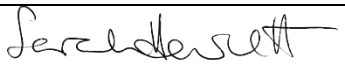
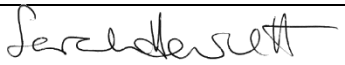
Bristol Randomised Trials Collaboration (BRTC)

RAFT: Reducing Arthritis Fatigue – clinical Teams using cognitive behavioural approaches

Statistical and Health Economic Analysis Plan

Version 10.0 (March 2015)

The following people have reviewed the Statistical Analysis Plan and are in agreement with the contents:			
Name	**Role**	**Signature**	**Date**
Keeley Tomkinson	Author	KT	21/04/15
Pete Blair	Supervisor		21/04/15
Lee Shepstone	Statistical Reviewer		15 – 4 - 2015
Sarah Hewlett	Chief Investigator		24 April 2015
The following people have reviewed the Health Economic Analysis Plan and are in agreement with the contents:			
Name	**Role**	**Signature**	**Date**
Joanna Thorn	Author		23/4/15
Will Hollingworth	Supervisor (Health Economics)		23.4.15
Lee Shepstone	Statistical Reviewer		15-4-2015
Sarah Hewlett	Chief Investigator		24 April 2015

Kimme Hyrich. Trialist Chair of the RAFT Data Monitoring and Ethics Committee.

Signature……  **……** **Date ……**16/04/2015

Abbreviations

RA	Rheumatoid Arthritis
CBT	Cognitive behavioural therapy
CB	Cognitive behavioural
CFS	Chronic fatigue syndrome
SCT	Social cognition theory
RCT	Randomised controlled trial
NHS	National health service
NRS	Numerical rating scale
BRAF - MDQ	Bristol Rheumatoid Arthritis Fatigue Multi-Dimensional Questionnaire
BRAF - NRS	Bristol Rheumatoid Arthritis Fatigue Numerical Rating Scales
OT	Occupational therapist
MHAQ	Modified Health Assessment Questionnaire
DAS28	Disease activity score
PDAS2	Patient self-reported Disease Activity Score
HAD	Hospital Anxiety and Depression scale
VLA	Valued Life Activities
EQ-5D-5L	EuroQol - Quality of life questionnaire
AHI	Arthritis Helplessness Index
RASE	Rheumatoid Arthritis Self-Efficacy
WPAI	Work Productivity and Activity Impairment scale
VAS	Visual Analogue Scale
CRP	C-reactive protein (inflammation biomarker)
GP	General practitioner

Contents

1. Introduction & Purpose

This document details the rules proposed and the presentation that will be followed, as closely as possible, when analysing and reporting the main results from the RAFT clinical trial: A multicentre randomised controlled trial of a group cognitive-behavioural (CB) intervention for fatigue self-management in rheumatoid arthritis (RA), delivered by the rheumatology clinical team in addition to usual care; compared to usual care (Arthritis research UK Fatigue booklet).

The purpose of the plan is to:

1. Ensure that the analysis is appropriate for the aims of the trial, reflects good statistical practice, and that interpretation of a priori and post hoc analyses respectively is appropriate.
2. Explain in detail how the data will be handled and analysed to enable others to perform the actual analysis in the event of sickness or other absence.

Additional exploratory or auxiliary analyses of data not specified in the protocol are permitted but fall outside the scope of this analysis plan (although such analyses would be expected to follow Good Statistical Practice).

The analysis plan will be made available if required by journal editors or referees when the main papers are submitted for publication. Additional analyses suggested by reviewers or editors will, if considered appropriate, be performed in accordance with the Analysis Plan, but if reported the source of such a post-hoc analysis will be declared.

Amendments to the statistical analysis plan will be described and justified in the final report of the trial.

2. Trial Synopsis

The information in this section is extracted from the study protocol in order to place the analysis plan within the context of the trial aims and methods. Please see the protocol for a full rationale.

2.1 Research Objectives

2.1.1 Primary

To assess whether there is a clinically important difference in the impact of fatigue 26 weeks after baseline (i.e. starting the intervention) between patients participating in a group cognitive-behavioural self-management course for RA fatigue delivered by the clinical rheumatology team using a detailed manual, in addition to usual care; compared to patients receiving usual care, which includes written fatigue self-management information.

2.1.2 Secondary

2. To compare differences between groups for secondary outcomes of fatigue severity, coping, anxiety and depression, sleep, helplessness, self-efficacy, pain, disability, valued activities, quality of life, work, health service use, acceptability, and cost-effectiveness for the NHS, patients and society.

3. To evaluate and control for potential demographic, psychological, and clinical predictors of fatigue change.

4. To evaluate persistence of effect (both short-term and long-term) over 2 years.

5. To explore whether clinical teams trained in cognitive behavioural approaches, perceive any positive or negative outcomes, particularly on their wider clinical practice.

2.2 Trial Design

Hospital based, multicentre (7 centres), 2 armed, randomised controlled trial. Individual participants are the unit of randomisation with fatigue assessed at 0, 6, 10, 18, 26, 52, 78, and 104 weeks.

2.3 Trial Centres

Trial patients will be recruited in 7 hospital outpatient rheumatology departments. Bristol Royal Infirmary will be the main centre where trial management is based.

