2023-03-29

Mental imagery to reduce alcohol-related harm in patients with alcohol-related liver damage (MIRAGE)

Musicha, C

Pilot randomised trial of functional imagery training plus treatment as usual versus treatment as usual alone to reduce alcohol-related harm in patients with alcohol-related liver disease admitted to hospital

Mental imagery to reduce alcohol-related harm in patients with alcohol-related liver damage (MIRAGE)

Statistical Analysis Plan

ISRCTN: 293042
IRAS PROJECT ID: 293042

CONTENTS

ADMINISTRATIVE INFORMATION

Full title	Pilot randomised trial of functional imagery training (FIT) plus treatment as usual (TAU) versus TAU alone to reduce alcohol-related harm in patients with alcohol-related liver disease admitted to hospital
Short title	Mental imagery to reduce alcohol-related harm in patients with alcohol-related liver damage
Protocol Version	Version 3.1 (03 February 2022)
SAP Version	Version 1.0 (30 March 2023)

SAP Revisions	Version 1.1 (30 March 2023)
	Changes were made to the SAP after database lock and export of the data. Statisticians were not blinded to the allocations at this stage. The following changes were made to the SAP:
	• Added number and proportion of participants who reported not taking alcohol 7 days prior to the assessment day by allocation (page 11) and added Table 11b in Appendix (page 32).
	• Added text to explain use of mean (SD) or median (IQR) distribution of the data and sample size if N is small (Section 4.2 page 14). We also state that we only present summary statistics for those who complete all the visits if necessary
	• Kaplan-Meier survivor curves will be used to visualise the time of relapse to alcohol use by allocated group (Section 4.3 page 14).
	• Corrected MELD score formula in Appendix B (page 36) by replacing 132 with 1.32 in equation 2.
	• Modified Section 5.2: Analysis of pilot outcomes (page 18-19) to add:
	  ○ Number (percentage) of participants re-hospitalised
	  ○ Number (percentage) of participants who relapsed to alcohol use

	Name	Signature	Date
Statistical Analysis Plan authored by:	Senior Trial Statistician: Siobhan Creanor		27/03/23
	Trial Statistician: Joe Lomax		29/03/2023
	Trial Statistician: Crispin Musicha		27/03/2023
	(Kara Stevens, previously UoP)	N/A	

Approved by:	Chief Investigator: Ashwin Dhanda		28/03/2023
	Independent Statistician: Christopher Partlett		28-03-2023

ABBREVIATIONS

ALNs	Alcohol Liaison Nurses
ArLD	Alcohol-related liver disease
AUDIT	Alcohol Use Disorder Identification Test
CI	Confidence Interval
EQ-5D-5L	EuroQol 5 Dimension 5 Level
FIT	Functional Imagery Training
HEAP	Health Economic Analysis Plan
ITT	Intention To Treat
MELD	Model of End-stage Liver Disease
MI	Motivational Interviewing
QoL	Quality of Life
RCT	Randomised Controlled Trial
RUQ	Resource Use Questionnaire
SADQ	Severity of Alcohol Dependence Questionnaire
SAP	Statistical Analysis Plan
SD	Standard Deviation
SWEMWBS	Short Warwick-Edinburgh Mental Wellbeing Scale
TAU	Treatment as Usual
WEMWBS	Warwick-Edinburgh Mental Wellbeing Scale

1 INTRODUCTION

1.1 BACKGROUND AND RATIONALE FOR THE TRIAL

Alcohol-related liver disease (ArLD) is caused by long-term alcohol consumption, usually with physiological and psychological dependence, characterised by liver damage (fibrosis) leading to cirrhosis, which affects patients' quality of life (QoL) and survival[1]. The only effective treatment to prevent progression of liver damage is reducing or ceasing alcohol consumption[2]. In 2019/18 there were over 13,500 emergency hospital admission with a primary diagnosis of ArLD[3]. This is a crisis point in the ArLD patient's journey and research indicates this time point provides an opportunity for intervention where behaviour change is more likely to result.

Treatment as usual (TAU) is a brief intervention, a form of motivational interviewing (MI), conducted by a trained health professional, usually an alcohol liaison nurse, during the in-patient stay, lasting less than 20 minutes and signposting patients to community services, as recommended by NICE[4]. However, early relapse after hospital admission remains a challenge[5]. Functional Imagery Training (FIT) is a new treatment that combines the benefits of MI with evidence-based techniques of mental imagery exercises and training to further strengthen motivation, combat craving, and train self-management skills[6].

As a trial of this new intervention has not been completed before in the target population, this pilot study will be used to gather the necessary operational and feasibility data to inform the design and implementation of a definitive randomised controlled trial (RCT) to assess the effectiveness of FIT in patients with ArLD admitted to hospital.

1.2 PURPOSE OF STATISTICAL ANALYSIS PLAN

The study protocol includes an outline of the statistical methods to be employed in the analysis of the trial data. The purpose of the Statistical Analysis Plan (SAP) is to provide full details of the planned statistical methods to be used in the primary report of the trial results. The SAP has been drafted following the SAP Guidelines[7], CONSORT extension for Pilot and Feasibility Studies[8] and also taking cognisance of the CONSORT extensions for reporting patient-reported outcomes[9] and non-pharmacologic treatment interventions[10]. However, it is worth noting that as this is a pilot trial, formal/inferential statistical analysis and hypothesis testing of the outcome measures are not appropriate and thus will not be undertaken. There is a separate health economic analysis plan (HEAP).

2 TRIAL OBJECTIVES

2.1 PRIMARY OBJECTIVES

To conduct a randomised pilot trial of FIT and TAU versus TAU alone. This pilot study will provide high quality data:

1. To estimate rates of screening, recruitment, randomisation, retention, adherence to FIT/TAU and possible contamination
2. To allow a preliminary assessment of FIT intervention in the ArLD population.

2.2 Secondary Objectives

1. To estimate the resource use and costs associated with delivery of intervention, and to pilot methods for the cost-effectiveness framework in a full trial
2. To identify if there is a need to improve FIT training and delivery by Alcohol Liaison Nurses (ALNs) within the NHS and if so, methods for improvement.

3 Trial Design

3.1 General Design

Multicentre randomised controlled pilot trial of FIT+TAU (intervention group) vs TAU alone (control group).

