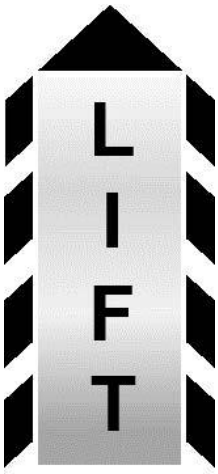


Lessening the impact of fatigue in inflammatory rheumatic diseases: a randomised clinical trial

Statistical Analysis Plan

Funded by Versus Arthritis

1 Administrative information

This SAP is based as far as is appropriate on guidelines given in JAMA. 2017;318(23):2337-2343. doi:10.1001/jama.2017.18556

TRIAL FULL TITLE	**Lessening the impact of fatigue in inflammatory rheumatic diseases: a randomised clinical trial**
EUDRACT NUMBER	n/a
SAP VERSION	Version 1 (based on Protocol LIFT (*08.10.2020 version 11)*)
IRAS ID	216267
Clinicaltriaslgov Number	NCT03248518
SAP VERSION DATE	6th January 2021
TRIAL STATISTICIAN	Dr Lorna Aucott
TRIAL CHIEF INVESTIGATOR Co-Chief Investigator	Prof Gary Macfarlane Dr Neil Basu
SAP AUTHOR	Dr Lorna Aucott (Senior Statistician CHaRT)

1.1 SAP Signatures

I give my approval for the attached statistical analysis plan (SAP) for the randomised controlled trial entitled LIFT, Version: 1 Dated: 06/01/2021

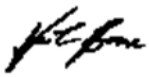
Chief Investigator

Name: Prof Gary Macfarlane

Signature: 

Date: 6 January 2021

Name: Co-Chief Investigator Dr Neil Basu

Signature: 

Date: 6 January 2021

Statistician

Name: Dr Lorna Aucott

Signature: 

Date: 6 January 2021

Table of Contents

1.2 Abbreviations and Definitions

AE	Adverse events
ASDAS	Ankylosing Spondylitis Disease Activity Score
AxSpA	Axial Spondyloarthritis
BILAG	British Isles Lupus Activity Group
BRAF-MDQ	Bristol Rheumatoid Arthritis Fatigue Multi-Dimensional Questionnaire
CACE	Complier Average Causal Effect
CBA	Cognitive Behavioural Approach
CF	Chalder Fatigue Scale
CHaRT	Centre of Healthcare Randomised Trials
CI	Chief Investigator
CONSORT	Consolidated Standards of Reporting Trials
CRF	Case Report Form
CRP	C-reactive protein
CTU	Clinical Trials Unit
DAS28	Disease Activity Score 28
DMC	Data Monitoring Committee
FSS	Fatigue Severity Scale
HADS	Hospital anxiety and depression scale
IRD	Inflammatory Rheumatic Disease
ITT	Intention to treat
MAR	Missing at random
PA	Physical Activity
PEP	Personalised Exercise Programme
QOL	Quality of Life
RA	Rheumatic Arthritis
RCT	Randomised Controlled Trial
rHCPs	Rheumatology Health Care Professionals
S-VLA	Valued Life Activities short form
SAE	Serious adverse events
SLE	Systemic Lupus Erythematosus
TSH	Thyroid Stimulating Hormone
VBM	Voxel-based morphometry
VO2	Volume oxygen
WPAI:SHP	Work Productivity and Activity Impairment Questionnaire: Specific Health Problem

2 Introduction

Despite advances in management of inflammatory rheumatic diseases (IRDs), patients remain burdened by their disease and cite fatigue as a principal problem, equal to pain in terms of burden. Fatigue is a crucial determinant of impaired quality of life (QOL) and a predictor of work disability and indeed the main barrier to remaining in employment. Patients feel this symptom is clinically ignored with rheumatologists admitting ignorance regarding its management.

There is now considerable consensus across the health care community that non-pharmacological interventions, specifically cognitive behavioural approaches (CBA) and programmes designed to support increased physical activity, are valuable treatments to help IRD patients manage the functional challenges such as fatigue.

This statistical analysis plan (SAP) documents the planned analysis for the main Lift Trial

2.1 Study Aims and Objectives

- To test our hypothesis that usual care in addition to either standardised cognitive behavioural approach (CBA) or personalised exercise programme (PEP) interventions is more effective than usual care alone to lessen the impact and severity of fatigue after 56 weeks from baseline. Please see the protocol for the primary here and then add secondary hypotheses/research questions.

2.2 Study Design

- The LIFT study is a multi-centre, three-arm pragmatic randomised controlled trial testing usual care alone versus usual care with additional adapted CBA or PEP therapies, figure 1

2.3 Interventions to be evaluated (All arms are <u>fully</u> defined in the protocol)

- *Usual care:* Arthritis Research UK's information booklet[18] for self-management of fatigue represents usual care in almost all UK rheumatology centres and is freely available. It covers the major relevant topics underpinned by goal-setting and self-monitoring of activity. It encourages that patients ask for support to work through the booklet.
- Both active interventions will last 14 weeks with a booster at 22 weeks. (protocol section 4.1 and figure 2)
- The *Cognitive behavioural approach (CBA)* is a structured psychological intervention, aiming to replace unhelpful beliefs and behaviours with more adaptive ones. It will use patient-centred strategies and behavioural activities, supported by written materials and regular consultations with rheumatology health care professionals. Participants will receive additional leaflets and diaries about making changes to manage fatigue. The times and duration of keeping the diary as well as the exchange of content will be set individually for each patient in collaboration with the allocated therapist.

- The *Personalised Exercise Programme (PEP)* is theoretically based on the premise that chronic fatigue relates to physical activity (PA) intolerance, supported by unhelpful illness beliefs and deconditioning, thus increased perception of effort. PEP aims to disrupt this cycle with graded exposure to behaviour therapy contingent on symptoms, to gradually optimise patients levels of PA so as to modify altered perceptions of effort, improve tolerance of PA, fitness and function, reverse the deconditioning and ultimately reduce the severity and impact of fatigue. Participants will receive a tailored graded exercise programme, initially delivered according to physical capacity, gradually increasing in duration and intensity. Participants will receive additional information and diaries. The times and duration of keeping the diary as well as the exchange of content will be set individually for each patient in collaboration with the allocated therapist. The intervention will utilise pedometers and/or heart rate monitors for goal-setting and to enhance motivation.