2.4 Randomisation Procedure

Whenever a centre has completed the baseline visit for 10-16 participants randomization will be performed for that cohort. Randomisation will be managed by BRTC and is stratified by centre. Each centre will recruit 4 cohorts over 2 years. Allocation will be 1:1 but in the event of an odd number, the CB arm will receive an additional patient. If a patient randomised to the intervention group is not able to make the first date they will be offered the next course and have a new baseline recorded when the next group baselines are being conducted.

2.5 Sample Size Calculation

$$n = \frac{2\sigma^2}{d^2} f(\alpha, \beta) \qquad (1)$$

A standard sample size calculation uses equation 1, where σ^2 is the within group variance, d^2 is the standardised minimum clinically significant difference (MCSD) detectable between the two groups, and f(α,β) is the multiplying factor determined by the alpha level and the power.

The MCSD detectable is 1.46 units on a 0-10 VAS. A standard deviation of 2.7 minutes has been assumed from previous trial. With 90% power and 5% 2-sided alpha level, f(α,β) is equal to 10.5, thus the un-inflated sample size required is n = ((2*2.7^2/(1.46^2)*10.5) = 72 per arm.

[stata code: sampsi 0 1.46, sd(2.7) power(0.9) alpha(0.05)]

$$Design\ Effect = 1 + (m - 1) * ICC \qquad (2)$$

Due to the clustering nature of the RAFT trial, an intra-class correlation coefficient is required to calculate the design effect which is the inflation factor for the sample size. m is the minimum number of patients per group and an ICC of 0.01 has been used. Thus the design effect is:

(1 + (4-1)*0.01) = 1.03

Inflating 72 by the design effect gives n = (72*1.03) = 75 per arm. This is based on the minimum sample size per cluster of 4 and so this value of 75 needs to be inflated by 25% to give n = 94 per arm. However, the attrition rate over two years in unknown and so the trial has planned to recruit 150 patients per arm to ensure that the trial has enough power to conduct a 2 year analysis**. Therefore, the study requires a total of 300 patients, 150 per arm, with 7 centres running 4 cohorts each with average cohort size 10-12.**

(Please see study protocol for further details.)

2.6 Fidelity

Fidelity will be monitored in a random session of each of the 4 courses run in each centre by an independent observer.

2.7 Blinding

Blinding of patients and clinicians is not possible because of the need to engage patients in making cognitive and behavioural changes. Analysis will be performed blind to allocation by the primary statistician (KT) and health economist (JT). Interim results can be unblinded by the advisory statistician (PSB) if findings need to be presented to the Steering committee.

2.8 Safety Considerations/ Adverse Events

We do not anticipate any serious adverse events (SAE) to occur related to the intervention, but any events will be recorded in accordance with UH Bristol's Research Related Adverse Event Reporting Policy and reported to the CI, R&D departments, ethics committee, and TSC as appropriate.

3. Statistical Analysis

3.1 Software

It is anticipated that Stata will be used for exploratory analysis and Stata and/or MLWin for the main statistical analysis involving multilevel multivariable regression. SPSS may also be used for part of the analysis.

3.2 Outcome Measures

These outcome measures correspond directly to the research objectives defined in section 2.6.

Primary

1) The primary outcome measure is the fatigue impact at 26 weeks (after adjusting for baseline fatigue impact and centre).

Secondary

2) The main secondary outcome measures are the questionnaire scores assessing fatigue severity, coping, anxiety, depression, sleep, helplessness, pain, disability, valued activities, quality of life, work, health service use, acceptability at weeks 26 and 104.

3) Other secondary outcome measures include demographic, psychological, and clinical predictors of fatigue change, and will be measured at weeks 26 and 104.

4) The questionnaire score that determines fatigue impact at each time point will be used to assess persistence of effect over 2 years.

5) The evaluation of whether clinical teams trained in cognitive behavioural approaches, perceive any positive or negative outcomes, is examined through qualitative methods after the last course is delivered. Focus groups and one-to-one interviews will be held with tutors and data subjected to inductive thematic analysis.

Distributions

Where the distribution of the outcomes is approximately normal, mean values with standard deviations will be presented. For baseline characteristics, where the distribution of the outcome is not approximately normal, medians and interquartile ranges (IQR) will be presented, the t-test or Mann Whitney test will be used where appropriate in the univariable analysis. For binary/categorical variables, a number and percentage will be presented and the chi-square test (Fisher's Exact test for expected values < 5) will be used in the univariable analyses. For the continuously measured outcomes in the primary and secondary analyses, where outcomes are clearly non-normal, transforming to improve the normality of the residuals in the regression models would be explored. The choice of whether or not to transform variables, and if so which transformation to use, will be decided by considering: (1) the distribution of the variable, (2) the distribution of residuals from regression models, (3) the ease of interpreting results following any given transformation compared with no transformation and (4) whether main results/conclusions are influenced by the transformation or not. For skewed variables that result in markedly non-normal residuals in regression models, a natural log transformation would be used and the results compared with and without this transformation. If the overall conclusion was not altered by whether the variable is transformed or not, the untransformed (easier to interpret) version would be used. Where variables have been log-transformed, the resulting coefficients will be converted to differences in means on a % scale.

