAHS: $\geq$ 4% Oxygen desaturation index > 7.5 events/hour and an Epworth sleepiness scale $\geq$ 9
- Ability to give written informed consent

2.3 Exclusion Criteria

- Previous exposure to CPAP therapy
- Arterial oxygen saturation <90% on room air
- FEV_1 / FVC <60%

- Substantial problems with sleepiness driving (in those who are still driving)
- Currently using HGV or PSV driving licence (where applicable - annual application is required for drivers > 65 years)
- Shift work
- Any very severe complication of OSAHS such that CPAP therapy is mandatory
- Inability to give informed consent or comply with the protocol e.g. the patient must be able to see to be able to participate in the wakefulness test
- Enrolled in another intervention study

2.4 Randomisation

Once a participant has given written consent to the trial, an enrolment form is completed and the participant is randomised using the telephone computerised randomisation service of the MRC CTU. Randomisation is by minimisation with a random element of 80%. The minimisation criteria are:

- Subjective day time sleepiness (Epworth sleepiness score, > 13 or ≤ 13)
- Functionality (Townsend disability index, >1 or ≤1)
- Recruiting centre

3. OUTCOME MEASURES

3.1 Primary outcome measures

Difference between the two treatment arms in:

- Subjective Sleepiness: the mean change in the mean of the Epworth Sleepiness Scale (ESS) scores measured at months 3 and 4 compared to baseline. The ESS assesses the tendency to fall asleep during eight typical daytime scenarios (1). Each component is given a score of 0, 1, 2 or 3 to represent no, slight, moderate or high chance of dozing respectively. The ESS score is then the sum of its eight components. If at least one of the components is missing the ESS will also be set to missing. Should non-integer values be given, these should be included in the sum and the final ESS rounded up to the next integer.
- Cost effectiveness at 12 months: Described by the EQ-5D, valued using UK population tariffs. This will be used to estimate the cost per QALY gained by providing CPAP in comparison to Best Supportive Care. The analysis will incorporate health care utilisation, including in patient and out patient hospital visits and GP visits during the trial. The cost-effectiveness analysis will be performed by the Centre for Health Economics, University of York and a separate analysis plan will be written for this outcome.

3.2 Secondary outcome measures

Difference between the two treatment arms in:

- Subjective sleepiness: the mean of the ESS scores measured at months 10, 11 & 12 compared to baseline

In addition, the change from baseline in the following outcomes will be analysed at 3 and 12 months:

Difference between the two treatment arms in:

- Objective sleepiness: OSLER (Oxford Sleep Resistance Test). This test assesses a patient's ability to resist sleep for 40 minutes. Two tests are conducted at each visit (baseline, 3m and 12m) and the average time taken to fall asleep at each visit will be used for analysis.
- Self reported health status (quality of life and mood):
  - Short Form 36 questionnaire (SF-36) consists of 36 quality of life related questions. Answers to questions are condensed into 8 summary scores which can be reduced further to the mental component summary (MCS) and physical component summary (PCS) scores. Each summary score will be calculated using the formulae proposed by Jenkinson et al (2). Should any of the 36 questions not be answered, the MCS and PCS will be set to missing along with any of the 8 summary scores which are dependent on the missing answers.
  - Sleep Apnoea Quality of Life Index (SAQLI; a disease specific sleep apnoea questionnaire which includes CPAP side effects). The SAQLI is scored by averaging the answers to 14 sleep apnoea related questions and, if applicable, adjusting for side effects attributable to CPAP therapy (see 10.1). Should any of the answers to 14 questions be missing, the SAQLI will also be set to missing)
  - Hospital Anxiety & Depression Scale (HADS). The anxiety and depression aspects of the HADS will be scored by summing the scores from the relevant questions, each of which is scored on a 0-3 scale (7 questions for each aspect).
- Functional index of activities of daily living: Townsend Disability Index (TDI). Each of the 9 items of the TDI is scored with either 0 (Yes, with no difficulty), 1 (Yes, with some difficulty) or 2 (No, need help). Items are then summed to give a total score (3). If at least one of the components is missing the TDI will also be set to missing.
- Frequency of nocturia: The average number of times that patients get up to pass urine at night is reported at the study visits
- Mobility: The Timed up and go test measures, in seconds, the time taken by an individual to stand up from a standard arm chair, walk a distance of 3 metres, turn, walk back to the chair and sit down. There is no upper time limit and the time in seconds is rounded up or down to a whole second.
- Road, and domestic accidents: the number of domestic accidents are self-reported at the follow-up visits (3 and 12 months). The proportion of patients experiencing each accident and any accident will be analysed.
- Cognitive function: Mini-mental state score, Trail making B time, Digit Symbol Substitution test score and simple and four-choice reaction time (see section 0 for a description of these tests).
- Cardiovascular Risk factors: systolic and diastolic blood pressures (SBP & DBP), fasting glucose, fasting lipids, HbA1c.
- Adverse cardiovascular events: Myocardial infarction, stroke, transient ischemic attack, new angina, new atrial fibrillation and new peripheral vascular disease. At each follow-up visit (3 and 12 months) patients report whether they have been newly diagnosed or experienced any of these events since the last visit.