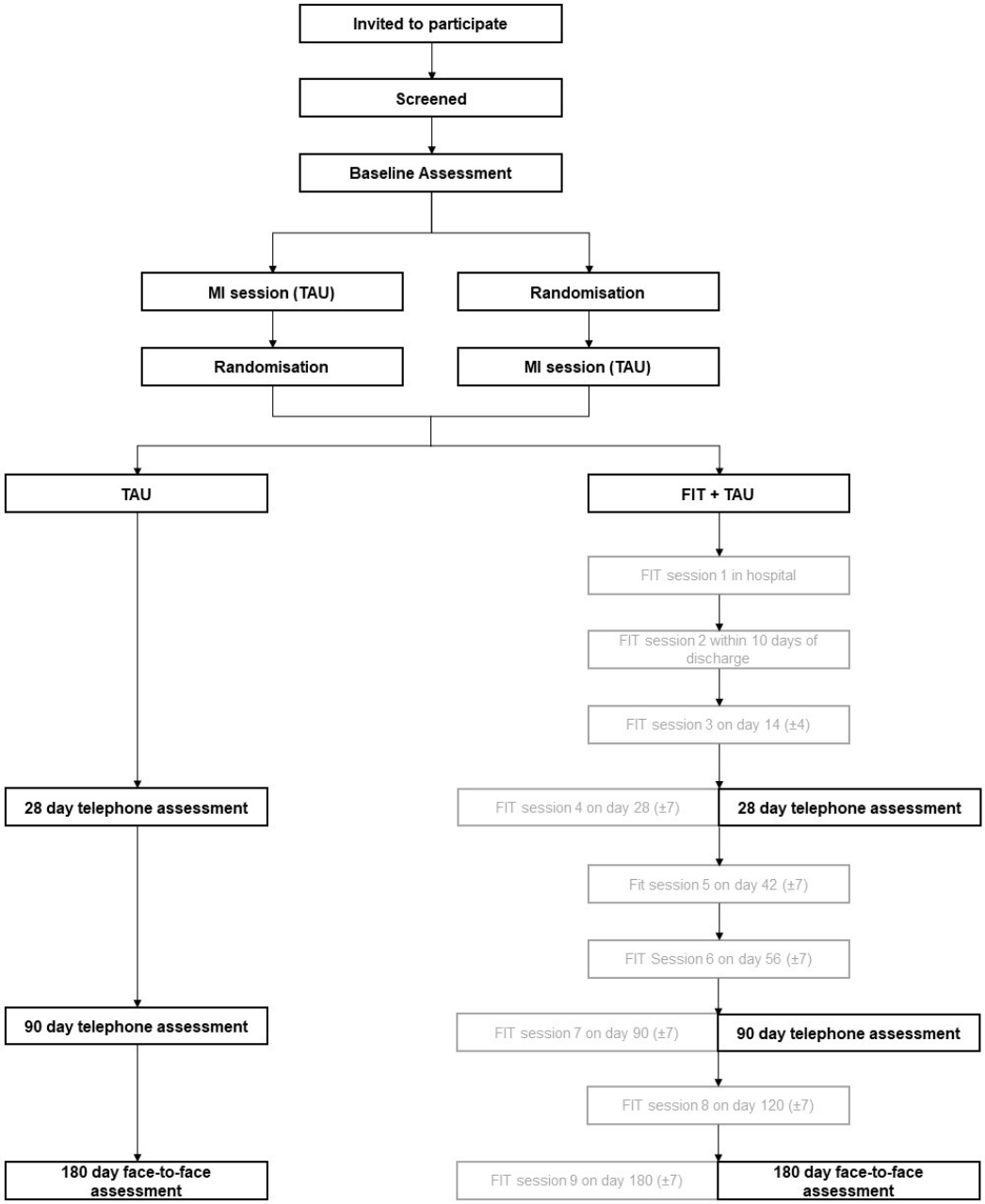

Figure 1: Flow of participants' progression through the MIRAGE pilot trial.

TAU comprises of one brief MI-based session given in hospital by the ALN. All participants should receive TAU. Due to local hospital practices, participants may have received TAU prior to being approached about this study, prior to giving informed consent, or prior to completing the baseline measures and being randomised. The FIT intervention comprises of one session given face-to-face to participants before discharge from hospital, a second session given, if possible, face-to-face to patients in an outpatient clinic or via telephone and then a further seven sessions delivered by telephone over a period of 6 months.

All participants will be scheduled for follow-up at 28 (±7), 90 (±7) and 180 (±14) days post-baseline. Figure *1* shows participants' progression through the study and time points they receive the FIT intervention (if allocated to TAU + FIT) and their follow-up assessments.

3.2 Blinding

This trial is non-blinded to ALNs and participants, as it is not possible to conceal the active psychological FIT intervention from them. The outcome assessors (i.e. research team members conducting research visits) will be blinded to treatment allocation. The trial statistician undertaking the analyses will not be blinded to the treatment allocations. This SAP will be finalised prior to the end of the recruitment period, limiting any potential risk arising from the statistician not being blinded.

3.3 Inclusion and Exclusion Criteria

3.3.1 Inclusion Criteria

Patients must satisfy all of the following criteria to be enrolled in the study:

- Adult patients ≥18 years
- Able and willing to provide informed consent
- Clinical diagnosis of ArLD by at least one of the following methods
    - radiological appearance of fatty infiltration of the liver or cirrhosis
    - histological findings of cirrhosis or alcoholic steatohepatitis
    - signs consistent with chronic liver disease on physical examination
- High risk alcohol consumption (>50 units/week for males and >35 units/week for females) within 4 weeks prior to hospital admission
- Alcohol Use Disorder Identification Test (AUDIT) score[11] >15 during current hospital admission
- Diagnosis of alcohol dependence documented by clinician in medical records. This should be with reference to the ICD-10 meeting at least three of the following conditions:
    - strong desire or sense of compulsion to take alcohol
    - difficulties in controlling alcohol-consuming behaviour in terms of its onset, termination, or levels of use
    - a physiological withdrawal state when alcohol use has ceased or been reduced, as evidenced by: the characteristic withdrawal syndrome; or use of alcohol with the intention of relieving or avoiding withdrawal symptoms
    - evidence of tolerance, such that increased doses of alcohol are required in order to achieve effects originally produced by lower doses
    - progressive neglect of alternative pleasures or interests because of alcohol use, increased amount of time necessary to obtain or consume alcohol or to recover from its effects
    - persisting with alcohol use despite clear evidence of overtly harmful consequences.

3.3.2 Exclusion Criteria

Patients who meet any of the following criteria will be excluded from study participation:

- Any condition with an estimated life expectancy of less than 6 months
- Patients participating in concurrent interventional research
- Participants who have significant difficulties in adequate understanding of English such that they are unable to benefit from the trial intervention or sufficiently understand the trial documentation
- Prisoners
- Patients who do not have access to a telephone so would be unable to participant in FIT sessions.

3.4 Outcome Measures

3.4.1 Study Population

Information to describe the study population will include:

- Number of eligible participants
- Number willing/consenting participants
- Number of participants who withdraw/ are lost to follow-up
- Reasons for decline, ineligibility, withdrawal where available.

Collected baseline data will include:

- demographic information including gender, age, ethnicity, place of residence, living arrangements, marital status, age at completing full time education, employment status, and occupational status
- Liver disease data including
  - date of diagnosis,
  - stage of liver disease (e.g. fatty liver only, fibrosis, cirrhosis)
  - Model of End-stage Liver Disease (MELD) score[12]
  - Child's Pugh score[13]
  - known co-factors (e.g. Hepatitis B or C, non-alcoholic liver disease, haemochromatosis, Alpha-1 antitrypsin deficiency, autoimmune liver disease)
- Alcohol use data including:
  - AUDIT score
  - Severity of Alcohol Dependence Questionnaire (SADQ)[14]
- Participant characteristics including:
  - Substance misuse history including type, frequency and whether currently being used.
  - Comorbidities including type, year of diagnosis and whether this is ongoing
  - Surgical history including type, year and whether this is ongoing
  - Mental health history including type, year and whether this is ongoing
  - Medications including type, dose and frequency

3.4.2 Pilot Trial Outcome Measures

To facilitate the design and planning of a future definitive trial, we will gather the following outcome measures:

- Recruitment rate (overall and by site)
- Retention rate at 90 and 180 days (overall and by site)
- Intervention engagement – number/proportion of scheduled FIT sessions attended
- Duration of each FIT session
- Number/proportion of completed FIT sessions prior to each follow-up assessment
- Number/proportion of instances where:
    - 28 day follow-up occurred before FIT session 4
    - 90 day follow-up occurred before FIT session 7
    - 180 day follow-up occurred before FIT session 9
- Completeness of potential outcome measurements for definitive trial
- Number/proportion of instances the outcome assessor was un-blinded at each follow-up visit.