3.3 Variables

Centre Level

The variables recorded at the centre level include:

1. Centre location;

2. Centre tutors;

Subject Level

The variables recorded at the subject level include:

1. Gender;
2. Disease duration;
3. Ethnicity;
4. Comorbidities at baseline;
5. Age;
6. Socio economic status at baseline;
7. Medication change during trial;
8. Fatigue Score measurements from questionnaires;
9. Clinical status (anxiety, depression, sleep, helplessness, pain, disability, ability to undertake valued activities, quality of life VAS, disease activity; social interaction);
10. Group attended (CB arm only);
11. Attendance at each of 7 CB sessions (CB arm only);
12. Course evaluation (satisfaction, recommendation to a friend);
13. Quality of Life Adjusted Years (via EQ-5D-5L)
14. Costs of NHS care
15. Costs to patients, including work

3.4 Adherence

Every patient who attends session 1 of their CB course is defined as having received the intervention, regardless of whether they continue to attend. Adherence is measured by the number of sessions attended out of the full course of 7 sessions.

3.5 Quality Control

Any potential anomalous results that present in the dataset will be traced back to the raw data form and checked against this. Rigorous and conscientious methods will be maintained throughout the analysis of this trial to ensure that the results observed are of the highest possible quality. For example, we will minimise the amount of missing data by contacting the participants by telephone for all primary outcome data (fatigue impact NRS) and again if a returned questionnaire package has missing data. We will encourage all intervention participants to attend all sessions. We will check data in the initial stages of the trial to ensure that data are being collected in the right format and to pick up any systematic mistakes that might occur. Finally, we will also check 100% of entered data against the original questionnaire forms using two people and correct any errors, prior to closing the data set for analysis.

3.6 Baseline Characteristics

Baseline characteristics will be compared between the two arms by reporting relevant summary statistics in order to determine whether any potentially influential imbalance occurred, by chance, between the two arms. Characteristics will be reported as means (sd), medians (IQR) or number (%) depending of the nature of the data and its respective distribution as defined in section 3.2.. P-values will not be reported for differences between the two groups at baseline since appropriate randomisation methods have been used. Therefore, any differences identified would be due to chance such that a significant p-value would in reality be representative of a type 1 error (an incorrect rejection of the null hypothesis of no relationship when it is in fact true). Instead, it is better practice to identify differences in the two groups at baseline by their standard deviations; if the baseline characteristics of the two groups differ by more than half a standard deviation then the effect of this variable on the outcome will be investigated in a sensitivity analysis of the primary analysis of effectiveness.

3.7 Analysis of Effectiveness

Analysis Populations

• **Full Analysis set:** All randomised participants who received usual care, and (in those randomized to the CB arm) attended at least the first CB group session, and for whom at least one post-baseline assessment of the primary endpoint is available. It is analysis of this population which is in accordance with the "intention to treat" (ITT) principle.

• **Per protocol set:** All participants in the Full Analysis set who are deemed to have no major protocol violations (for example the unlikely event of a participant moving to another hospital and receiving CBT for RA fatigue) that could interfere with the objectives of the study.

Primary Outcome

Multivariable Analysis

The primary intention-to-treat analysis will involve between-arm comparisons for fatigue impact (BRAF-NRS Impact) at 26 weeks. These analyses of covariance will be implemented using mixed effect multivariable linear regression models with centre as a random effect. Standardised effect sizes for the primary outcome will be reported as adjusted mean difference divided by pooled baseline SD.

As well as the total sample results, descriptive data for 0 and 26 weeks will be presented by anonymised centres. The aim was not to compare between centres, therefore no inferential statistics will be applied, but if centre differences appear to exist, the potential clinical reasons and future implications will be discussed and reported.

Following this, covariates that differ at baseline by more than half a standard deviation will be added to the model to investigate their effect on the difference observed between the two groups in a sensitivity analysis.

Secondary Outcomes

Multivariable Analysis

Secondary outcomes will be analysed in the same way, including analysis of the total BRAF-MDQ fatigue score and then the 4 subscales, including preliminary multivariable analysis of

the effect of different attendance rates and patterns (per protocol complier average causal effect (CACE) analysis which includes only people who have adhered to the intervention [31]). Further analyses using repeated measures mixed effects ANCOVA models will examine the effect of interventions over time by including up to 4 follow-up scores (26, 52, 78, 104 weeks) per participant for the primary outcome, adjusting for baseline scores. Convergence/divergence between trial arms over time will be investigated by including appropriate interaction terms in the model. Clustering effect of delivery will be again investigated by including CB group and centre identifiers as additional levels.

Sub-group analyses: Where numbers allow, we will explore co-efficients of predictors of outcome. We will also explore potential cohort effects (ie whether patients in the 7 x 4th cohorts had a better fatigue outcome than those in the 7 x 1st cohorts) which might imply increasing tutor skills by cohort 4).

3.8 Exploratory Analysis

The dataset will be used to explore other RA fatigue questions, although it cannot give definitive answers as the study was not designed to answer such questions. We anticipate that exploratory analysis will include the exploration of three questions. First, which particular components of self-efficacy are changed by the intervention and whether there is a relationship with changes in fatigue or pain (the RASE has 28 items that ask about beliefs regarding self-management tasks, and many of these are covered intensively during the intervention and loosely in usual care)[32]. Second, the Impact Triad[17] is the concept that impact from a condition is a combination of severity, coping and personal context. These data could be used to examine whether disease severity (pain, DAS28), fatigue coping (BRAF NRS coping) and personal context (using surrogates of co-morbidities, work status, age) combine to explain fatigue impact. Third, we will investigate how multiple causes interact to drive fatigue[18].