The proportion of patients experiencing any adverse cardiovascular event listed above will be compared between treatment arms.

3.3 Tertiary outcome measures

- Treatment compliance: Measured objectively by smartcards in the machines and downloaded at 3 and 12 month clinic visits.

4. COGNITIVE FUNCTION TESTS

- The Mini Mental State Examination (MMSE) is a widely used screening tool for cognitive function. The MMSE provides a measure of orientation, registration (immediate memory), short-term memory (but not long-term memory) as well as language functioning. It is scored out of 30. Scores of 25-30 are considering normal; 18-24 indicate mild-to-moderate impairment; scores of 17 or less indicate severe impairment.
- The Trail Making Test B (TMT-B) provides information on visual search, scanning, speed of processing, mental flexibility, and executive functions. It requires individuals to draw a line sequentially connecting 25 encircled numbers and letters, distributed on a piece of paper alternating between numbers and letters (e.g. 1, A, 2, B, 3, C etc.). The score represents the amount of time required to complete the task. Performance on the TMT decreases with increasing age and lower levels of education.
- The Digit Symbol Substitution is a coding exercise. It requires an individual to copy a code at the top of the piece of paper. Each symbol in the code corresponds to a single digit number. The individual must write each code under each number and complete as many as possible in 90 seconds. The total number they get correct in this time is recorded.
- The simple and four choice reaction time is a two part test which measures reaction time and the number of correct responses and errors and is completed on a computer. The first test measures the time to react to a symbol appearing in a white box on the screen by pressing any button on a keyboard. The second part requires the individual to respond to the symbol appearing in any 1 of 4 white boxes at random. They have to respond using the allocated key on the keyboard.

5. SAMPLE SIZE CALCULATIONS

The primary analysis will be the difference between the two treatment arms in the mean change of the Epworth Sleepiness Score (ESS) from baseline to the mean of the 3 and 4 month scores. In the recent NICE/HTA Technology Appraisal of CPAP for OSAHS in middle-aged patients (4), the effect of CPAP treatment on the difference in ESS in middle-aged patients with mild sleep apnoea was -1.07 (SD 2.4). The inclusion criterion for this trial lies in the range of "moderate" sleep apnoea by OSAHS severity, but since sleepiness is often less pronounced in older people, power calculations are performed assuming a treatment response similar to that seen in mild disease in the middle-aged. A mean change of 1 point on the ESS is the minimum clinically significant change since it is indicative of one symptom state shift on one domain of

the score. To detect a one point change in Epworth score (SD of change 2.4), requires 244 patients randomised in a 1:1 ratio (alpha=0.05, power 90%).

In previous randomised trials with a similar design a loss to follow-up rate of 5% was found. Since PREDICT is a 12-month trial we have assumed the loss to follow-up rate will be 10%. Patients who cease CPAP therapy will be followed-up through the normal trial systems. Therefore, the sample size for this trial will be 270 patients in total randomised in a 1:1 ratio.