3.4.3 Participant-Reported and Clinical Outcome Measures

Outcome Measure	Baseline	28 days post-baseline	90 days post-baseline	180 days post-baseline
Alcohol use	X	X	X	X
SADQ	X	X	X	X
WEMWBS	X	X	X	X
SWEMWBS	X	X	X	X
Re-hospitalisation		X	X	X
Alcohol relapse		X	X	X
Alcohol metabolites				X
EQ-5D-5L*	X	X	X	X
Resource Use Questionnaire*	X		X	X

* please see separate HEAP

Table 1: Table detailing the timing of participant reported and other clinical outcomes

The participant-reported and clinical outcome measures listed in Table 1 will be collected in the pilot trial to help inform appropriate primary and secondary outcome measures for the definitive trial and to allow a preliminary assessment of FIT intervention in the ArLD population. The proposed primary outcome for a future definitive trial is change in self-reported alcohol use (grams of pure alcohol/week) between baseline and 180 (±14) days post-baseline. The Timeline Follow-Back technique[15] will be used to determine a participant's alcohol use during the 7 days immediately prior to: their hospital admission (baseline) and 28, 90 and 180 days post-baseline. The completeness of the Timeline Follow-Back data will also be assessed. We will report, for each visit after baseline, the number and proportion of participants who reported not taking alcohol 7 days prior to the assessment day by allocation.

Proposed secondary outcomes for a future definitive trial are the:

- *Severity of Alcohol Dependence Questionnaire (SADQ)*[14] – This questionnaire consists of 20 items and covers the following areas of the dependency syndrome:
    - physical withdrawal symptoms; affective withdrawal symptoms
    - relief drinking
    - frequency of alcohol consumption
    - speed of onset of withdrawal symptoms.

  Each item of the questionnaire is scored between 0 (almost never) and 4 (nearly always)
- *Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS)*[17] – This questionnaire is a measurement of mental wellbeing, consisting of 14 items each with a score between 1 (None of the time) and 5 (all of the time)
- *Short Warwick-Edinburgh Mental Wellbeing Scale (SWEMWBS)*[18] - This questionnaire is the short version of WEMBWS, consisting of 7 of the items from the full WEMWBS
- *Re-hospitalisation within 180 days post baseline* – Self-reported by participants or determined using hospital records at participating sites
- *Time to relapse* – Self-reported time to relapse to regular alcohol use, defined as ≥5 drinking days per week, or ≥5 units in a single day
- *Alcohol metabolites* – a urinary sample will be used to measure ethyl glucuronide (µg/L) and ethyl sulphate (µg/L), which provide a highly sensitive and specific objective quantitative measure of alcohol consumption within the preceding 72 hours
- *EuroQol 5 Dimension 5 Level (EQ-5D-5L)*[16] – This quality of life questionnaire, commonly used in health economics, consists of five questions, where each question/item can be scored between 1 (no problems) and 5 (extreme problems). Further information is provided in the HEAP
- *Resource Use Questionnaire (RUQ)* – see the separate HEAP for details.

3.5 Sample Size Calculation

One of the key objectives of this pilot study, to inform the definitive trial, is whether patients can be successfully recruited and followed-up. We estimate that across all sites, 24 potentially eligible ArLD patients are admitted per month. Our total recruitment target is 90 participants.

This recruitment target will allow estimation of the overall retention rate with a 95% confidence interval (CI) with precision of at least ±11%. Assuming a non-differential retention rate of 75% at the 6-month follow-up (the anticipated primary endpoint for a definitive trial), indicates primary outcome data will be available from a minimum of 33 participants within each allocated group. This will enable appropriate estimation of key components, such as the variability in the proposed primary outcome, to inform the sample size calculations for the definitive trial.

4 STATISTICAL PRINCIPLES

4.1 RANDOMISATION, ALLOCATION CONCEALMENT AND STRATIFICATION

Participants will be allocated to receive TAU or TAU+FIT, in a 1:1 ratio, using random permuted blocks, stratified by recruiting site and the participant's baseline SADQ total score, dichotomised as ≤30 (moderate) or >30 (severe)[19].

4.2 ANALYSIS POPULATIONS

Primary analysis (in the form of summary statistics such as means and standard deviations (SDs) or medians and interquartile ranges (IQR), where appropriate), will be undertaken on an Intention To Treat (ITT) basis, where participants are analysed according to their allocated group, regardless of adherence to the protocol or lack of participation or completion of FIT sessions if allocated to the intervention group. No formal (i.e. no hypothesis testing/inferential) analysis will be performed. Missing responses will be imputed according to the guidance set out in section 4.8.1. Any participants with non-imputable items/scores, will not be included in the ITT analysis for the outcome/time-point under consideration. We will present summary statistics of outcomes for participants who complete all visits in both groups. the changes between baseline and each of the 28, 90, and 180 day assessments, a participant will need to have non-missing or imputable data at both of the time points to be included in the analysis. Kaplan-Meier survivor curves will be used to characterise the time of relapse to alcohol use by group.

The safety population will include all participants who consent to participate in the study, with safety data collected from the time of recruitment until a participant completes or withdraws from the study. Any participant who does not attend at least one FIT session will be analysed as TAU and any participant who has received at least one FIT session will be analysed as TAU + FIT.

4.3 STATISTICAL SIGNIFICANCE LEVELS

As this is a pilot trial, no inferential between-group comparisons will be undertaken (i.e. there will be no between-group hypothesis testing). Feasibility outcomes, such as recruitment rates, will be presented with two-sided 95% confidence intervals. Between-group differences for proposed trial outcomes will be summarised descriptively and presented with two-sided 75%, 85% and 95% confidence intervals, as recommended by Lee et al (2014)[23].

Estimates that may be used for future sample size calculations (e.g. standard deviation of proposed primary outcome, correlation between baseline and follow-up measure) will be presented with two-sided 80% and 90% confidence intervals (Browne (1995)).

4.4 INTERIM ANALYSIS

There is no planned interim analysis for this pilot trial. If, for any reason, the Trial Steering Committee (TSC) requests an interim analysis of the data, efforts will be made to appoint a statistician independent of the trial team to undertake such work, in order to retain the blinding of the trial statistician.

4.5 Time Points of Statistical Analysis

Statistical analysis will be undertaken once the final group of participants has completed the final assessment at 180 (±14) days post-randomisation and the database is locked.