We will also investigate how the BRAF scores change between the final session (week 6), week 10, the consolidation session (week 14) and week 18 to see if there is a drop in BRAF score after the course finishes and whether this increases again after the consolidation session.

3.9 Missing Data

Missing data may arise because some participants will not return their questionnaires at different data collection periods. It is anticipated that proportions with missing data for any particular measure will be similar in the two randomised groups as it was in the original RCT but this will be checked in case loss to follow-up is greater in the control arm. Multiple imputation techniques[z] will be used to generate missing values in the dataset if more than 5% of the data is missing. Potential methods include using the mode (conservative estimate) or using variables to predict most likely values by drawing 1000 samples and taking the mean.

3.10 Sensitivity Analysis

Several sensitivity analyses will be conducted to test the robustness of the results from the statistical analyses, and in some cases, to increase understanding of the relationship between the dependent and independent variables. Sensitivity analyses of primary outcome will be conducted by:

S1) Participants with any covariate values that are greater than +/- 4 standard deviations will be examined and removed from the dataset if no explanation can be found and the statistical analyses repeated;

S2) A per protocol CACE (complier average causal effect) analysis to assess whether the intervention was more effective to those who adhered to the intervention;

S3) Additional adjustment for any variables displaying imbalance at baseline;

S4) Fitting multi-level mixed effects models to investigate any clustering effect from delivery in groups and centres;

S5) Multiple imputation techniques to investigate the impact of missing data, based on 20 imputed datasets, with baseline fatigue severity, impact, pain and disease activity added to the imputation model as variables predictive of missingness.[33]

(Reproduced from Hewlett et al,[12] Self-management of fatigue in Rheumatoid Arthritis: a randomised controlled trial of group cognitive-behavioural therapy. Ann Rheum Dis 2011;70:1060-7 with permission from BMJ Publishing Group Ltd.)

4. Health Economic Analysis

4.1 Health economic analysis plan purpose

The purpose of this health economic analysis plan is to set out in detail the analysis and reporting procedure intended for the economic analyses to be undertaken in the RAFT trial. While the intentions outlined in this plan will be followed as closely as possible, the plan also describes the circumstances under which amendments are permitted and the documentation of such changes; any deviations from this plan will be justified in the final report.

The analysis plan is designed to ensure that there is no conflict with the protocol and statistical analysis plan (SAP) described in section 3 and should be read in conjunction with them.

4.2 Trial overview

A synopsis of the trial design, centres, study population, intervention and sample size is provided in section 2.5. The primary outcome measure for effectiveness is the impact of fatigue measured using the BRAF-NRS Impact instrument (a 0-10 Numerical Rating Scale)[15,16] at 20 weeks after the delivery of session 6 (approximately 26 weeks post-randomisation).

4.3 Economic analysis background

Aim

The aim is to determine the cost-effectiveness of delivering a cognitive behavioural programme in a group setting to RA patients with fatigue alongside usual care, compared with patients receiving usual care only.

Perspective

The primary economic analysis will be from the societal perspective, including costs incurred by the NHS, personal social services (PSS), personal travel costs and costs arising from loss of productivity. A secondary analysis will restrict the perspective to NHS and PSS costs.

Time horizon

The primary economic analysis will compare the costs and benefits of each arm over the first 26 weeks of follow up. A secondary analysis will extend this to compare costs and benefits over the two-year follow-up period from randomisation.

4.4 Economic measurements

Measurement of outcomes

The primary economic outcome measure will be Quality Adjusted Life Years (QALYs) derived from utility scores, obtained using the EQ-5D-5L quality of life instrument[19]; measurements will be recorded at baseline and at 6, 26, 52, 78 and 104 weeks post-randomisation. At baseline and week 26, the patient will complete the instrument during a visit to the research nurse; at all other time points, the questionnaires will be distributed and returned by post, with a telephone reminder.

Fatigue impact at 26 weeks will be measured as the primary effectiveness outcome of the trial, and will be used in a secondary economic analysis (using an NHS and PSS perspective). Fatigue impact will be measured using the BRAF-NRS Impact instrument at the same time points as the utility measures, and additionally at 10 and 18 weeks post-randomisation; all fatigue impact data will be collected by the trial secretary by telephone.

Identification of relevant resource use

Data will be collected for resource use that is related to RA or RA-related fatigue only. Routine monthly blood monitoring visits will be excluded, whether performed in primary or secondary care as these are considered not to be related to the intervention. For the NHS and PSS perspective, data will be collected on arthritis medication, steroid injections, primary and community medical care, physiotherapy, occupational therapy, podiatry or nursing care, secondary care, use of rheumatology telephone helpline and social care by the patient. Staff

time and other expenses incurred during the training and delivery of the CB intervention will be recorded.

For the analysis from the societal perspective, we will additionally collect data on work productivity (presenteeism and absenteeism) and patient travel costs.

Measurement of resource use

Staff training

Time logs and expense forms (see appendix A parts 1 and 2) will be used to track all resources used in the delivery of the 4 day training programme including trainee and trainer time (and preparation time), travel costs and course materials to calculate the fixed cost of training (including support during the observed practice course).