6. ANALYSIS PRINCIPLES

6.1 Minimisation factors

Randomisation will be by minimisation with a random element of 80%. The minimisation factors are:

- Epworth sleepiness score, ESS (13 or less, or above 13)
- Townsend disability index, TDI (1 or less, or above 1)
- Recruiting centre

All analyses will be adjusted for these factors to optimise power and reduce bias. The ESS and TDI will be entered into models as fixed effects continuous variables. Recruiting centre will be adjusted for using random effects in order to avoid dropping centres that may only recruit a single patient.

6.2 Other covariates

Age, gender, ODI and BMI will also be adjusted for in addition to the minimisation factors in an additional analysis of the primary efficacy endpoint.

6.3 Other principles

- All analyses will be intention-to-treat incorporating all randomised patients who have data recorded on the outcome of interest (complete case analysis).
- No adjustments for multiple testing will be made, but cautious interpretations will be made of statistically significant secondary outcomes due to the large number of secondary analyses being performed.

7. ANALYSIS DETAILS

7.1 Patient flowchart

Patient throughput, from those screened for entry through those who are eligible (meet all inclusion criteria and no exclusion criteria) for the trial will be reported. The throughput of patients from those eligible to be randomised to those that are included in the ITT primary analyses will be summarised in a CONSORT flowchart.

The number of patients who are excluded at screening (failure to satisfy inclusion and exclusion criteria, refusal to participate), discontinued from treatment, and discontinued from follow-up will be reported.

7.2 Baseline characteristics

Baseline characteristics will be summarised by treatment arm. Categorical variables will be summarised by number and percentage in each category and continuous variables will be summarised by mean and standard deviation or by median, 25th and 75th percentiles as appropriate. No formal statistical tests will be performed since any differences should be the result of chance rather than bias.

7.3 Primary endpoint analysis

Primary effectiveness outcome

The mean of the 3 and 4 month ESS scores will be calculated for each patient and used as the follow-up ESS score. Should either score be missing, the single observed score will be used in the analysis. If both scores are missing the patient will be excluded from the primary analysis. Any 3 or 4 month ESS scores which are obtained before 2 months or after 5 months of follow-up has been completed will be excluded from the analysis. The difference between the follow-up ESS and the ESS used for randomisation will then be calculated for each patient and compared between treatment groups using a multivariable linear regression model. The analysis will be adjusted for the minimisation factors as outlined in section 6.1.

7.4 Secondary endpoint analyses

ESS

The mean of the observed 10, 11 and 12 month ESS scores will be calculated for each patient and will be taken to be the 12 month subjective sleepiness score. Similar principles to those described in section 7.3 for calculating the mean score will be used. The difference between the two treatment arms in the change in subjective sleepiness at 12 months compared to baseline will then be analysed using a multivariable linear regression model adjusting for the minimisation factors.

OSLER

Each patient participates in two OSLER tests at baseline, 3m and 12m. Kaplan-Meier plots will be used to summarise the mean time taken to fall asleep (the event of interest) at baseline, 3 and 12 months. The difference in the mean time taken to fall asleep at each follow-up visit compared to baseline will be compared between treatment groups using multivariable linear regression models. Analyses will be adjusted for the mean time taken to fall asleep at baseline in addition to the minimisation factors.

Other continuous outcomes

Continuous outcomes (SF36, SAQLI, HADS, TDI, cognitive function tests, cardiovascular risk factors, mobility test, frequency of nocturia) will be analysed using multivariable regression models and will be adjusted for their corresponding baseline score/measurement and the minimisation factors. Non-normal (skewed) data should not be an issue and can be analysed using this method due to the implications of the Central Limit Theorem that for a large sample size the mean will be approximately normally distributed.

Binary outcomes

For binary outcomes (accidents, adverse cardiovascular events) the odds of experiencing the outcome will be compared between treatment arms using logistic regression. The comparison of the odds of patients having an accident (at home or while driving) will be adjusted for the accident history at baseline (whether had an accident at home in the month before enrolment or while driving in the three months before enrolment). All analyses with be adjusted for the minimisation factors.

7.5 Tertiary endpoint analyses

Treatment usage is taken to be the mean number of hours that CPAP is used per night during follow-up (total number of hours used divided by total number of days follow-up). CPAP usage will be summarised using the median and 25th and 75th percentiles since the data are likely to be skewed.