3 Randomisation, Allocation and Blinding

After consent, participants will be randomised via Centre for Healthcare Randomised Trials (CHaRT) based within the University of Aberdeen. The CHaRT provides a 24 h randomisation web-based service. Using a computer-generated sequence, participants will be allocated to one of the two treatments or usual care (1:1:1 ratio).

Randomisation will be minimised by diagnosis (Rheumatic Arthritis [RA], Systemic Lupus Erythematosus [SLE], Axial Spondyloarthritis [AxSpA] or other Inflammatory Rheumatic Disease [IRD]) and the presence/absence of depressive symptoms (Hospital Anxiety & Depression Scale (HADS-D) depression subscale $>10^9$ and will include a random element set at 20%. That is, 20% of all the allocated randomisations will be randomly re-allocated 50:50 to the remaining two treatment options.

Full blinding will not be possible due to the need to engage people in behavioural change. However, we will aim to blind research personnel undertaking outcome assessments to participants' treatment allocation – including the trial statistician with the data being analysed blind to allocation, until the final analyses.

4 Data Monitoring

While there are no planned interim analyses for efficacy or futility, an independent Data Monitoring Committee (DMC) will monitor trial progress and specifically any safety issues. The data available at each DMC will be preserved, along with all documentation of analysis plans, programming code and reporting provided.

For this relatively simple design, the biases should be minimal with the biggest threat being due to data missingness. However, to minimise bias:

- Only the DMC will see any data or analyses for their decisions making, prepared by the trial statistician (blinded –a colleague will re-run the code to reveal the true allocations for each interim report)
- The trial statistician will perform the final analyses, remaining blinded until the final follow-up and data entry has been completed

5 Timing of final Analyses

The final analyses will be performed after the last participants' final follow-up information has been collected and data entered.

6 Timing of Outcome Measurements

The outcome measurements have been planned be taken within a one-week period at defined times (10, 28 and 56 weeks) post randomisation. The actual times will be summarised in the results.

6.1 Primary Outcomes (Specifically at 56 weeks)

- Chalder Fatigue Scale (CF)[16] assessing the physical and mental symptoms of fatigue as a total score using the Likert scale version and not as sub-domains
- Fatigue Severity Scale (FSS)[17] assessing the impact of fatigue.

If the effect of intervention is positive on the CF, then the FSS outcome will be formally analysed. Should the intervention have no effect on the CF, then an explorative analysis of the FSS outcome will be performed.

(prior to 56 weeks these are also monitored and will be included in the final model but are considered as secondary outcomes)

6.2 Other Secondary Outcomes (at all time periods see section 7)

- *Fatigue:* Bristol Rheumatoid Arthritis Fatigue Multi-Dimensional Questionnaire (BRAF-MDQ)[6] assessing physical, living, cognition and emotional aspects of fatigue
- *Quality of life & health utility index:* SF-12[7] assessing functional health and wellbeing from the patient's perspective
- *Pain:* Pain numerical rating scale (10 point) assessing pain intensity[8]
- *Anxiety and depression*: Hospital anxiety and depression scale (HADS)[9]
- *Sleep:* Sleep Problem Scale[10]
- *Impact on work:* Work Productivity and Activity Impairment Questionnaire: Specific Health Problem (WPAI:SHP)[11]
- *Impact on activities:* Valued Life Activities Scale (short form S-VLA)[12]
- *Global outcome*: change of global health[14]

6.3 Additional Demographic and Mediator/moderator variables

- *Demographic:* Age, gender, marital status, employment status, level of education
- *Cognitions and behaviours:* Brief Illness Perception Questionnaire; Behavioural Response to Illness Questionnaire
- *Clinical:* Presence of fibromyalgia; Disease activity (self-reported)
- *Physical:* Physical activity profiles, over a 7 day period; Quantifying aerobic fitness (step) test (weight, VO_2 max and Borg Rating of Perceived Exertion)

6.4 Quantitative evaluation (Qualitative evaluation not covered here)

- Patient preference (only at baseline)
- Patient acceptability (assessed at week 28)