4.6 Derived Variables

The following variables will require coding to calculate scores for analysis:

- *MELD Score* – Specifications of how to calculate the MELD score are in detailed in **Error! Reference source not found.**. N.B. the maximum MELD score is 40
- *Child Pugh Score* – This score consists of five items each with a score of 1 to 3. The Child Pugh score is the total of all five items, giving a total score of between 5 and 15 points. Further details of how each item is scored is detailed in **Error! Reference source not found.**
- *SADQ Score* – The score is the sum of all 20 items of the SADQ and can be between 0 and 80. A score of:
    - <16 usually indicates only a mild physical dependency
    - 16 – 30 indicates moderate dependence
    - >30 indicates severe alcohol dependence[19]
- *WEMWBS* – The score is sum of all the items from the questionnaire and can be any value from 14 to 70[20]
- *SWEMWBS* – The raw score (sum of all 7 items of the SWEMBS) needs to be converted to the metric score using the conversion table from Stewart-Brown, et al.[21] in **Error! Reference source not found.** (Table 21).

4.7 Missing Data

One of the objectives of this pilot trial is to assess the completeness of potential outcome measures for the definitive trial, at the level of both item and outcome measure. Missing outcome data will be noted and used to inform the likely pattern of missing data in a full-scale trial. If a considerable amount of outcome data is missing, this may suggest a need to reconsider the choice of outcome measures and in particular inform the choice of primary outcome measure for any future definitive trial. This may also provide an insight into how missing data can be minimised in any subsequent full-scale trial.

4.7.1 Imputation

Imputation will only be performed for missing questionnaire items when a small proportion of items are missing from an individual response. Where possible, validated, published guidelines will be followed. Any participant who misses a follow-up visit or the whole outcome measure/questionnaire, will not be imputed for that time-point or measurement. Any missing clinical measures, such as the Child Pugh score and MELD score, will not be imputed.

The published guidelines for imputation of the proposed outcomes are summarised below:

- WEMWBS – Impute missing items if no more than 3 are missing. The guidance suggests using the mean of all responses for the missing item or the mean of the non-missing items for the participant as possible imputation methods[24].

- SWEMWBS – The transformations set out in Table 17 are only valid when the response is complete, i.e. no missing values, therefore no imputation will be performed for this measure[24].

5 STATISTICAL ANALYSES

As the study is a pilot trial, it is not suitably powered to be able to support or justify any conclusions regarding treatment efficacy/effectiveness realised from hypothesis testing, and indeed that is not the purpose of the study. As such, the analysis of the results of this trial will not involve formal, inferential statistical comparisons or hypothesis testing between groups. Analyses will be descriptive with the view to informing the design of a fully powered MIRAGE randomised controlled trial.

Continuous measures will be summarised as means, standard deviations and ranges where the distribution appears approximately normally distributed, and as medians, inter-quartile ranges and ranges otherwise. Categorical data will be summarised by frequencies and percentages. Where appropriate, parameter estimates (e.g. between-group differences) will be presented with confidence intervals.

Analysis of quantitative data will be conducted to summarise pilot outcomes, evaluate acceptability of and concordance with the FIT intervention, and the completion and summary statistics of the planned primary and secondary patient-reported and clinical outcomes measures. In addition, appropriate plots will be used to illustrate key data and assess potential between-group differences.

5.1 STUDY POPULATION

Data from the screening process through to the completion of the trial will be recorded and presented in a CONSORT-style flow diagram (see Figure 2). In particular, the following data will be provided:

- Number of people invited to participate
- Number of people screened for eligibility
- Number (percentage) of people ineligible
- Number (percentage) of people eligible and asked to participate
- Number (percentage) of people who declined to participate
- Number (percentage) of people consented to participate
- Number (percentage) of participants who consented to participate but did not proceed to randomisation
- Number (percentage) of participants randomised to each allocated group
- Number (percentage) of participants who did not receive their allocated treatment
- Number (percentage) of participants who did receive their allocated treatment
- Number (percentage) of participants who completed the 28 (±7) days post baseline assessment
- Number (percentage) of participants who completed the 90 (±7) days post baseline assessment
- Number (percentage) of participants who completed the 180 (±14) days post baseline assessment
- Number (percentage) of participants lost to follow up
- Number (percentage) of participants analysed

5.1.1 Baseline Characteristics and Demographics

Baseline characteristics, collected prior to randomisation, will be summarised by allocated group to informally check for balance between groups and provide an overview of the study sample (see **Error! Reference source not found.**). No inferential analysis baseline data by of allocated groups will be undertaken[22], but any considerable imbalance will be noted to inform the design of the full trial.

5.1.2 Participants who Discontinue, Withdraw or are Lost to Follow-up

It is possible that participants will withdraw consent part way through the trial, or their treatment may be discontinued due to medical reasons. It is unlikely in this trial that a participant will be discontinued on medical grounds (for either allocated group), but for reasons such as injury, some participants may not be able to complete the trial. Participants who discontinue will be categorised as follows:

- Continue to consent for follow-up and data collection
- Consent to use pre-collected data only
- Complete withdrawal of consent to use any data

Reasons for withdrawal or loss to follow up will be summarised where reported, at each stage of the process (**Error! Reference source not found.**), including withdrawal prior to randomisation, participants who did not receive their allocated treatment, non-completion of treatment, lost to follow-up.

Participants who withdraw from the study, or whose treatment is discontinued on medical grounds, will not be replaced although their available data will be used unless they have specifically requested for it to be removed from the database. The extent of discontinuation, withdrawal and loss to follow up will be used to inform the design of the fully powered subsequent study, predominantly to ensure a sufficiently powered trial after allowing for losses to follow-up.

5.2 Analysis of Pilot Outcome Measures

In addition to the summary statistics detailed in section 5.1, the following data will be summarised:

- Time to recruit total study sample
- Number (percentage) of participants who completed TAU before recruitment or after recruitment but before baseline assessment, and number (percentage) of participants who completed TAU post-baseline assessment, in total and by site
- Time elapsed, by allocated group, of:
  - baseline assessment to randomisation
  - baseline to 28 (±7) day assessment
  - baseline to 90 (±7) day assessment
  - baseline to 180 (±14) day assessment
- Number of assessments completed within the 28, 90 and 180-day assessment windows
- Acceptability of and adherence of FIT programme:
  - Number (percentage) of participants in the intervention group who completed:
    - all FIT sessions
    - ⩾2 FIT sessions
    - FIT session 1 along with at least one of FIT session 2 or FIT session 3
    - each of the nine FIT sessions
- Number (percentage) of participants re-hospitalised

- Number (percentage) of participants who relapsed to alcohol use.
- Time elapsed between randomisation and start of the FIT sessions
- Number (percentage) of FIT sessions delivered within the specified time frame (see Figure 1 for specified time frames)
- Number (percentage) of participants who attended at least four FIT sessions before the 28 day assessment
- Number (percentage) of participants who attended at least seven FIT sessions before the 90 day assessment
- Number (percentage) of participants who attended at least nine FIT sessions before the 180 day assessment
- Time elapsed between the last FIT session participant attended and the 180 day assessment
- Number who attended final fit session after 180 day assessment
- Feasibility of using the proposed patient- and clinician-reported outcome measures. For each outcome measure we will report (by allocated group and overall):
    - Number (percentage) of participants who completed all items of the outcome measure (if relevant)
    - Number (percentage) of participants with a valid total score for each outcome measure, allowing for appropriate imputation methods if available.