Patients who have stopped CPAP during follow-up and are missing adherence data will be assumed to have zero hours/night usage. The number of patients stopping CPAP or swapping to CPAP from BSC will be summarised along with reasons.

7.6 Sensitivity analyses

Patients who are randomised to the control and who start CPAP therapy during follow-up may dilute the results of the ESS comparisons. Sensitivity analyses of the primary and secondary ESS outcomes will be performed in which ESS observations in control arm patients will be excluded from analysis if CPAP therapy is started before the visit at which the observation is recorded.

7.7 Multiple Imputation

Under the Missing at Random (MAR) assumption

The missing at random (MAR) assumption assumes that the probability that the missing data depends on the values of the observed data but does not depend on the values of the missing data.

Under the MAR assumption multiple imputation can be used to impute missing ESS scores over follow-up and produce an unbiased analysis on all randomised individuals. The plausibility of the MAR assumption will be explored by comparing observed data in those patients with and without the outcome of interest.

All 12 ESS follow-up scores will be entered into an imputation model along with the minimisation variables and the variables listed in section 6.2. Imputations will be performed separately within treatment groups. CPAP compliance at the 3 month and 12 month visits will also be included in the imputation model for the CPAP arm. For each treatment arm fifty imputation models will be created using the 'ice' command in Stata. In analyses secondary to those described above the primary and secondary ESS outcomes will be reanalysed on the imputed datasets and the results combined using Rubin's rules.

Sensitivity Analysis

The MAR assumption is untestable and may be inappropriate so the probability that data are missing could depend on values of the missing data (missing not at random, MNAR). The ESS outcomes will therefore be reanalysed on all randomised individuals under a range of "missing not at random" scenarios. This will be done using the

formula $\Delta=\Delta_{CC} + (\delta_1 p_1 - \delta_0 p_0)$, where Δ_{CC} is the adjusted treatment effect in the complete case scenario (primary analysis), p_1 and p_0 are the proportion of missing outcomes, and δ_1 and δ_0 are the differences between the mean unobserved outcomes and mean observed outcomes in the CPAP and Best Supportive Care arms respectively. The standard error for Δ is approximately equal to the standard error for Δ_{CC} and so a confidence interval and p-value for Δ can be calculated.

Positive and negative values of δ_1 and δ_0 will be considered and varied simultaneously and separately. The resulting Δ will be displayed graphically with its confidence interval. The aim of this technique is to determine how sensitive the observed results are to different assumptions on the unobserved outcomes in the two treatment arms.

8. EXPLORATORY ANALYSES

8.1 Effect of CPAP adherence on ESS

Patients who were allocated to the CPAP arm at randomisation will be split into tertiles by their average CPAP usage in the last month of follow-up before the 3 month visit. Each group will then be compared to the BSC arm in a single model on the change in the primary ESS outcome. The minimisation variables will be adjusted for. A global test will be used to determine whether the treatment effect in each of the three CPAP groups differs.

A similar analysis will take place on the secondary ESS outcome, splitting patients into tertiles by their average CPAP usage in the last 3 months of follow-up before the 12 month visit.

The effect of CPAP usage on ESS at each timepoint will also be modelled using multivariable fractional polynomial models (5) adjusting for the minimisation variables. Since the BSC arm will not have compliance data the mean change in ESS in this arm will be displayed on the fractional polynomial plot.

8.2 Subgroup analyses

The effect of CPAP therapy on the primary ESS outcome will be compared separately by age, BMI and ESS and ODI at baseline. Each baseline variable will be categorised by its quartiles. The treatment effect in each subgroup will be estimated and compared using a global test for interaction. A continuous treatment effect plot will also be obtained from a fractional polynomial model (`mfpi` command in STATA) to show the treatment-covariate interaction in more detail (5, 6). The results from the two methods of analysis should be consistent; however, should the two models not agree this may be an indictor of an erroneous fractional polynomial model and so the results from the subgroup analysis will be used.

The effect of CPAP therapy on cognitive function (simple and four-choice reaction time) will be estimated in drivers and non-drivers. The treatment effects in the two subgroups will be formally tested for equality using an interaction test. Age and gender will also be adjusted for in this analysis.