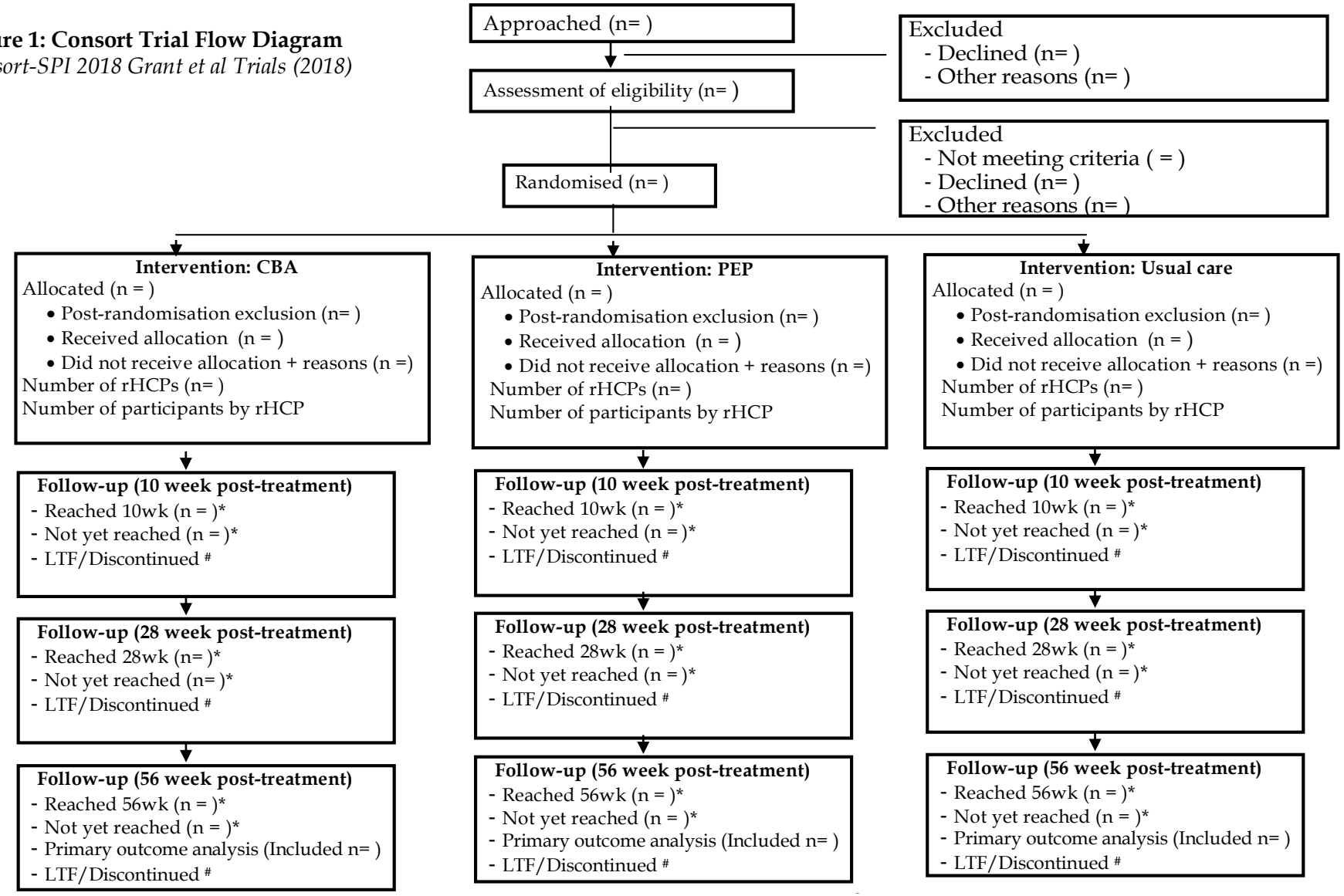
7 Timing of Outcome Measures

	Screening	Proposed assessment [wks]			
		0	10	28	56
Demographic data					
Date of birth, gender, marital status, employment status, level of education		✓			
Characteristics of study population					
Overall health (from domain in SF-12)		✓			
Physical activity (typical self-reported)		✓			
Experience of fatigue for more than 3 months	✓	✓			
Average level of fatigue(self reported- scale 1-10)	✓	✓			
Thyroid function test		✓			
Urea and electrolytes		✓			
Full blood count		✓			
Serological status [S]		✓			
Erosive status		✓			
Disease duration		✓			
Presence of other co-morbidities (Charlson Index) [D]		✓	✓	✓	✓
History of Suicide attempts		✓			
Disease activity DAS28, ASDAS and BILAG for RA, AxSpA and SLE respectively [S]		✓	✓	✓	✓
Inflammation (CRP/ESR)		✓	✓	✓	✓
Previous and current pharmacological therapies		✓	✓	✓	✓
Hypertension / Blood pressure [S]		✓			

Primary Outcome					
Chalder Fatigue Scale (Likert scoring) D		✓	✓	✓	✓
Fatigue Severity Scale (FSS) D		✓	✓	✓	✓
Secondary Outcomes					
BRAF-MDQ (fatigue) D		✓	✓	✓	✓
HADS (anxiety and depression) D		✓	✓	✓	✓
Short Form-12 D		✓	✓	✓	✓
Pain numerical rating scale D		✓	✓	✓	✓
Sleep problem scale		✓	✓	✓	✓
Work Productivity and Activity Impairment Questionnaire D		✓	✓	✓	✓
Valued Life Activities Scale (short 14 items) D		✓	✓	✓	✓
Global outcome D			✓	✓	✓
Additional mediator/moderator data					
Cognitions and behaviours					
Brief Illness Perception Questionnaire D		✓	✓	✓	✓
Behavioural Response to Illness Questionnaire D		✓	✓	✓	✓
Clinical					
Presence of fibromyalgia D		✓			✓
Disease activity (self-reported)		✓	✓	✓	✓
Physical					
Physical activity profiles, over a 7 day period $		✓	✓	✓	✓
Quantifying aerobic fitness (step) test (weight, VO_2 max and Borg Rating of Perceived Exertion) $		✓	✓	✓	✓
Quantitative evaluation					
Patient preference		✓			
Patient adherence (attendance records)		x	x	x	
Patient engagement and adherence (telephone)			x	x	
Patient engagement and adherence (therapist view)			x	x	
Patient acceptability (Client Satisfaction Questionnaire)				✓	

S –Secondary analyses phase
$ PA summarised data to be threaded for the secondary analyses phase
D – derived variables

Figure 1: Consort Trial Flow Diagram
Consort-SPI 2018 Grant et al Trials (2018)



* For monitoring purpose only during the course of the trial - not in final analysis
# the number of participants Lost to Follow-up (LTF) and/or who Discontinued the intervention will also be monitored along with reasons
rHCPs: rheumatology health care professionals CBA: Cognitive Behavioural Approach PEP: Personalised Exercise Programme

8 Trial Population

Patients with rheumatologist diagnosed IRDs (e.g. Rheumatoid Arthritis [RA], Systemic Lupus Erythematosus [SLE] and AxSpA, psoriatic arthritis, vasculitis or Sjogren's Syndrome).

9 Adverse events:

An adverse event (AE) is defined as any untoward medical occurrence in a participant, not necessarily being intervention related. Adverse events are collated according to the protocol (defined by the appropriate SOP). An adverse event is defined as "serious" (SAE) if it

- results in death
- is life threatening
- requires or prolongs inpatient hospitalisation
- results in persistent/significant disability/incapacity
- is otherwise considered medically significant by the investigator.