5.3 Preliminary Assessment of FIT

As this is not a fully powered trial, it is inappropriate to perform any formal hypothesis tests comparing the allocated groups. However, to provide a preliminary assessment of whether FIT is a potentially useful treatment for patients with ArLD, we will calculate summary statistics for each of the participant-reported and clinical outcomes listed in section 0, by allocated group.

For continuous outcomes, we will produce boxplots and calculate the mean and corresponding confidence interval by allocated group for the measure at 28, 90 and 180 days and the change between baseline and 28, 90 and 180 days.

The simple, unadjusted estimated between-group difference of the change between baseline and 28, 90 and 180 days for each continuous outcome will be presented with the corresponding confidence intervals. The between-group differences, and corresponding 95% confidence intervals, adjusted for stratification variables will also be calculated using multivariable linear regression. The model will include both of the stratification variables (dichotomised baseline SADQ score and site).

For binary outcomes, we will produce calculations of the frequency and percentage (with corresponding exact 95% confidence interval) by allocated group. The unadjusted and adjusted between-group differences will be reported, derived from logistic regression models, with dichotomised baseline SADQ score and site included in the adjusted model.

5.4 Safety Data

The likelihood of participants being harmed by either the FIT intervention or any of the trial procedures is very low. As such, the reporting of adverse events in the MIRAGE trial is restricted to only those events that are serious, which

- results in death
- Is life-threatening
- Requires inpatient hospitalisation or prolongation of existing hospitalisation
- Results in a persistent or significant disability/incapacity
- Is a significant or important medical event

All serious adverse events reported after the participant has consented to take part in the trial will be reported by treatment group, whereby any participant who has received at least one FIT session will be included in the TAU + FIT treatment group.

5.5 Progression to Definitive Trial

The progression criteria in Table 2 are proposed for consideration as part of the decision on whether or not to progress to planning a definitive trial. If the study delivers results on all of the progression criteria in the green column, it is intended that this will result in the planning of a definitive trial. If some criteria are in the amber column, modification of the trial design will be required before considering progression. If some criteria fall in the red column, the trial team will consider all options, including not progressing to a definitive trial.

Criteria	Red	Amber	Green
Percentage recruited from patients approached by research team (for discussion of informed consent)	<40	40 – 60	>60
Percentage of intervention participants completing FIT session 1 **and** at least one of FIT session 2 or FIT session 3	<50	50 – 70	>70
Percentage of all participants followed-up at proposed primary endpoint of 180 days	<60	60 – 80	>80
Percentage of all participants providing valid data for the proposed primary outcome of alcohol use at proposed primary endpoint of 180 days	<55	55 – 75	>75

Table 2: Proposed progression criteria for planning a definitive trial

5.6 Sample Size for Definitive Trial

One of the key purposes of a pilot study is to obtain data to inform the sample size calculation for a definitive trial. To assist with future sample size calculations, we will calculate, overall and by allocated group:

- the standard deviation for the proposed primary outcome of alcohol consumption (grams consumed of pure alcohol in the previous week)
- the correlation coefficient between baseline and (i) 180 day assessment and (ii) change between baseline and 180 day assessment for the proposed primary outcome of grams consumed of pure alcohol in the previous week

Point estimates will be presented alongside two-sided 80% and 90% confidence intervals.

Indicative sample sizes for a definitive trial with alcohol use at 180 days as the primary outcome will then be produced, conservatively using the upper bounds from the confidence intervals for the estimated standard deviation. Additional sample size scenarios will consider adjustment for baseline alcohol use (i.e. using an analysis of covariance approach) and will use the lower bound from the confidence intervals for the correlation coefficient.

5.7 Statistical Software

The statistical analyses will be undertaken using StataSE version 16 or later, supplemented where required by R.

6 REFERENCES

1. Foster JH, Powell JE, Marshall EJ, et al. Quality of life in alcohol-dependent subjects - a review. *Qual Life Res* 1999;8(3):255-61. doi: 10.1023/a:1008802711478

2. European Association for the Study of the Liver. Electronic address eee, European Association for the Study of the L. EASL Clinical Practice Guidelines: Management of alcohol-related liver disease. *J Hepatol* 2018;69(1):154-81. doi: 10.1016/j.jhep.2018.03.018 [published Online First: 2018/04/10]

3. CCG. 1.8 Emergency admission for alcohol related liver disease. June 2020. NHS Digital, 2020.

4. National Collaborating Centre for Mental Health. Alcohol use disorders. *The NICE guideline on the diagnosis, assessment and management of harmful drinking and alcohol dependence* 2011

5. Thursz M, Forrest E, Roderick P, et al. The clinical effectiveness and cost-effectiveness of STeroids Or Pentoxifylline for Alcoholic Hepatitis (STOPAH): a 2 x 2 factorial randomised controlled trial. *Health Technol Assess* 2015;19(102):1-104. doi: 10.3310/hta191020

6. Kavanagh DJ, Andrade J, May J. Imaginary relish and exquisite torture: the elaborated intrusion theory of desire. *Psychol Rev* 2005;112(2):446-67. doi: 10.1037/0033-295X.112.2.446

7. Gamble C, Krishan A, Stocken D, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. *JAMA* 2017;318(23):2337-43. doi: 10.1001/jama.2017.18556

8. Eldridge SM, Chan CL, Campbell MJ, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *BMJ* 2016;355 doi: 10.1136/bmj.i5239

9. Calvert M, Blazeby J, Altman DG, et al. Reporting of Patient-Reported Outcomes in Randomized Trials The CONSORT PRO Extension. *JAMA* 2013;309(8):814-22. doi: 10.1001/jama.2013.879

10. Boutron I, Altman DG, Moher D, et al. CONSORT Statement for Randomized Trials of Nonpharmacologic Treatments: A 2017 Update and a CONSORT Extension for Nonpharmacologic Trial Abstracts. *Ann Intern Med* 2017;167(1):40-47. doi: 10.7326/M17-0046

11. Saunders JB, Aasland OG, Babor TF, et al. Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption - II. *Addiction* 1993;88(6):791-804.

12. Malinchoc M, Kamath PS, Gordon FD, et al. A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts. *Hepatology* 2000;31(4):864-71. doi: 10.1053/he.2000.5852 [published Online First: 2000/03/25]

13. Pugh RN, Murray-Lyon IM, Dawson JL, et al. Transection of the oesophagus for bleeding oesophageal varices. *Br J Surg* 1973;60(8):646-9. doi: 10.1002/bjs.1800600817 [published Online First: 1973/08/01]

14. Stockwell T, Hodgson R, Edwards G, et al. The development of a questionnaire to measure severity of alcohol dependence. *Br J Addict Alcohol Other Drugs* 1979;74(1):79-87. doi: 10.1111/j.1360-0443.1979.tb02415.x [published Online First: 1979/03/01]

15. Sobell LC, Sobell MB. Timeline Follow-Back. In: Litten RZ, Allen JP, eds. Measuring Alcohol Consumption: Psychosocial and Biochemical Methods. Totowa, NJ: Humana Press 1992:41-72.

16. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;20(10):1727-36. doi: 10.1007/s11136-011-9903-x [published Online First: 2011/04/12]

17. Tennant R, Hiller L, Fishwick R, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. *Health Qual Life Outcomes* 2007;5:63. doi: 10.1186/1477-7525-5-63 [published Online First: 2007/11/29]

18. Shah N, Cader M, Andrews WP, et al. Responsiveness of the Short Warwick Edinburgh Mental Well-Being Scale (SWEMWBS): evaluation a clinical sample. *Health Qual Life Outcomes* 2018;16(1):239. doi: 10.1186/s12955-018-1060-2 [published Online First: 2018/12/24]

19. Stockwell T, Murphy D, Hodgson R. The severity of alcohol dependence questionnaire: its use, reliability and validity. *Br J Addict* 1983;78(2):145-55. doi: 10.1111/j.1360-0443.1983.tb05502.x [published Online First: 1983/06/01]

20. WEMWBS Resource. Collect, score, analyse and interpret WEMWBS Warwick Medical School: University of Warwick; 2021 [updated 06/07/2021. Available from: https://warwick.ac.uk/fac/sci/med/research/platform/wemwbs/using/howto/ accessed 27/07/2021 2021.

21. Stewart-Brown S, Tennant A, Tennant R, et al. Internal construct validity of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS): a Rasch analysis using data from the Scottish Health Education Population Survey. *Health and Quality of Life Outcomes* 2009;7(1):15. doi: 10.1186/1477-7525-7-15

22. Senn S. Testing for baseline balance in clinical trials. *Stat Med* 1994;13(17):1715-26. doi: 10.1002/sim.4780131703 [published Online First: 1994/09/15]

23. Lee, E.C., Whitehead, A.L., Jacques, R.M. *et al.* The statistical interpretation of pilot trials: should significance thresholds be reconsidered?. *BMC Med Res Methodol* **14,** 41 (2014) doi:10.1186/1471-2288-14-41

24. Taggart, F., Stewart-Brown, S., & Parkinson, J. (2015). Warwick-Edinburgh Mental Well-being Scale (WEMWBS) User Guide, Version 2. NHS Health Scotland

25 Browne RH. On the use of a pilot sample for sample size determination. Stat Med. 1995 Sep 15;14(17):1933-40. doi: 10.1002/sim.4780141709. PMID: 8532986.

APPENDIX A: PLOTS AND TABLES

Examples of potential tables and plots to be used in the statistical analysis report.

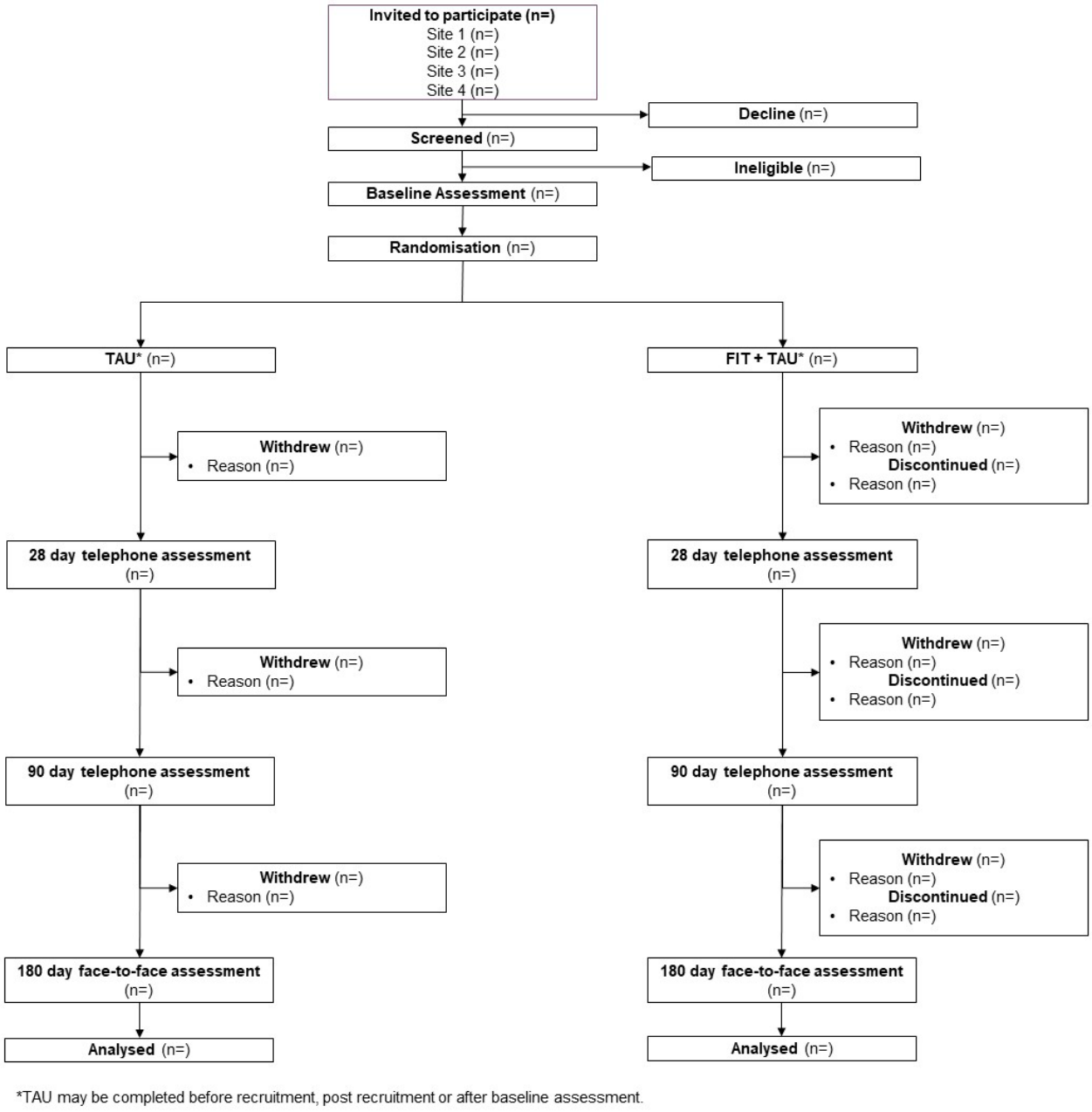

Figure 2: CONSORT diagram of participant flow through the MIRAGE Pilot Trial

Table 3: Number of participants who completed TAU before recruitment, post recruitment or after baseline assessment

Site	Before recruitment N (%)	Post recruitment but before baseline assessment N (%)	Post baseline assessment N (%)
1			
2			
3			
4			
Total			

Table 4: Number of assessments completed within the scheduled 28, 90, and 180-day assessment windows

Assessments Completed	TAU N (%)	FIT + TAU N (%)	ALL N (%)
28-day window (±7 days)			
90-day window (±7 days)			
180-day window (±14 days)			

Table 5: Time differences between key events

Time Elapsed (days) Between	TAU Mean (SD) [range]	FIT + TAU Mean (SD) [range]	ALL MEAN (SD) [range]
Baseline and randomisation			
Randomisation and start of the FIT sessions*	N/A		
Randomisation and 28 (±7) days post baseline assessment			
Randomisation and 90 (±7) days post baseline assessment			
Randomisation and 180 (±14) days post baseline assessment			
Last FIT session attended and 180 (±14) days post baseline assessment*	N/A		