Delivery of intervention

For the delivery of the CB sessions, NHS resources will be captured in staff time logs (see appendix A parts 1 and 2 and appendix B) recording session preparation, delivery, de-briefing, supervision time and materials. Group size and identities of attendees at each practice and CB session will be recorded so that these shared costs can be allocated per patient.

Health care utilisation

NHS primary and secondary care and patient personal resource use during the 2 year follow-up will be captured using patient-reported questionnaires and nurse completed forms at weeks 0, 6, 26, 52, 78 and 104 (see appendix C part 1 and 2). Patients will complete the questionnaires at their research nurse visits at weeks 0 and 26. At all other time points the patient questionnaires will be administered by post; the research secretary will remind patients by telephone to expect the questionnaire, and will remind patients by telephone that the questionnaire needs returning if it has not been received within 2 weeks. Medications will be recorded by the research nurse at the baseline visit; in each follow up questionnaire patients will be asked to indicate changes from the previous date. Time for delivery of the usual care session prior to randomisation will be recorded on researcher timesheets (see appendix D).

Transport

At the baseline visit the research nurse will document the patient's normal transport method for hospital, GP or other NHS appointments and the cost (for public transport) or mileage (for private transport), to use as a multiplier for calculating costs.

Productivity

Work disability, presenteeism and absenteeism will be captured by the Work Productivity and Activity Impairment scale (WPAI)[20] (see appendix E).

Valuation of resource use

Unit costs for NHS staff time to train for and deliver the intervention will be based on the most recently available national estimates (*e.g.* Curtis[21]). Actual expenses incurred for training materials, refreshments and staff travel will be recorded. The costs of medications, community, primary and secondary care (including usual care) during follow up will be based on the most recently published national tariffs where available (*e.g.* DOH[22]; BNF[23]), supplemented by micro-costing or local estimates if necessary. Productivity costs due to RA and fatigue will be estimated based on average weekly earnings stratified by age (*e.g.* ONS[24]). Resource use will be combined with unit costs to estimate the incremental cost or savings of the group CB programme over the 2 year period.

Costs and outcomes occurring during the second year of follow up will be discounted in line with NICE guidance (currently 3.5%)[25].

4.5 Economic analyses

The primary (societal) and secondary (NHS and PSS) analyses will follow the same analysis plan unless mentioned otherwise. All analyses will be conducted using intention-to-treat principles, comparing the two groups as randomised and including all patients in the analysis wherever practical.

Data cleaning and missing costs and outcomes

Data cleaning and imputation will be undertaken prior to unblinding by the economic researcher.

Data cleaning will include correction of obvious 'free text' response errors (e.g. misspelt drug names), group coding of similar resource items (e.g. 'orthopaedics' and 'trauma & orthopaedics' clinics) to enable unit costing and simple imputation of data missing minor details (e.g. missing drug dose) based on reasonable assumptions (e.g. the most commonly prescribed dose). Any areas of uncertainty will be discussed between two health economists and, where necessary, referred for adjudication by a clinical expert. The number of uninterpretable responses (e.g. illegible drug names) will be reported and not included in the cost analysis. However, questionnaires will not be classed as 'missing data' for the cost analysis, unless the questionnaire is not returned or the majority of responses are uninterpretable.

The primary analysis will include all participants using imputation to predict missing costs and outcomes[26]. For EQ-5D-5L scores, multiple imputation techniques[31] will be used to generate missing values in the dataset if more than 5% of patients have missing data. Cruder methods (e.g. treatment group mode or last value carried forward) may be used if missing values are less frequent. Multiple imputation of missing NHS & PSS, travel costs, and productivity losses will be performed separately and for each follow up time point.

Analysis of costs

All costs will be reported in 2016/2017 pounds sterling. The cost of each resource item will be calculated by multiplying the number of resource units used by the unit cost. The total cost for each individual patient will then be estimated as the sum of the cost of resource-use items consumed.

Overall mean costs, stratified by NHS & PSS, patient and productivity costs, and standard deviations for both arms of the trial will be calculated. We will estimate incremental mean difference in total costs between the two arms of the trial and 95% confidence intervals. Bootstrapping techniques will be used to derive confidence intervals[27].

Analysis of economic outcome measures

EQ-5D-5L utility scores will be used to estimate Quality Adjusted Life Years (QALYs) over the 2 year period, adjusting for any imbalances in baseline scores[28]. We will report the incremental mean difference in QALYs between the two arms of the trial and 95% confidence intervals.

Analysis of relative costs and outcomes

Cost and QALY data will be combined to calculate an incremental cost-effectiveness ratio (ICER) and net monetary benefit (NMB) statistic[30].

$$NMB_i = \lambda E_i - C_i$$

For each individual *i*, the NMB statistic is given as the societal willingness to pay for a QALY, λ, multiplied by the patient outcome E_i (i.e. QALYs), from which the total cost C_i is subtracted. In the primary analysis we will estimate whether the group CB programme is cost-effective at the established NICE threshold of £20,000 per QALY gained.

Uncertainty in the point estimate of cost per QALY will be quantified using bootstrapping methods to calculate confidence intervals around the ICER and NMB. The probability that the group CB programme is cost-effective at various 'willingness to pay for a QALY' thresholds will be depicted using a cost-effectiveness acceptability curve[28a].