There are no related serious AEs expected in this trial. However, any serious related AEs that do occur will be recorded following specific Standard Operating Procedure (SOP) for adverse events in non-CTIMP studies. Hospitalisations for elective treatment of a pre-existing condition are not considered as an AE or SAE. Complications occurring during such hospitalisation are also not AEs or SAEs.

10 Sample Size and Power Calculation

Our planned primary Intention-to-Treat analyses (ITT) will compare PEP + usual care versus usual care alone, and CBA + usual care versus usual care alone. This was based on a standardised effect size of 0.50 (considered credible in other pragmatic effectiveness studies), which would equate to being powered to detect a minimal important clinical difference of 2 units in the CF Scale, assuming a common standard deviation across the randomised groups of 4 units, as with PACE [19]. Assuming an overall significance level of 5% (by calculating the two pre-specified randomised groups comparisons, PEP + usual care vs. usual care alone and CBA + usual care vs usual care alone, at 2.5%, to maintain an overall level of not more than 5%) and a power of 90%, we require 100 evaluable participants in each of the three groups.

11 Statistical Methods

11.1 General Methods

All the main analyses will be based on the ITT principle and utilise all available follow-up data from all randomised participants who provide consent. Any post-randomisation exclusions will be removed. Final analysis will take place after full recruitment and follow-up. The results of the trial will follow the guidelines of the CONSORT statement developed specifically for social and psychological intervention trials [3] when presenting and analysing the data. Baseline characteristics of the study population will be summarised separately using the appropriate descriptive statistics and graphical summaries within each randomised group. Baseline characteristics will also be presented for dropouts and completers within each intervention group.

Treatment effects will be tested at the 2-sided 5% significance level with any estimates displayed with 95% confidence intervals (CIs) and p-values. There will be no

adjustment to secondary outcomes CIs for multiple testing. (See section 3 for statistician blinding)

11.2 Statistical Analysis

LIFT has repeated measures on individual participants nested within site suggesting a multilevel model with an appropriate link function depending on the outcome. The analysis will adjust for the outcome variable at baseline as a covariate (when available) as well as the design factors also at baseline [diagnosis (RA, SLE, AxSpA or other IRD), the presence/absence of depressive symptoms HADS depression subscale >10)]. Centre clustering will be accounted for using a random effects robust variance.

11.2.1 Primary Outcome - Effectiveness Analysis.

We will test the primary hypothesis for between-group change in the primary outcome for each of the two pre-specified comparisons (CBA + usual care vs usual care alone and PEP + usual care vs usual care alone) using treatment and its interaction with time fitted as fixed effects, and we apply standard regression diagnostics. The main analysis will focus on the 56 weeks after baseline – providing effect sizes for each of the active arms compared to usual care. Standard regression diagnostics will be applied.

A Complier Average Causal Effect (CACE) analysis will be considered as a sensitivity analysis. Patent engagement & adherence (therapist view) at 8 weeks forms the CACE variable (if missing the 4 weeks reported value will be substituted) as a continuous instrumental variable in the CACE analysis

11.2.2 Secondary outcome Analysis

The secondary outcomes will be analysed using analogous methods to also test for between-group change for each the secondary outcome for the two interventions compared to usual care using treatment, time and treatment/time interaction fitted as fixed effects.

11.3 Mediation and Moderator analyses: (These analyses will be a secondary phase)

If the effectiveness analysis shows significant between group differences on the measures considered as putative mediators (i.e. significant ITT effects when these measures are considered as outcomes), then we will test for mediation of the effect of interventions on primary outcome(s) at 56 weeks through these putative mediators. The analysis will use causal mediation analysis based on parametric regression models (Landau et al, 2013).

This involves estimating a linear model for the mediator with group assignment, baseline CFS (or FSS), baseline mediator, diagnosis and presence/absence of depressive symptoms as covariates, and separately estimating a linear model for CFS (FSS) with the mediator, group assignment, baseline CFS, baseline mediator, diagnosis and presence/absence of depressive symptoms as covariates. The effect of group assignment on the mediator is multiplied by the effect of mediator on CFS (FSS) to estimate the indirect effect, and the effect of interventions on CFS (FSS) in the model including mediator is an estimate of the residual direct effect. The

indirect and direct effects sum to the total effect, and bootstrapping with 1000 replications will be used to obtain valid standard errors for the causal mediation effects. The proportion mediated is the indirect effect divided by the total effect. We will test for moderation of the mediation pathways by primary diagnosis.

Exploratory moderation analyses examined whether the between-group effect on CFS (FSS) was moderated by the following baseline variables: XX. The primary analysis models will be extended by including the moderator, its interaction with group assignment and a three-way interaction with group assignment and time as fixed effects. The difference in between-group effects at each level of the moderator will be calculated using the -margins- command in Stata.

The Moderation and Mediation analyses will be in place of any Subgroup analyses and is planned as a secondary analysis paper.

11.4 Quantitative evaluation Analysis

The main analysis to assess preference on the treatment effect whereby 'no preference' will be considered as being 'not matched' i.e. did not get their preferred treatment.

Two sensitivity analyses will be considered regarding those who 'had no preference'.

- To drop them from the analysis
- To include in the 'matched' group.

Another set of sensitivity analyses will assess the impact of adjusting for 'how positive' participants were about receiving their preferred option, summarised in the table see dummy tables below*. *post randomisation moderator effects such as therapist/HCP effect analyses, patient adherence will be a secondary analysis phase.*

11.5 Missing Outcome Data

The sensitivities of treatment effect estimate to missing outcome data will be explored; these models will explore the robustness of the treatment estimate to whatever small amount of missing data there is. We will follow the strategy outlined in White *et al* (2). The analysis will use all available data that we believe are valid under the assumption of missing at random. The multilevel models used to account for follow-up over time will also internally impute any covariate missingness assuming they are MAR. However, the models only require the outcome variable at baseline as a covariate (when available) as well as the design factors also at baseline and so may be imputed as described above if missingness is substantial. In a trial this is unlikely. Our final estimates at each follow-up will be only for the actual numbers obtained for each of the Outcomes for the primary ITT analyses. If required, that is if the missingness for the primary outcome is >10%, sensitivity analyses will include multiple imputation such as MICE and/or we will explore a range of values for missing data imputed under missing not at random assumptions (such as pattern mixture models); the extent of

missingness will be assessed along with a determination of if the data are MAR or MCAR. In addition, a comparison of baseline characteristics of the responders and non-responders will be conducted with respect to the primary outcome.