Sleepiness during driving (whether nodded off whilst driving or pulled off the road due to sleepiness) will be compared at 3 and 12 months between treatment arms by driving habits (frequency of short local journeys and frequency long motorway journeys). Logistic regression models will be used.

Any reported road traffic accidents will be described in detail, with specific reference to the number of hours driving per week/month and the frequency of short local journeys vs motorway journeys > 1 hour.

In all subgroup analyses the minimisation variables will be adjusted for.

8.3 Exploratory analyses

Monthly diaries

A longitudinal analysis of the effect of CPAP therapy compared to BSC over the whole follow-up period will be performed by using the ESS scores from the monthly diaries. A multilevel model for repeated measures will be used with ESS as the response variable and month and baseline ESS as fixed effects with participant-specific and month-specific random intercepts (with the latter nested within the former). The model will make the assumption that all study visits and monthly diaries are completed on the expected dates. An unstructured covariance matrix will be used. From the model a plot of the treatment effect and its 95% CI at each month will be constructed.

9. REFERENCES

1. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep. 1991;14(6):540-5. Epub 1991/12/01.
2. Jenkinson C, Layte R, Wright L, Coulter A. The UK SF-36: An analysis and interpretation manual. 1996.
3. Townsend P. Poverty in the United Kingdom: a survey of household resources and standards of living: Harmondsworth: Penguin Books; 1979.
4. McDaid C, Griffin S, Weatherly H. The Continuous Positive Airway Pressure for the treatment of obstructive sleep apnoea-hypopnoea syndrome: a systematic review and economic analysis. Assessment report for the National Institute for Health and Clinical Excellence, 2008.
5. Royston P, Sauerbrei W. Multivariable model-building: a pragmatic approach to regression analysis based on fractional polynomials for modelling continuous variables: John Wiley; 2008.
6. Royston P, Sauerbrei W. A new approach to modelling interactions between treatment and continuous covariates in clinical trials by using fractional polynomials. Stat Med. 2004;23(16):2509-25. Epub 2004/08/03.

10. APPENDIX

10.1 Short SAQLI scoring manual

The Short Sleep Apnea Quality of Life Index (Short SAQLI) measures the effects of sleep apnea on a person's quality of life. The Short SAQLI has two parts; part I is completed by all subjects, but part II is completed by patients only if their sleep apnea condition has been treated.

Short SAQLI – Part I Scoring (Pre-Treatment)

A. Domains:
The 14 questions represent four quality of life domains.
Domain A – Questions #1,2,3,4 – Daily Activities
Domain B – Questions #5,6,7,8 – Social Interactions
Domain C – Questions #9,10,11 – Emotions
Domain D – Questions #12,13,14 – Symptoms

B. Scoring for Questions # 1 - 14:
Each questions has 7 response options that are scored as follows:

Score		Rating
7	=	not at all / no difficulty
6	=	a small amount
5	=	a small to moderate amount
4	=	a moderate amount
3	=	a moderate to large amount

2 = a large amount
1 = a very large amount

C. Total Score Calculation:
Add the individual scores for each of the 14 questions and divide the total by 14.

Short SAQLI – Part I and Part II Scoring (Post-Treatment)

A. The 14 questions in Section I represent four quality of life domains. The 4 questions in Section II represent treatment related problems.

Section I
Domain A – Questions #1,2,3,4 – Daily Activities
Domain B – Questions #5,6,7,8 – Social Interactions
Domain C – Questions #9,10,11 – Emotions
Domain D – Questions #12,13,14 – Symptoms

Section II
Domain E – Questions #15,16,17,18 – Treatment Related Side Effects

B. Scoring for Questions #1 - 14:
Each question has 7 response options that are scored as follows:

Score Rating
7 = not at all / no difficulty
6 = a small amount
5 = a small to moderate amount
4 = a moderate amount
3 = a moderate to large amount
2 = a large amount
1 = a very large amount

C. Scoring for Questions #15 – 17:
Each of these questions has 7 response options that are scored as follows:

Score Rating
0 = no problem
1 = a small problem
2 = a small to moderate problem
3 = a moderate problem
4 = a moderate to large problem
5 = a large problem
6 = a very large problem