* FIT participants only

Table 6: Number of participants attending the recommended number of FIT sessions before the assessments

Number of Participants	FIT + TAU N (%)
Attended ≥4 FIT sessions before 28 (±7) day assessment	
Attended ≥7 FIT sessions before 90 (±7) day assessment	
Attended ≥9 FIT sessions before 180 (±14) day assessment	
Attended final FIT session after 180 (±14) day assessment	

Table 7: Acceptability and adherence to the FIT intervention

Fit Session	Completed total N (%)	Completed within specified timeframe N (%)	Cumulative total of completion N (%)	Cumulative total of completion within specified timeframe N (%)
1 (in-patient)				
2 (face-to-face or telephone)				
3 (telephone)				
4 (telephone)				
5 (telephone)				
6 (telephone)				
7 (telephone)				
8 (telephone)				
9 (telephone)				
Total				

Table 8a: Completeness of proposed primary and secondary outcomes

Outcome	Time Point	TAU Number (%) of Participants			FIT + TAU Number (%) of Participants			All Number (%) of Participants		
		Attended Visit	Completed All Items	With Valid Score	Attended Visit	Completed All Items	With Valid Score	Attended Visit	Completed All Items	With Valid Score
Alcohol Use (grams of pure alcohol/ week)*	Baseline									
	28 (±7) days									
	90 (±7) days									
	180 (±14) days									
SADQ Score	Baseline									
	28 (±7) days									
	90 (±7) days									
	180 (±14) days									
WEMWBS	Baseline									
	28 (±7) days									
	90 (±7) days									
	180 (±14) days									
SWEMWBS	Baseline									
	28 (±7) days									
	90 (±7) days									
	180 (±14) days									
Re-hospitalisation	180 (±14) days		N/A			N/A			N/A	

* completeness of Timeline Follow-Back data, used by the Research Nurses to calculate Alcohol Use, is summarised below in Table 8b.

Table 9b: Completeness of timeline follow-back data (used to calculate proposed primary outcome of alcohol use)

Timeline Follow-Back data collection	Days before visit	TAU Number (%) of Participants				FIT + TAU Number (%) of Participants				All Number (%) of Participants			
		With Completed Units of Alcohol Data	Reason not collected			With Completed Units of Alcohol Data	Reason not collected			With Completed Units of Alcohol Data	Reason not collected		
			Could not remember	Not asked about alcohol	Other		Could not remember	Not asked about alcohol	Other		Could not remember	Not asked about alcohol	Other
Baseline	1												
	2												
	3												
	4												
	5												
	6												
	7												
28 (±7) days	1												
	2												
	3												
	4												
	5												
	6												
	7												
90 (±7) days	1												
	2												
	3												
	4												
	5												
	6												
	7												
180 (±14) days	1												
	2												
	3												
	4												
	5												
	6												
	7												

Table 10: Summary statistics of baseline and demographic participant characteristics

	TAU (n=)	**FIT + TAU (n=)**	**All (n=)**
Age			
Mean (SD) [range] *Median (IQR)*			
AUDIT Score			
Mean (SD) [range] *Median (IQR)*			
Gender n (%)			
Male Female			
Living Arrangements n (%) *			
Alone Spouse/Partner Parent/s Sibling/s Child/ren Rather not say Other			
Ethnicity n (%)			
White Asian Black Mixed/Multiple Other			
Place of Residence n (%) *			
Flat/Apartment House/Bungalow Residential/Care Home Other			

Occupation Status n (%) *			
Unemployed Student Part time work Full time work Age retired Medically Retired Other			
Year of diagnosis			
Mean (SD) [range] *Median (IQR)*			
Stage of Liver Disease n (%)			
Fatty *Fibrosis* *Cirrhosis*			
Child-Pugh Score			
Mean (SD) [range] *Median (IQR)*			
MELD Score			
Mean (SD) [range] *Median (IQR)*			

*Participants can provide data in more than one category – percentages will not necessarily add to 100%

Table 10: Summary statistics for the proposed primary and secondary outcome measures by allocated group

Variable	Time point	TAU			FIT + TAU		
		N	Mean (SD) (95% CI)/Median [IQR]	Range	N	Mean (SD) (95% CI) /Median [IQR]	Range
Alcohol Use (grams of pure alcohol/ week)	baseline						
	28 (±7) days						
	90 (±7) days						
	180 (±14) days						
SADQ Score	baseline						
	28 (±7) days						
	90 (±7) days						
	180 (±14) days						
WEMWBS	baseline						
	28 (±7) days						
	90 (±7) days						
	180 (±14) days						
SWEMWBS	baseline						
	28 (±7) days						
	90 (±7) days						
	180 (±14) days						
Ethyl Glucuronide (µg/L)	180 (±14) days						
Ethyl Sulphate (µg/L)	180 (±14) days						
Re-hospitalisation*	180 (±14) days						

* Rates and corresponding confidence intervals given for re-hospitalisation

Table 11: Summary statistics for the changes between baseline and 28, 90, and 180 day assessments for proposed primary and secondary outcomes including only those participants who complete all follow-up visits

Variable	Time point	TAU			FIT + TAU		
		N	Mean (SD) Change From Baseline (95% CI)/Median [IQR]	Range	N	Mean (SD) Change From Baseline (95% CI) /Median [IQR]	Range
Alcohol Use (grams of pure alcohol/ week)	28 (±7) days						
	90 (±7) days						
	180 (±14) days						
SADQ Score	28 (±7) days						
	90 (±7) days						
	180 (±14) days						
WEMWBS	28 (±7) days						
	90 (±7) days						
	180 (±14) days						
SWEMWBS	28 (±7) days						
	90 (±7) days						
	180 (±14) days						

Table 11b: Proportion of participants who completed TLFB assessment and with no alcohol consumption as calculated from TLFB

Variable	Time point	TAU	FIT + TAU
		N (% of participants with complete TLFB)	N (% of participants with complete TLFB)
Participants with zero alcohol consumption	Baseline		
	28 (±7) days		
	90 (±7) days		
	180 (±14) days		

Table 12: Re-hospitalisations by allocated group

Re-hospitalisation	TAU N (%)				FIT + TAU N (%)			
No. of re-admissions	0	1	2	>2	0	1	2	>2
By 28 (±7) days								
By 90 (±7) days								
By 180 (±14) days								

Table 13: Proportion of participants who relapse and time to relapse to alcohol use by allocated group

	TAU	FIT + TAU
N (%) who relapsed by:		
28 (±7) days		
90 (±7) days		
180 (±14) days		
Time to relapse (days)	**Median (Inter-quartile Range) [Range]**	**Median (Inter-quartile Range) [Range]**

Table 14: Unadjusted between-group differences and confidence intervals for proposed primary and secondary outcome measures, with unadjusted confidence intervals