11.6 Missing Baseline Data

Data missing at baseline will be reported as such. If required primary and/or secondary outcome data will be imputed with centre specific mean for continuous data and missing binary/categorical data will include a missing indicator, as indicated by current practice[20]

11.7 Missing items for Derived Variables - Patient Reported Outcome Measures (PROMs):

There are a number of PROM trial data collected using validated questionnaires, some of which are combined into an overall score and/or domain scores. These are indicated by [D] in table above in the Timing of Outcome Measures (Section 5). Codes developed in-house are checked and validated by an independent statistician using dummy data. Missingness for amalgamated scores will be treated according to decisions made by the Project Team on 21/06/2019 [See section 14 - Appendix] informed by a review of how others have treated missingness for these derived variables.

11.8 COVID-19

The effect of COVD-19 will be explored. In the first instance, periods before, during and after COVID-19 will be summarised using appropriate descriptive statistics and graphical summaries. If need be, formal analysis will be carried out to explore the effect of COVID-19, that may include time of recruitment in relation to UK lock-down (23rd March 2020) and local conditions at the time each outcome is measured. Attempts will be made to account/adjust for the multiple lockdowns and variations of that around the country using emerging methodologies.

12 Technical Details

Protocol version (vs 11) will be consulted for this SAP. All statistical analyses will use stata (vs 15 for DMC's – and vs 16 for the final analyses). All results will be processed directly into PDF/Word from Stata via LaTEX (MiKTeX 2.9 at time of writing) for the DMC's, the use of putdocx commands in Stata 16 for the final Statistical Report.

13 Dummy Tables

13.1 Descriptive Tables

Table 1: Baseline Demographics (potential moderator variables *)

		measures	CBA N =	PEP N=	Usual Care N =
	age	*Continuous			
	female	(Y) n/N (%)			
	marital status* Single Married Widowed Divorced Separated Living with partner/spouse				
	employment status* Working full-time (30 hrs or more per week) Working part-time (less than 30 hrs per week) Unemployed and looking for work Unable to work because of illness or disability At home and not looking for paid employment Student Retired other	(Y) n/N (%)			
	level of education * Secondary school Apprenticeship Further education college University degree Further degree	(Y) n/N (%)			
	ethnicity * Scottish Other British Irish Other White missing	(Y) n/N (%)			

*Continuous data: n; mean (sd), median (IQR) and (min, max)

Table 2: Baseline population health characteristics (potential moderator variables *)

		measures	CBA N =	PEP N=	Usual Care N =
	Overall Health	*Continuous			
	Fatigue for > 3 months	(Y) n/N (%)			
	Average level of fatigue	Continuous			
	Physical Activity (Typical self-reported)	*Continuous			
	Thyroid function test	*Continuous			
	Urea and electrolytes	*Continuous			
	Full blood count	*Continuous			
	Serological status Rheumatoid Factor positive Anti-cyclic citrullinated protein (CCP) positive Anti-citrullinated protein (ACP) positive Anti-bodyDna Positive Anti-bodyNuclear Positive Anti-Sm Positive Anti-Ro Positive Anti-La Positive HlaB27 Positive Serum complement C3 (g/L) Serum complement C4 (g/L)	(Y) n/N (%) Continuous Continuous			
	Erosive status	(Y) n/N (%)			
	Disease duration Summary >=6 wk	 *Continuous (Y) n/N (%)			

*Continuous data: n; mean (sd), median (IQR) and (min, max)

Table 3: Baseline variable outcome measures

	Measures	CBA N =	PEP N=	Usual Care N =
Primary				
Chalder Fatigue Scale (Likert score) 0-33	*Continuous			
Fatigue Severity Scale (FSS)	*Continuous			
Secondary				
BRAF-MDQ (fatigue) (0-70)	*Continuous			
HADS (anxiety and depression)	*Continuous			
Short Form-12	*Continuous			
Pain numerical rating scale (0-11)	*Continuous			
Sleep problem scale (0-20)	*Continuous			
Work Productivity and Activity Impairment Questionnaire (for all 4 domains)	*Continuous			
Valued Life Activities Scale (short 14 items)	*Continuous			

*Continuous data: n; mean (sd), median (IQR) and (min, max)

Table(s) 4a-c: Variable outcome summaries at follow-up [at weeks a)10, b) 28 and c) 56]

	Measures	CBA N =	PEP N=	Usual Care N =
Primary				
Chalder Fatigue Scale (Combined Likert scores) (0-33)	*Continuous			
Fatigue Severity Scale (FSS)	*Continuous			
Secondary				
BRAF-MDQ (fatigue) (0-70)	*Continuous			
HADS (anxiety and depression)	*Continuous			
Short Form-12	*Continuous			
Pain numerical rating scale (0-11)	*Continuous			
Sleep problem scale (0-20)	*Continuous			
Work Productivity and Activity Impairment Questionnaire (for 4 domains- all)	*Continuous			
Valued Life Activities Scale (short 14 items)	*Continuous			
Global outcome	*Ordinal/ Continuous			

*Continuous data: n; mean (sd), median (IQR) and (min, max)