4.6 Further economic analyses

Subgroup analyses

Currently no subgroup analyses are planned. Such analyses may be undertaken for any clinical subgroups where there is a strong a priori belief of effect modification. Any such analyses will be identified prior to unblinding, and recorded in this section or clearly reported as post-hoc analyses.

Sensitivity analyses

One way sensitivity analyses will be used to judge the potential impact of other sources of uncertainty. The following analyses will be conducted.

- Rare high-cost events *(e.g.* inpatient stays, high cost medication use) will be identified. Two clinicians (unaware of randomised treatment allocation) will discuss and adjudicate on whether the events are likely to be related to fatigue or not; a sensitivity analysis dropping events thought unlikely to be related to fatigue will be conducted.
- A sensitivity analysis dropping supervisory travelling costs will be conducted to estimate the impact of local psychologists taking over supervision in a roll-out situation.

Value of information analysis

There is no plan to conduct a value of information analysis within the scope of this study.

4.7 Updating the economic analysis plan

Changes to existing analyses

Dated changes to the analysis plan will be documented in this section. Circumstances under which changes will be permitted are as follows.

- Development of novel statistical methods that are deemed more appropriate for this analysis.
- Clarification of currently debated issues.
- Preliminary data cleaning or analysis (conducted prior to unblinding) suggesting that planned analyses are sub-optimal.

Post hoc analyses

Any suitable analyses that are identified after unblinding or during the refereeing process will be listed in this section, dated and the source will be identified. Such analyses will be identified clearly as *post hoc* analyses in trial reports.

References

1 Suurmeijer T, Waltz M, Moum et al. Quality of life profiles in the first years of RA. Arth Rheum (Arth Care Res) 2001;45:111-21

2 Luqmani R, Hennell S, Estrach C, et al. British Society of Rheumatology & British Health Professionals in Rheumatology guidelines for the management of RA (first 2 years). Rheumatology 2006:45:1167-9

3 Luqmani R, Hennell S, Estrach C et al. British Society for Rheumatology and British Health Professionals in Rheumatology guideline for the management of RA (after the first two years). Rheumatology 2009;48:436-9

4 Symmons D, Turner G, Webb R et al. The prevalence of RA in the United Kingdom. Rheumatology 2002;41:793-800

5 Wolfe F, Hawley D, Wilson K. The prevalence and meaning of fatigue in rheumatic disease. J Rheumatol 1996;23:1407-17

6 Van Hoogmoed D, Fransen J, Bleijenberg G et al. Physical and psychosocial correlates of severe fatigue in RA. Rheumatology 2010;49:1294-1302

7 Hewlett S, Cockshott Z, Byron M, Kitchen K, Tipler S, Pope D, Hehir M. Patients' perceptions of fatigue in RA: Overwhelming, uncontrollable, ignored. Arthritis Rheum (Arthritis Care Res) 2005;53:697-702

8 Cramp F, Hewlett S, Almeida C, Kirwan JR, Choy EHS, Chalder T, Pollock J, Christensen R. Non-pharmacological interventions for fatigue in RA. Cochrane Database of Systematic Reviews 2013 (in press)

9 Van Kessel K, Moss-Morris R, Willoughby E et al. An RCT of CBT for MS fatigue. Psychosom Med 2008;70:205-13

10 O'Dowd H, Gladwell P, Rogers CA et al. CBT in chronic fatigue syndrome. Health Technology Assessment 2006;10:37

11 Sage N, Sowden M, Chorlton E, Edeleanu A. What is the Cognitive Behavioural Approach? In: Cognitive behaviour therapy for chronic illness and palliative care. Chichester, Wiley, 2008. pp3-9

12 Hewlett S, Ambler N, Cliss A, Almeida C, Knops B, Pope D et al. Self-management of fatigue in Rheumatoid Arthritis: a randomised controlled trial of group cognitive-behavioural therapy. Ann Rheum Dis 2011;70:1060-7

13 Dures E, Ambler N, Fletcher D, Pope D, Knops B, Hewlett S. Why do we need to pilot interventions? Essential refinements identified during pilots of a fatigue intervention [abstract]. Rheumatology 2012;51(S3):iii167-8

14 Arnett F, Edworthy S, Bloch D et al. The ARA 1987 revised criteria for the classification of RA. Arth Rheum 1988;31:315–24

15 Nicklin J, Cramp F, Kirwan J, Urban M, Hewlett S. Measuring fatigue in RA: A cross-sectional study to evaluate the BRAF Multi-Dimensional Questionnaire, VAS and NRS. Arthritis Care Res 2010:62:1559-68

16. Nicklin J, Cramp F, Kirwan J, Urban M, Hewlett S. Collaboration with patients in the design of patient reported outcome measures: Capturing the experience of fatigue in RA. Arthritis Care Res 2010:62:1552-8

17. Sanderson T, Hewlett S, Flurey C, Dures E, Richards P, Kirwan J. The impact triad (severity, importance and self-management) as a method of enhancing the measurement of the personal life impact of rheumatic diseases (Editorial). J Rheumatology 2011; 38: 2: 191-4

18. Hewlett S, Chalder T, Choy E, Cramp F, Davis B, Dures E, Nicholls C, Kirwan J. Fatigue in Rheumatoid Arthritis: Time for a conceptual model (Editorial). Rheumatology 2011; 50: 1004-6

19. Herdman M, Gudex C, Lloyd A et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Quality of Life Research 2011;20:1727-36

20. Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. PharmacoEconomics. 1993;4:353-65

21. Curtis L. Unit Costs of Health & Social Care 2011. PSSRU, Univ Kent, 2011 http://www.pssru.ac.uk/pdf/uc/uc2011/uc2011.pdf

22. Dept Health. Reference Costs 2009-10 (ref 15423). 2011 http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_123459

23. British National Formulary http://www.bnf.org/bnf/index.htm

24. Office for National Statistics. Annual Survey of Hours and Earnings Pension Tables, 2011 (SOC 2010) http://www.ons.gov.uk/ons/rel/ashe/annual-survey-of-hours-and-earnings-pension-tables/2011-provisional-results--soc-2010-/index.html

25.NICE http://www.nice.org.uk/media/D45/1E/GuideToMethodsTechnologyAppraisal2013.pdf

26. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward M et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;338:b2393
27. Ng E S-W, Grieve R and Carpenter HR. Two-stage nonparametric bootstrap sampling with shrinkage correction for clustered data," Stata Journal 2013;13(1):141-164
28. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. Health Economics 2005;14:487–96
29. Hoch J, Briggs A, Willan A. Something old, something new, something borrowed, something blue: a framework for the marriage of health econometrics and cost-effectiveness analysis. Health Econ 2002;11:415-30
30. Löthgren M and Zethraeus N. Definition, interpretation and calculation of cost-effectiveness acceptability curves. Health Economics 2000;9(7):623-630
31. A Connell. Employing Complier Average Causal Effect Analytic Methods to Examine Effects of Randomised Encouragement Trials. Am J Drug Alcohol Abuse 2009;35(4):253-259
32. Hewlett S, Cockshott ZC, Kirwan JR et al. Development and validation of a self-efficacy scale for use in British patients with RA (RASE). Rheumatology 2001;40:1221-30
We are interested in which behaviours change.
33. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward M et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;338:b2393

Appendix A part 1: Trainer Weekly Time-Sheet for time spent on RAFT activities: Clinical only (not research)

Name …… **Week beginning ……**

	Preparing training. Minutes	Delivering Training. Eg 9-1 & 2-5	Practice run. Minutes	Clinical observation. Minutes	Travel time (to & from). Minutes	Travel costs. Miles/ public transport	Clin supervision (eg phone call). Minutes	Extra tutor support Explain & Minutes	Other clinical RAFT activity. Minutes	Comments
Monday			Session No: Centre: Mins	Session No: Centre: Mins	To: Return:					
Tuesday			Session No: Centre: Mins	Session No: Centre: Mins	To: Return:					
Wednesday			Session No: Centre: Mins	Session No: Centre: Mins	To: Return:					
Thursday			Session No: Centre: Mins	Session o: Centre: Mins	To: Return:					
Friday			Session No: Centre: Mins	Session No: Centre: Mins	To: Return:					
Overtime			Session No: Centre: Mins	Session No: Centre: Mins	To: Return:					

Appendix A part 2: Tutor Weekly Time-Sheet for time spent on RAFT activities: Clinical only (not research)

Name …… Centre …….. Week beginning ……

	Preparing RAFT session Minutes	Delivering RAFT Session Minutes	De-briefing RAFT session Minutes	Clinical supervision (eg phone call) Minutes	Extra support for patients for fatigue Minutes	RAFT training (clinical only, not research Eg 9-1 & 2-5	Travel time (to & from) Minutes	Travel costs Miles/ public transport	Any other clinical RAFT activity; or other comments
Monday		Session No: 1 2 3 4 5 6 7			RAFT pt: Control pt:		To: Return:		
Tuesday		Session No: 1 2 3 4 5 6 7			RAFT pt: Control pt:		To: Return:		
Wednesday		Session No: 1 2 3 4 5 6 7			RAFT pt: Control pt:		To: Return:		
Thursday		Session No: 1 2 3 4 5 6 7			RAFT pt: Control pt:		To: Return:		
Friday		Session No: 1 2 3 4 5 6 7			RAFT pt: Control pt:		To: Return:		
Overtime		Session No: 1 2 3 4 5 6 7			RAFT pt: Control pt:		To: Return:		

Appendix B: Costs per RAFT course (clinical, not research) Tutor to complete

Name …… Centre…….. Course: Training 1 2 3 4 5 6 7 (circle)

	Participants **Expected & Actual**	**Photocopying** **No of sheets**	**CDs given** **Number**	**Tea/Coffees** **Number**	**Tutors present** **Number**	**Other expenses** (out of pocket expenses, minor purchases) **Explain & Cost**	**Comment or explanation**
Session 1	Exp Act						
Session 2	Exp Act						
Session 3	Exp Act						
Session 4	Exp Act						
Session 5	Exp Act						
Session 6	Exp Act						
Session 7	Exp Act						
Other Event							

Appendix C part 1: Nurse completed secondary care data

Date _____ **Patient ID Number**_____

RAFT STUDY

Secondary Care Data

Date randomised _____ **Date of last check** _____

Week Number: 0 26 52 78 104 (please circle)

- **List all events related to arthritis/arthritis fatigue since the last check.**
- **Do not include monthly blood monitoring visits.**

Event	In-Patient / Out-Patient	Who	Start Date	End Date (for IP)
E.g. Rheumatology Appt / Orthopaedic	*e.g. IP*	*E.g. Dr/ Nurse/ OT/ PT*	*1/1/14*	*3/1/14*

Please send this sheet to Celia Almeida in Bristol

Appendix C part 2: Primary care questions taken from patient questionnaire weeks 0, 6, 26, 52, 78 and 104

We are interested in any care you have received for your <u>arthritis</u> outside the hospital (that is, in the community).