Variable	Time Point	Mean Between-group Difference	Unadjusted 95% CI	Unadjusted 85% CI	Unadjusted 75% CI
Alcohol Use (grams of pure alcohol/ week)	28 (±7) days				
	90 (±7) days				
	180 (±14) days				
SADQ Score	28 (±7) days				
	90 (±7) days				
	180 (±14) days				
WEMWBS	28 (±7) days				
	90 (±7) days				
	180 (±14) days				
SWEMWBS	28 (±7) days				
	90 (±7) days				
	180 (±14) days				
Ethyl Glucuronide (µg/L)	180 (±14) days				
Ethyl Sulphate (µg/L)	180 (±14) days				
Re-hospitalisation*	180 (±14) days				

* Rates and corresponding confidence intervals given for re-hospitalisation

Table 15: Between-group differences and confidence intervals for proposed primary and secondary outcome measures, adjusted for stratification variables (dichotomised baseline SADQ score and site)

Variable	Time Point	Adjusted Mean Between-group Difference	Adjusted 95% CI	Adjusted 85% CI	Adjusted 75% CI
Alcohol Use (grams of pure alcohol/ week)	28 (±7) days				
	90 (±7) days				
	180 (±14) days				
SADQ Score	28 (±7) days				
	90 (±7) days				
	180 (±14) days				
WEMWBS	28 (±7) days				
	90 (±7) days				
	180 (±14) days				
SWEMWBS	28 (±7) days				
	90 (±7) days				
	180 (±14) days				
Ethyl Glucuronide (µg/L)	180 (±14) days				
Ethyl Sulphate (µg/L)	180 (±14) days				
Re-hospitalisation*	180 (±14) days				

* Rates and corresponding confidence intervals given for re-hospitalisation

Table 16: Unadjusted between-group differences and confidence intervals of change from baseline for proposed primary and secondary outcomes

Variable	Time Point	Mean Between-group Difference of Change From Baseline	Unadjusted 95% CI	Unadjusted 85% CI	Unadjusted 75% CI
Alcohol Use(grams of pure alcohol/ week)	28 (±7) days				
	90 (±7) days				
	180 (±14) days				
SADQ Score	28 (±7) days				
	90 (±7) days				
	180 (±14) days				
WEMWBS	28 (±7) days				
	90 (±7) days				
	180 (±14) days				
SWEMWBS	28 (±7) days				
	90 (±7) days				
	180 (±14) days				
Ethyl Glucuronide (µg/L)	180 (±14) days				
Ethyl Sulphate (µg/L)	180 (±14) days				

Table 17: Between-group differences and confidence intervals for of change from baseline for proposed primary and secondary outcome measures, adjusted for stratification variables (dichotomised baseline SADQ score and site)

Variable	Time Point	Adjusted Mean Between-group Difference of Change From Baseline	Adjusted 95% CI	Adjusted 85% CI	Adjusted 75% CI
Alcohol Use (grams of pure alcohol/ week)	28 (±7) days				
	90 (±7) days				
	180 (±14) days				
SADQ Score	28 (±7) days				
	90 (±7) days				
	180 (±14) days				
WEMWBS	28 (±7) days				
	90 (±7) days				
	180 (±14) days				
SWEMWBS	28 (±7) days				
	90 (±7) days				
	180 (±14) days				
Ethyl Glucuronide (µg/L)	180 (±14) days				
Ethyl Sulphate (µg/L)	180 (±14) days				

Table 18: Estimates of standard deviations of proposed primary outcome (alcohol use) for informing sample size calculations for a definitive trial

Parameter	Point Estimate	80% CI*	90% CI*
Standard deviation of alcohol use at baseline			
Standard deviation of alcohol use at 180 (±14) days assessment			
Standard deviation of change in alcohol use between baseline and 180 (±14) days assessment			

* upper bound of two-sided confidence interval relevant as conservative estimate for sample size calculations

Table 19: Estimates of correlation coefficients for proposed primary outcome (alcohol use) for informing sample size calculations for a definitive trial

Parameter	Point Estimate	80% CI*	90% CI*
Correlation coefficient between baseline and 180 (±14) days assessment of alcohol use			
Correlation coefficient between baseline and change between baseline and 180 (±14) days assessment of alcohol use			

* lower bound of two-sided confidence interval relevant as conservative estimate for sample size calculations

APPENDIX B: DERIVED VARIABLES

MELD SCORE

To calculate the MELD score:

- Cr and bilirubin must be in mg/dL, Na in mEg/L and INR is unitless
- $bilirubin = \begin{cases} 1.0 \\ bilirubin \end{cases} \quad if \quad \begin{matrix} bilirubin < 1.0 \\ bilirubin \geq 1.0 \end{matrix}$
- $INR = \begin{cases} 1.0 \\ INR \end{cases} \quad if \quad \begin{matrix} INR < 1.0 \\ INR \geq 1.0 \end{matrix}$
- $Cr = \begin{cases} 1.0 \\ Cr \\ 4.0 \end{cases} \quad if \quad \begin{matrix} Cr < 1.0 \\ 1.0 \geq Cr \geq 4.0 \\ Cr > 4.0 \mid \text{dialysis } \geq 2 \text{ past week} \mid 24\text{hrs CVVHD in past week} \end{matrix}$
- $Na = \begin{cases} 125 \\ Na \\ 137 \end{cases} \quad if \quad \begin{matrix} Na < 125 \\ 125 \geq Na \geq 137 \\ Na > 137 \end{matrix}$

The process of calculating the MELD score is described below:

$$MELD(i) = 10 \times (0.957\ x\ \ln(Cr) + 0.378 \times \ln(bilirubin) + 1.120 \times \ln(INR) + 0.643)$$

Equation 1

Round to 9 decimal places, then

$$MELD = \begin{cases} MELD(i) & if\ MELD(i) \leq 11 \\ MELD(i) + 1.32 \times (137 - Na) - [0.033 \times MELD(i) \times (137 - Na)] & if\ MELD(i) > 11 \end{cases}$$

Equation 2

The maximum MELD score is 40.

CHILD PUGH SCORE

The Child Pugh score is the total sum of all the individual items. The scoring of each item is shown below.

Table 20: Child Pugh Scoring of each item

Item		Score
Bilirubin	< 2 mg/dL (< 34.2 µmol/L)	1
	2 – 3 mg/dL (34.2 – 51.3 µmol/L)	2
	> 3 mg/dL (> 51.3 µmol/L)	3
Albumin	> 3.5 g/dL (> 35 g/L)	1
	2.8 – 3.5 g/dL (28 – 35 g/L)	2
	< 2.8 g/dL (< 28 g/L)	3
INR	< 1.7	1
	1.7 – 2.2	2
	> 2.2	3
Ascites	Absent	1
	Slight	2
	Moderate	3
Encephalopathy	No Encephalopathy	1
	Grade 1 – 2	2
	Grade 3 – 4	3

SWEMWBS

Table 21: Conversion table for the SWEMWBS score

Raw Score	Metric Score
7	7.00
8	9.51
9	11.25
10	12.40
11	13.33
12	14.08
13	14.75
14	15.32
15	15.84
16	16.36
17	16.88
18	17.43
19	17.98
20	18.59
21	19.25
22	19.98
23	20.73
24	21.54
25	22.35
26	23.21
27	24.11
28	25.03
29	26.02
30	27.03
31	28.13
32	29.31
33	30.70
34	32.55
35	35.00