Table(s) 5a-d: Moderators summaries at a) baseline [time=0] b) 10, c) 28 and d) 56 weeks as appropriate

	measures	time	CBA N =	PEP N=	Usual Care N =
Characteristics					
Overall Health	Categories n/N (%)	all			
Other co-morbidities (Charlson Index)	*Continuous	all			
Disease Activity		all			
DAS28		all			
ASDAS		all			
BILAG for RA		all			
Inflammation CRP ESR	*Continuous	all			
Cognitions and behaviours					
Brief Illness Perception Questionnaire BIPQ (9 items) Item 9 (text) ‡	*Continuous	all			
Behavioural Response to Illness Questionnaire - BRIQ Scale 1 Scale 2 Total	 *Continuous *Continuous *Continuous	all			
Clinical					
Presence of fibromyalgia y/n (And the WPI + SSI Score + TOTAL Score) Disease activity (self-reported, 0-10)	(Y) n/N (%) *Continuous *Continuous	0, 56 all			
Quantitative evaluation					
Patient preference: Option CBA Option PEP Option Usual No preference How positive about receiving this option	(Y) n/N (%) *Continuous	0 0			
Patient adherence (attendance records)	Secondary	0			
Patient engagement & adherence (telephone)	Secondary	0,10,28			
Patent engagement & adherence (therapist view)†	Secondary	10, 56			
Patent acceptability	(Y) n/N (%)	28			
How satisfied with service received? 1:Very satisfied; 2:Mostly satisfied; 3:Indifferent or mildly dissatisfied; 4:Quite dissatisfied; 99:Not answered;					
Come back to this program? 1:No, definitely not; 2:No, I don't think so; 3:Yes, I think so; 4:Yes, definitely; 99:Not answered;					
Get the kind of service wanted? 1:No, definitely; 2:No, not really; 3:Yes, generally; 4:Yes, definitely;					

99:Not answered; To what extent did the program meet needs? 1:Almost all of my needs have been met; 2:Most of my needs have been met; 3:Only a few of my needs have been met; 4:None of my needs have been met; 99:Not answered; Recommend this program to a friend? 1:No, definitely not; 2:No, I don't think so; 3:Yes, I think so; 4:Yes, definitely; 99:Not answered; Satisfied with the amount of help received? 1:Quite dissatisfied; 2:Indifferent or mildly dissatisfied; 3:Mostly satisfied; 4:Very satisfied; 99:Not answered; Have the services received helped to deal more effectively with problems? 1:Yes, they helped a great deal; 2:Yes, they helped; 3:No, they really didn't help; 4:No, they seemed to make things worse; 99:Not answered; How satisfied with the service received overall? 1:Very satisfied; 2:Mostly satisfied; 3:Indifferent or mildly dissatisfied; 4:Quite dissatisfied; 99:Not answered; Would you come back to this program if needed? 1:No, definitely not; 2:No, I don't think so; 3:Yes, I think so; 4:Yes, definitely; 99:Not answered;					

*Continuous data: n; mean (sd), median (IQR) and (min, max)
‡ Summary of Item 1-8 plus overall Score. Also Item 9 indicates causality, BUT will need to be coded and analysed separately (text data) SG : TBC by Stuart Grey
† Patent engagement & adherence (therapist view) at 8 weeks forms the CACE variable (if missing the 4 week) reported value to be substituted

13.2 Serious adverse Events

Table 6: Serious adverse Events

Adverse Events n(%)	**CBA N =**	**PEP N=**	**Usual Care N =**	**Total**
People				
Male				
Female				
AEs				
Type of Adverse Event	**SAE**		**SAE**	
Expected				
Death				

There are SAE's related to any of the interventions expected

13.3 Follow-up timings

Table 7: Summaries of actual follow-up timings

Time period		CBA N =	PEP N=	Usual Care N =	Total
10 weeks	mean (sd)				
	median (IQR)				
	min(max)				
28 weeks	mean (sd)				
	median (IQR)				
	min(max)				
56 weeks	mean (sd)				
	median (IQR)				
	min(max)				

13.4 Primary outcome summaries and model Estimates

Table 8: Primary outcome for Fatigue: Summaries*# and Model results

	CBA N =	PEP N=	Usual Care N =	Effect size[a]	95% CI	p-value	Effect size[b]	95% CI	p-value
Chalder Fatigue Scale (Combined Likert score) 0-33 *[c]									
Baseline									
10wks									
28wks									
56wks (P)									
Fatigue Severity Scale (FSS) *[c]									
Baseline									
10wks									
28wks									
56wks (P)									
Baseline									
10wks									
28wks									
56wks (P)									

All models adjusted for their baseline outcome measure, HADS depression subscale >10 at baseline as fixed effects fixed effect with Centre clustering and individuals nested within centres as random effects

*Continuous data: n; mean (sd), median (IQR) and (min, max)

Binary x/n (%)

[a] Mean difference between CBA and Usual care

[b] Mean difference between PEP and Usual care

P Primary time point

[c] Multilevel mixed-effects generalized linear model (glm) accounting for different time points.

Recall:

- The primary intention to treat analyses will compare PEP + usual care versus usual care alone, and CBA + usual care versus usual care giving effect sizes [a] and [b]
- The main estimate of treatment effect will focus on the 56 weeks after baseline.

- If the effect of intervention is positive on the CF, then the FSS outcome will be formally analysed. Should the intervention have no effect on the CF, then an explorative analysis of the FSS outcome will be performed.
- All analyses and reporting will follow the guidelines of the CONSORT statement developed specifically for social and psychological intervention trials [3]
- CACE analysis as a sensitivity analysis will be considered for the primary outcome - Chalder Fatigue

13.5 Secondary outcome summaries and model Estimates

Table 9: Summaries*# and Model results

	CBA N =	PEP N=	Usual Care N =	Effect size[a]	95% CI	p-value	Effect size[b]	95% CI	p-value
BRAF-MDQ (fatigue) *[c]									
Baseline									
10wks									
28wks									
56wks (P)									
HADS *[c]									
Anxiety									
Baseline									
10wks									
28wks									
56wks (P)									
Depression									
Baseline									
10wks									
28wks									
56wks (P)									
Short Form-12 *[c]									
SF-12 PCS									
Baseline									
10wks									
28wks									
56wks (P)									
SF-12 MCS									
Baseline									
10wks									
28wks									
56wks (P)									
Pain numerical rating scale *[c]									
Baseline									
10wks									
28wks									
56wks (P)									
Sleep problem scale *[c]									
Baseline									
10wks									
28wks									