D. Scoring for Question #18:
This score provides a weighting mechanism to reflect the trade-off between treatment related side effects compared with treatment benefits. It is scored as follows:

Score Rating
0.25 = no problem compared to the benefits
0.50 = a small problem compared to the benefits

0.75 = a small to moderate problem compared to the benefits
1.00 = about equal
1.00 = a moderate to large problem
1.00 = a large problem
1.00 = a very large problem compared to the benefits

E. Total Score – Calculation:
Step 1: Add the individual scores for each of questions #1 - 14.
Step 2: Add the individual scores for each of questions #15 - 17 and multiply this total by the weighting factor (score) for question #18.
Step 3: Subtract the Step 2 (questions #15-18) score from the Step 1 (questions #1-14) score.
Step 4: Divide the Step 3 score by 14.

Effect of Treatment:
Subtract the pre treatment total SAQLI score from the post treatment total SAQLI score.

10.2 Changes from version 1 to version .1.1

- The analysis of the OSLER (secondary outcome) has been simplified to allow an easier and more clinically meaningful interpretation of the results. An analysis using a survival model (as in version 1) is not necessary as censoring is not an issue and so treating OLSER time as a continuous measure is appropriate
- Exploratory treatment interaction analyses with baseline ODI and baseline ESS have been added
- Several exploratory analyses in section 8.3 have been removed

11. SIGNATURES OF APPROVAL

Date: 10/06/2013
Version: 1.1

Signatures

Name	Trial Role	Signature	Date
Mary Morrell	Chief Investigator		
Andrew Nunn	Trial Statistician		

Oxford Respiratory Trials Unit

PREDICT

Positive Airway Pressure in Older People: A Randomised Controlled Trial

HEALTH ECONOMIC ANALYSIS PLAN

Version 1

Version	Date	Comments
1	16/11/2011	Final draft

1. INTRODUCTION

This document details the planned health economic analysis for an investigator-blind, randomised controlled trial that compares continuous positive airway pressure plus best supportive care (CPAP) against best supportive care only (BSC) for treatment of obstructive sleep apnoea hypopnoea syndrome (OSAHS) in patients aged 65 and over. Full details of the background to the trial and its design are presented in the trial protocol. The aim of the health economic analysis is to estimate the relative cost-effectiveness of CPAP compared with BSC in this patient population.

In general, cost-effectiveness analyses that include only the results of a single trial can form an incomplete analysis with limited usefulness for decision makers (Sculpher et al 2006). The concerns with such analyses are that they do not use all relevant evidence, that there is often a limited number of comparators and a restricted time horizon (dictated by the follow-up of the trial). Nevertheless, in some cases, a cost-effectiveness analysis based on a single study is appropriate. There may be no previous studies in which CPAP has been compared with BSC in patients with OSAHS who are aged over 65. Therefore it could be argued that there is no additional evidence comparing the efficacy of CPAP with BSC in this patient group. However, previous studies that have assessed the effectiveness of CPAP or other interventions for OSAHS in younger patients may provide additional information for particular cost or health outcomes that are not expected to differ according to patient age. Therefore the health economic analysis may incorporate information from additional data sources outside of the PREDICT trial.

In some studies, the follow-up period of the trial may be adequate to capture the differential costs and benefits of the intervention and comparators. However, in this instance, the time horizon of one year may be insufficient to capture all the costs and benefits associated with the treatment of a condition such as OSAHS with possible long-term sequelae. The PREDICT trial may provide information on surrogate outcomes at 12 months (e.g. the impact of CPAP on cardiovascular function or neurocognitive decline), that could be associated with long-term benefits (e.g. a reduction in cardiovascular events or improved cognitive function). Any impact of a reduction in sleepiness on the incidence of rare events, such as road traffic accidents among patients that drive, may also be difficult to characterise within a trial-based analysis.

For these reasons the primary analysis will extrapolate, using a decision analytic model. This decision analytic model will combine data from the PREDICT trial with information from additional sources where appropriate, in order to estimate mean costs and outcomes over a lifetime time horizon, and to calculate an incremental cost effectiveness ratio (ICER) in terms of incremental cost per quality adjusted life year (QALY) gained.