1) In the last 6 months, approximately how many times have you seen any of these people about your arthritis?

	Number of occasions
GP	
District nurse or practice nurse (do NOT include your routine blood monitoring)	
Community physiotherapist	
Community Occupational Therapist (OT)	
Community Podiatrist / Chiropodist (NHS)	
Other (state who, eg Private Podiatrist / Chiropodist)	

2) Have you used your rheumatology nurse helpline in the last 6 months?

No _____ Yes, approximately _____ times We don't have one_____

3) How many days sick leave have you taken from work as a result of your arthritis or arthritis fatigue in the last 6 months ?

None _____ Approximately _____ days Not applicable _____

4) We are interested in any social care you have received because of your <u>arthritis or arthritis fatigue</u>.

During the past seven days how much help or care have you received from the following:

	Number of hours in the 7 days
Professional help for everyday activities (e.g. bathing, dressing eating, housework, shopping etc.) provided or paid for by the Government	
Professional help for everyday activities (e.g. bathing, dressing eating, housework, shopping etc.) that you paid for	
Help for everyday activities (e.g. bathing, dressing eating, housework, shopping etc.) provided by friends or family	

5) We are interested in other hospital visits - have you been to hospital for either of the following since _____ because of your arthritis or arthritis fatigue?

Casualty (A&E) **No**

Yes **Date:** _____ **Reason**_____

Date: _____ **Reason**_____

Xray (imaging) **No**

Yes **Date:** _____ **Reason**_____

Date: _____ **Reason**_____

Appendix D: Researcher Weekly time-sheet for time spent on RAFT activities (clinical, not research)

Name …… Centre…… Week beginning ……

	Booklet session Minutes	Discussions with patient seeking fatigue help Minutes	Discussions with clinicians re: patient seeking help Minutes	Other clinical RAFT activity (state) Minutes	Travel time for CLINICAL RAFT work Minutes	Travel costs for CLINICAL RAFT work Miles/ public transport	Comment or explanation
Monday		RAFT pt: Control pt:	RAFT pt: Control pt:		To: Return:		
Tuesday		RAFT pt: Control pt:	RAFT pt: Control pt:		To: Return:		
Wednesday		RAFT pt: Control pt:	RAFT pt: Control pt:		To: Return:		
Thursday		RAFT pt: Control pt:	RAFT pt: Control pt:		To: Return:		
Friday		RAFT pt: Control pt:	RAFT pt: Control pt:		To: Return:		
Overtime		RAFT pt: Control pt:	RAFT pt: Control pt:		To: Return:		

Appendix E: Work Productivity and Activity Impairment scale (WPAI)[20] taken from patient questionnaires weeks 0, 6, 26, 52, 78 and 104

What are the effects of arthritis fatigue on your activities?

The following questions ask about the effect of your arthritis fatigue on your ability to work and perform regular activities. By arthritis fatigue problems we mean any physical or emotional problem or symptom.

Please fill in the blanks or circle a number, as indicated.

1. Are you currently employed (working for pay)? ____ NO ____ YES

   If NO, tick "NO" and skip to question 6.

The next questions are about the **past seven days**, not including today.

2. During the past seven days, how many hours did you miss from work because of your arthritis fatigue? *Include hours you missed on sick days, times you went in late, left early, etc., because of your arthritis fatigue. Do not include time you missed to participate in this study.*

   _____HOURS

3. During the past seven days, how many hours did you miss from work because of any other reason, such as vacation, holidays, time off to participate in this study?

   _____HOURS

4. During the past seven days, how many hours did you actually work?

   _____HOURS *(If "0", skip to question 6.)*

5. During the past seven days, how much did your arthritis fatigue affect your productivity <u>while you were working</u>?

Think about days you were limited in the amount or kind of work you could do, days you accomplished less than you would like, or days you could not do your work as carefully as usual. If arthritis fatigue affected your work only a little, choose a low number. Choose a high number if arthritis fatigue affected your work a great deal.

Consider only how much arthritis fatigue affected productivity <u>while you were working</u>.

Arthritis fatigue had no effect on my work	0 1 2 3 4 5 6 7 8 9 10	Arthritis fatigue completel prevented me from worki

CIRCLE A NUMBER

6. During the past seven days, how much did your arthritis fatigue affect your ability to do your <u>regular daily activities</u>, other than work at a job?

By regular activities, we mean the usual activities you do, such as work around the house, shopping, childcare, exercising, studying, etc. Think about times you were limited in the amount or kind of activities you could do and times you accomplished less than you would like. If arthritis fatigue affected your activities only a little, choose a low number. Choose a high number if arthritis fatigue affected your activities a great deal.

Consider only how much arthritis fatigue affected your ability to do your regular daily activities, other than work at a job.

Arthritis fatigue had no effect on my daily activities	0 1 2 3 4 5 6 7 8 9 10	Arthritis fatigue completely prevented me from doing my daily activities

CIRCLE A NUMBER