56wks (p)									
WPAI *[c] for all 4 domains									
Baseline									
10wks									
28wks									
56wks (p)									
Valued Life Activities Scale (short 14 items) *[c]									
Baseline									
10wks									
28wks									
56wks (p)									
Global Outcome [e]									
10wks									
28wks									
56wks (p)									

All models adjusted for their baseline outcome measure where appropriate (not Global outcome), HADS depression subscale >10 at baseline as fixed effects fixed effect with Centre clustering and individuals nested within centres as random effects

*Treated as Continuous data: n; mean (sd), median (IQR) and (min, max)

Binary x/n (%)

[a] Mean difference between CBA and Usual care

[b] Mean difference between PEP and Usual care

[c] Multilevel mixed-effects generalized linear model (glm) accounting for different time points.

[e] Multilevel mixed-effects glm ordinal regression and robust variance –(ref Zou 21004 (5), accounting for time points as interaction terms

[p] Primary time point

13.6 Mediation and Moderation analyses

Table 10: TBC by Prof Richard Emsley as secondary analyses

13.7 Quantitative evaluation

Table 11:

Actual allocation	CBA N =	PEP N=	Usual Care N =	Effect size[a]	95% CI	p-value	Effect size[b]	95% CI	p-value
Patient preference for treatment options #[a]									
CBA									
PEP									
Usual Care									
Patient Acceptability Score*									

Binary x/n (%) those who got their preferred treatment

[a] Mixed-effects glm as a Modified Poisson Regression with log link and robust variance –(ref Zou) adjusted for HADS depression subscale >10 at baseline as fixed effects fixed effect with Centre clustering and individuals nested within centres as random effects

* Additional adjusting variable as a sensitivity analysis

Recall:
The main analysis to assess preference on the treatment effect whereby 'no preference' will be considered as being 'not matched' ie did not get their preferred treatment.

Two sensitivity analyses will be considered regarding those who 'had no preference'.

- To drop them from the analysis
- To include in the 'matched' group.

Another set of sensitivity analyses will assess the impact of adjusting for 'how positive' participants were about receiving their preferred option.

14 References

1. Carpenter J, Kenward M. Missing data in clinical trials — a practical guide. In: Research NIfH, editor. Birmingham, 2008.
2. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. *BMJ* 2011; 342:d40.
3. Sean Grant, Evan Mayo-Wilson, Paul Montgomery, Geraldine Macdonald, Susan Michie, Sally Hopewell, David Moher, on behalf of the CONSORT-SPI Group 'CONSORT-SPI 2018 Explanation and Elaboration: guidance for reporting social and psychological intervention trials' *Trials (2018) 19:406*
4. G Dunn, R. A. Emsley, H Liu, S Landau, J Green, I White, and A Pickles. Evaluation and validation of social and psychological markers in randomised trials of complex interventions in mental health. Health Technology Assessment, 19(93), 2015.
5. Zou G. A Modified Poisson Regression Approach to Prospective Studies with Binary Data. Am J Epidemiol 2004;**159**:702–6. https://doi.org/10.1093/aje/kwh090.
6. J Nicklin, F Cramp, J Kirwan, R Greenwood, M Urban, S Hewlett. Measuring fatigue in rheumatoid arthritis: a cross-sectional study to evaluate the Bristol Rheumatoid Arthritis Fatigue Multi-Dimensional questionnaire, visual analog scales, and numerical rating scales. Arthritis Care Res (Hoboken). 62 (2010) 1559-1568.

7. JE Ware Jr., M Kosinski, SD Keller. A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity, Med.Care. 34 (1996) 220-233.
8. M McCaffery, A Beebe, Pain: Clinical manual for nursing practice, Pain: Clinical Manual for Nursing Practice. (1989) 1-353.
9. AS Zigmond, RP Snaith. The hospital anxiety and depression scale, Acta Psychiatr.Scand. 67 (1983) 361-370.
10. CD Jenkins, B- Stanton, SJ Niemcryk, RM Rose. A scale for the estimation of sleep problems in clinical research, J.Clin.Epidemiol. 41 (1988) 313-321.
11. MC Reilly, AS Zbrozek, EM Dukes. The Validity and Reproducibility of a Work Productivity and Activity Impairment Instrument, Pharmacoeconomics 4 (1993) 353-365.
12. PP Katz, DC Radvanski, D Allen, S Buyske, S Schiff, A Nadkarni, et al. Development and validation of a short form of the valued life activities disability questionnaire for rheumatoid arthritis, Arthritis Care Res. 63 (2011) 1664-1671.
13. E Bro adbent, KJ Petrie, J Main, J Weinman. The Brief Illness Perception Questionnaire, J.Psychosom.Res. 60 (2006) 631-637.
14. Marcus Beasley Gordon J Prescott, Graham Scotland, John McBeth, Karina Lovell, Phil Keeley,et al. Patient-reported improvements in health are maintained 2 years after completing a short course of cognitive behaviour therapy, exercise or both treatments for chronic widespread pain: long-term results from the MUSICIAN randomised controlled trial. RMD Open (2015) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4613171/pdf/rmdopen-2014-000026.pdf
15. Landau et al, 2013 - TBC
16. M Cella, T Chalder. Measuring fatigue in clinical and community settings, J.Psychosom.Res. 69 (2010) 17-22.
17. J Nicklin, F Cramp, J Kirwan, R Greenwood, M Urban, S Hewlett. Measuring fatigue in rheumatoid arthritis: a cross-sectional study to evaluate the Bristol Rheumatoid Arthritis Fatigue Multi-Dimensional questionnaire, visual analog scales, and numerical rating scales. Arthritis Care Res (Hoboken). 62 (2010) 1559-1568.
18. Arthritis Research UK, Fatigue and Arthritis - Self-help and daily living, 2269/FATIG/14-1. (2014) 1-32.
19. PD White, K Goldsmith, AL Johnson, L Potts, R Walwyn, JC Decesare, et al. Comparison of adaptive pacing therapy, cognitive behaviour therapy, graded exercise therapy, and specialist medical care for chronic fatigue syndrome (PACE): A randomised trial, Lancet. 377 (2011) 823-836.
20. Thomas R Sullivan, Ian R White, Amy B Salter et al. Should multiple imputation be the method of choice for handling missing data in randomized trials? Statistical Methods in Medical Research, Vol 27, Issue 9, 2018 https://journals.sagepub.com/doi/10.1177/0962280216683570