The analyses described in this document will be performed by a health economist at the Centre for Health Economics, University of York in collaboration with statisticians at the MRC Clinical Trials Unit. The health economist responsible for writing this document in discussion with the co-chief investigators and other principal investigators and conducting the final analyses are:
Susan Griffin, CHE; Mark Sculpher, CHE.

2. DESIGN

2.1 Summary

This study is a two-arm, investigator-blind, parallel group, multi-centre randomised controlled trial. A total of 270 participants will be recruited to the study and followed-up for 12 months. At baseline participants will be randomised to receive CPAP plus BSC (active) or BSC only (control).

The co-primary endpoints will be:

1. Change in Subjective Sleepiness recorded as a mean Epworth Sleepiness Scale (ESS) measured at the end of months 3 and 4, answering the question 'does CPAP work at 3 months?'
2. Change in health related quality of life: Described by the EQ-5D, valued using UK population tariffs. This will be used to estimate the cost per QALY gained by providing CPAP in comparison to Best Supportive Care. The analysis will incorporate health care utilisation, including inpatient and outpatient hospital visits and GP visits during the trial.

3. OUTCOME MEASURES

3.1 Primary health economic outcome measure

Difference between the two treatment arms in cost effectiveness to be estimated within a decision analytic model. This will be used to estimate the cost per QALY gained by providing CPAP in comparison to BSC. Data from the PREDICT trial will be used to inform the parameters in the decision analytic model relating to:

- Health related quality of life, which will be characterised utilising EQ-5D data collected during the trial, valued using UK population tariffs. The analysis will make use of EQ-5D data collected at all time points within the trial in order to estimate quality-adjusted survival using the area under the curve approach.
- Health service costs, which will be characterised in terms of health care utilisation, including inpatient and outpatient hospital visits, GP visits and medication use collected during the trial.
- Treatment costs for CPAP, which will be characterised in terms of the equipment and support supplied to patients in the PREDICT trial, to be informed by clinical opinion.

3.2 Secondary health economic outcome measures

An alternative source of data from the PREDICT trial will be used to inform parameters in the decision analytic model relating to:

- Health related quality of life, which will be characterised utilising SF-36 data collected during the trial to estimate the SF-6D, valued using UK population tariffs.

4. CALCULATION OF PARAMETER VALUES

- The five components of the EQ-5D can be assigned level 1, 2 or 3. The resultant health states described by the EQ-5D will be scored using UK value set estimated in Dolan et al. 1997. If at least one of the components is missing the EQ-5D will also be set to missing.
- A sub-set of 11 items from the SF-36 form the SF-6D. The health states described by the SF-6D will be scored using the UK value set for cost-utility analyses (Model 10) estimated in Brazier et al. 2002. If at least one of the components is missing the SF-6D will also be set to missing.
- Descriptive statistics will be reported for change in EQ-5D and SF-6D scores at 12 months.
- Unit costs for the resource use items recorded on the patient questionnaire will be derived from sources relevant to the UK NHS. GP and hospital visits will be costed according NHS Reference Costs, medication will be costed according to the British National Formulary.
- The cost of CPAP equipment determined by UK price lists for machines, masks and sundries.
- The cost of sleep studies and nurse time required by treatment with CPAP to be determined by expert/clinical opinion.
- Costs will be expressed in current year GBP. The Health Service Cost Index will be used to adjust costs to the current price year where necessary.
- Descriptive statistics will be reported for each resource use item.
- Where multiple sources of information exist that could be used to inform a single parameter value these will be assessed for heterogeneity and, where appropriate, pooled using meta-analytic techniques.

5. DECISION ANALYTIC MODEL

A Markov model describing a series of health states and health events experienced by patients with OSAHS, according to their treatment, will be developed. This model will be used as the basis for extrapolating the costs and health outcomes to a more appropriate time horizon. A similar model was developed for a cost-effectiveness analysis of CPAP for the treatment of younger patients with OSAHS (Weatherly et al. 2009). The uncertainty around the parameter values in the decision analytic model will be fully characterised and propagated through to the model results by conducting probabilistic sensitivity analysis. This is achieved by characterising parameter values using distributions (parametric or empirical based on bootstrapping) rather than point estimates. The decision analytic model is then evaluated multiple times, each time selecting a new random draw from the assigned distributions, producing a distribution of model outputs.

The strategies of CPAP versus no CPAP therapy will be evaluated using standard cost-effectiveness analysis. If one strategy is not found to be dominant (i.e. less costly and more effective) in comparison to the other, then an ICER will be determined. The ICER will be based on the mean costs and mean QALYs estimated within the probabilistic sensitivity analysis of the decision model. Uncertainty around cost-effectiveness will be described using cost-effectiveness acceptability curves which describe the probability that an intervention is cost-effective (Fenwick et al. 2001).

The primary analysis of the decision analytic model will include only those outcomes recorded in the PREDICT trial relating to change in health-related quality of life and health care resource utilisation associated with CPAP relative to BSC. Secondary analyses will incorporate additional health states such as cardiovascular and cerebrovascular events, cognitive function and road traffic accidents. The risk of cardiovascular and cerebrovascular events will be informed by linking the change in cardiovascular risk factors, such as blood pressure, observed within the PREDICT trial to longer term outcomes using existing published risk equations. The risk of road traffic accidents will be informed by linking the change in sleepiness observed within the PREDICT trial to risk of accidents using previously published studies.

A discount rate of 3.5% per annum will be applied to both costs and QALYs in line with NICE guidance.

6. ANALYSIS PRINCIPLES

6.1 Sub groups

The cost-effectiveness of CPAP may differ according to baseline patient characteristics. Sub-groups will be defined according to baseline disease severity or other baseline characteristics that would be known when assigning treatment where these may influence the expected ICER of CPAP relative to BSC. Definition of sub-groups will be informed by clinical opinion, but may include for example, Epworth sleepiness score, ESS (e.g. 13 or less, or above 13) and Townsend disability score, TDS (e.g. 1 or less, or above 1). The decision analytic model will be re-evaluated for all relevant sub-groups.

6.2 Scenario analyses

The use of a decision analytic model can introduce uncertainty around the assumptions used, including the health states described, the selection of data sources and the methods used to combine multiple data sources. These aspects of modelling uncertainty will be explored using scenario analysis. The decision analytic model will be re-evaluated utilising alternative assumptions in order to assess the sensitivity of the ICER to these modelling assumptions.

6.3 Multiple Imputation

- Patterns of missing data will be presented
- Sensitivity to missing data will be assessed by comparing the characteristics of patients with missing items to those with complete data. The assumption that data are "missing completely at random" will be assessed by checking whether complete cases differ systematically from the original sample (Briggs et al. 2003).
- If the assumption of "missing completely at random" is inappropriate regression analysis can be used to adjust for data that are "missing at random". Multiple imputation of missing items will be undertaken in Stata using the "ice" command.
- Where multiple imputation is undertaken, the estimation of parameter values within the decision model will be based on the appropriate pooled statistic from analyses on each of the multiply imputed datasets.

7. REFERENCES

Sculpher MJ, Claxton K, Drummond M, McCabe C. Whither trial-based economic evaluation for health care decision making? Health Econ. 2006;15(7):677-87.

Dolan P, Gudex C, Kind P, Williams A. A social tariff for Euroqol: Results from a UK general population survey. York Centre for Health Economics Discussion Paper 1995 Sep; 138

Brazier J, Roberts J, Deverill M. The estimation of a preference-based meausre of health from the SF36. Journal of Health Economics. 2002;21:271-92.

Briggs A, Clark T et al. Missing....presumed at random: cost-effectiveness of incomplete data 2003 Health Econ 12(5) 377-392

Weatherly HLA, Griffin S, McDaid C, Duree KH, Davies RJO, Stradling JR, Sculpher MJ et al. An economic analysis of continuous positive airway pressure for the treatment of obstructive sleep apnea-hypopnea syndrome. International Journal of Technology Assessment in Health Care. 2009;25(01):26-34.

Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. Health Economics 2001;10:779-89

8. SIGNATURES OF APPROVAL

Date:
Version:

Signatures

Name	Trial Role	Signature	Date
Susan Griffin	Health economist		
Mark Sculpher	Health economist		