15 Appendix – Rules for missing data in derived variables (for LIFT SAP)

Missingness rules (rational and decision)
Decision to adopt established rules where available and to be consistent within the LIFT study for other measures unless the outcome requires a different approach.

Demographic data

Item	Rational	MV Decision
Date of birth, gender, marital status, employment status, level of education	If any one of the baseline demographic items is missing then a basic mean value will be imputed	Baseline mean imputation

Characteristics of study population

Item	Rational	MV Decision
Overall health	Single item	Remains missing
Physical activity (typical self-report)	Single item	Remains missing
Average level of fatigue at screening	Mandatory, i.e. not missing	
Average level of fatigue at baseline	Mandatory, i.e. not missing	
Blood pressure	Mandatory, i.e. should not be missing	Remains missing
Thyroid function test (TSH)	Mandatory, i.e. should not be missing	Remains missing
Urea and electrolytes (eGFR)	Mandatory, i.e. should not be missing	Remains missing
Full blood count (Hb)	Mandatory, i.e. should not be missing	Remains missing
Serological status	Measure depending on completeness and up-to-date-ness of medical notes accessible to RN	Remains missing
Erosive status	Measure depending on completeness and up-to-date-ness of medical notes accessible to RN	Remains missing
Disease duration	Measure depending on completeness and up-to-date-ness of medical notes accessible to RN	Remains missing
History of suicide attempts	Measure depending on response by participant during visit	Remains missing
Charlson Index	Measure depending on completeness and up-to-date-ness of medical notes accessible to RN	Remains missing
Inflammation (CRP)	Minimal missing data as sample taken by RN during visit, except remote visits. Agreement to take blood value from medical notes if sample taken +/- 2 weeks from visit	Remains missing
Inflammation (ESR)	Minimal missing data as sample taken by RN during visit, except remote visits Agreement to take blood value from medical notes if sample taken +/- 2 weeks from visit	Remains missing

Primary Outcome

Item	Rational	MV Decision
Chalder Fatigue Scale	In line with major trials GETSET and Pace, validity	20% rule
Fatigue Severity Scale (FSS)	In line with other primary outcome, no fixed rules	20% rule

20% rule: for each sub scale use person specific mean imputation if <= 20%

Secondary Outcomes

Item	Rational	MV Decision
BRAF-MDQ (fatigue)	Provided by authors of this PROM (Hewlet et el) • Questions 1 and 2 are compulsory. • Only 1 question may be missing from each dimension (maximum of 3 in the overall BRAF-MDQ). Replace the missing question score with the average score for that dimension. • For the Physical Fatigue dimension, a weighted average score is used to account for the varying item score ranges: • Total the 3 completed scores, divide by the total max possible score for those 3 questions, then multiply by the maximum score possible for all 4 questions	See left
HADS (anxiety and depression)	Ad hoc rules, stick with internal 20% -	20% rule
Short Form-12	Long standing validated algorithm available S:/ProgStat/Secure/Statisticians/10Resources /ado/sf12v2	Use code prescribed
Pain numerical rating scale	Only one item NRS	Remains missing
Sleep problem scale	No missing data allowed as only 4 items, aim is to report overall score. Individual items could be used in secondary analysis	If >0%, whole measure missing
Work Productivity and Activity Impairment Questionnaire	Missing data not imputable as per developers	If >0%, whole measure missing
Valued Life Activities Scale	Initially administered in RA via phone True missingness needs assessing - ??	20% rule
Global outcome (change vs visit 1)	Only modelled in Health Economics – will be summerised in SAP	Remains missing

20% rule: for each sub scale use person specific mean imputation if <= 20% items are missing

Additional mediator/moderator data Cognitions and behaviours

Item	Rational	MV Decision
Brief Illness Perception Questionnaire	Each item (1-8) of the Brief IPQ assesses one dimension of illness perceptions Item 9 indicates causality, remains missing BUT will need to be coded and analysed separately (text data)	Remains missing for each question if missing
Behavioural Response to Illness Questionnaire	Follow advise from developers	20% rule

20% rule: for each sub scale use person specific mean imputation if <= 20% items are missing

Clinical

Item	Rational	MV Decision
Presence of fibromyalgia	Minimal missing data as completed by RN Continuous analysis and dichotomous. Made up of 2 subscales combined	Remains missing if any missing
Disease activity (self-reported)	Minimal missing data as completed by RN during visit	Remains missing
Disease activity DAS28 for RA	Minimal missing data as completed by RN during visit, except remote visits	Remains missing
ASDAS	Minimal missing data as completed by RN during visit, except remote visits as not CRP sample taken	Remains missing
BILAG	Completeness depending on medical notes	Remains missing

Quantitative evaluation

Item	Rational	MV Decision
Patient preference	Single item • Missing, set to "no stated preference" • 2 options ticked, set to "no stated preference"	See left
Patient adherence	Data derived from therapist notes if sessions took place (attendance records)	Remains missing
PEA -phone sessions 4 & 8	Single item	Remains missing
PEA - therapist sessions 4 & 8	Single item	Remains missing

PEA: Patient engagement & adherence; PAct: Patient acceptability
20% rule: for each sub scale use person specific mean imputation if <= 20% items are missing