SCENE WP5

Statistical Analysis Plan

Version: 1.0
Date: 15/09/2023

Person(s) contributing to the analysis plan	
Name(s) and position(s)	Fionn Cléirigh Büttner, trial statistician Thomas Hamborg, senior statistician Rianna Mortimer, trial statistician Stefan Priebe, co-chief investigator
Authorisation	
Position	**Co-Chief investigator**
Name	**Stefan Priebe**
Signature	Stefan Priebe
Date	18/09/2023
Position	**Senior trial statistician**
Name	**Thomas Hamborg**
Signature	Thomas Hamborg Digitally signed by Thomas Hamborg Date: 2023.09.18 15:19:39 +01'00'
Date	
Position	**Independent statistician**
Name	**Rebecca Walwyn**
Tick once reviewed	√
Date	12/09/2023

Table of Contents

1. Administrative Information

1.1 Trial registration number: ISRCTN 15815862

This SAP is based on protocol version 3.0 (21/09/2020)

1.2 SAP revision history

Protocol version	Updated SAP version no.	Section number changed	List of changes from previous version/protocol	Author of change	Date
1.0	0.1		Initial draft	TH	10/06/19
1.0	0.2		DG, AS & SE review	DG, SE	08/07/19
1.0	0.3		Amendments in response to v0.2 review	TH	12/07/19
1.0	0.4	3, 4 and 5.1	Updated sections 3,4 & 5.1 plus other minor amendments	TH	31/07/19
1.0	0.5	5.8 and 5.12	Updated sections 5.8 & 5.12 plus some minor comments	RM	11/12/19
1.0	0.6	2, 5.6, 5.12	Additional sensitivity analysis plus minor comments (in response to PMG discussion)	RM	21/02/20
1.0	0.7	1.3, 1.4, 3.3, 4.2, 5.1 5.2, 5.3, 5.12, Appendices C and D	Minor amendments to 1.3, 1.4, 3.3, 4.2, 5.1, 5.2, 5.3, 5.12, Appendices C and D plus some minor comments.	RM	02/07/20
1.0	0.8				
3.0	0.9			ET, TH	21/1/2021
3.0	0.10			FCB, TH	16/02/2023
3.0	0.10	All	Accepting changes, removing resolved comments	TH	11/06/2023

3.0	0.11	All	Incorporating independent statistician's edits and suggestions.	FCB	26/06/2023
3.0	0.12	All	Incorporating final review of CI and finalising tables in appendix	FCB/TH	15/09/23
3.0	1.0	NA	Sign off	TH	15/09/23

*If the SAP has been published, indicate which version.

1.3 Members of the writing committee

The SAP writing committee comprises or has comprised Thomas Hamborg (TH), Sandra Eldridge (SE), Rianna Mortimer (RM), Evangelia Tzorovili (ET), and Fionn Cléirigh Büttner (FCB). Early input was provided by Domenico Giacco. SE devised the initial design and analysis strategy. FCB and TH are primarily responsible for writing the SAP. RM and ET contributed to earlier drafts. FCB is responsible for implementing the statistical analysis strategy and TH for overseeing it. Stefan Priebe (blinded co-chief investigator) and Rebecca Walwyn (independent statistician) reviewed the SAP and provided critical feedback prior to sign-off.

1.4 Timing of statistical analyses

The analysis of 6-month follow-up data (i.e., primary outcome time point) will be conducted after 6 months follow-up has been completed. Data from subsequent follow-up time-points (i.e., 12 months and 18 months) will be analysed separately when the entire follow-up has been completed.

1.5 Timing of SAP revisions in relation to unblinding of data/results

Versions of the SAP up until version 1.0 were written whilst contributors did not have access to unblinded trial data or trial results by treatment group.

1.6 Analysis software

Analyses and data presentation described in this document will be performed using Stata version 17.0 or later, unless otherwise specified.

1.7 Remit of SAP

This document aims to detail statistical analyses and presentation of results of the effectiveness and mechanistic evaluation analysis of the SCENE trial. This SAP does not include health economic analyses, the process evaluation, or the fidelity study associated with the SCENE trial. These analyses will be/are described in separate documents.

2. Background and trial design

Study objectives	Aim: To assess the clinical effectiveness and cost-effectiveness of a psychosocial intervention to improve social networks of patients with psychosis compared with an active control condition (i.e., information on social activities available in the local area). Specific objectives: 1. Assess whether the intervention improves quality of life, as measured by the MANSA, of patients with psychosis (primary outcome) compared with the active control; 2. Assess whether the intervention improves secondary outcomes number of social contacts, mental health symptoms, social situation, feelings of loneliness, time spent in social activities, health-related quality of life and reduces service use; 3. Assess whether changes in quality of life are mediated by an increase in the number of social contacts; 4. Assess costs and cost-effectiveness of the intervention (not part of the SAP); 5. Evaluate implementation of the intervention and explore the processes which are associated with intervention effects (partially part of the SAP).
Study design	Individually randomised, parallel-group, controlled superiority trial. Patients will be individually randomised 1:1 to either the intervention or control arm. Randomisation will be stratified by site (NHS Trust). Permuted block randomisation with block sizes of m=6, 4, and 2 will be used within each stratum. Patients will be allocated to clinicians based on locality and availability – that is, not randomly.
Setting	Multi-centre trial across 14 Mental Health NHS Trusts in England. Data collection will be conducted in quiet rooms within NHS Trust facilities or at participants' homes/community facilities if the participant prefers.
Participants	**Inclusion Criteria (patients):** - 18-65 years old - Diagnosis of psychosis-related condition (ICD-10 F20-29) - Capacity to provide informed consent - Ability to communicate in English - Limited social network size (≤3 social contacts with non-first-degree relatives in the previous week outside of home, work and mental health services) - Low quality of life (≤ 5.0 on MANSA quality of life assessment) **Exclusion Criteria (patients):**

	- Does not meet inclusion criteria - Primary problem of current drug addiction - No capacity to provide written informed consent - An inpatient on a psychiatric ward at the time of recruitment
Interventions	**Intervention:** Psychosocial intervention comprising six sessions within 6 months with a clinician ("social contacts coach") to improve the social network of patients. A patient is classed as 'completing' the intervention if they attend three or more sessions. **Control:** Information about local options for social activities provided by the researcher. Usual mental health treatment will not be affected by participation in this study, neither in the intervention group nor in the control group.
Primary outcome measure(s)	Mean MANSA score at the end of the intervention period (i.e., six months follow-up) compared between intervention and control groups.

3. Outcome measures

3.1 Timing of outcome assessments

All patient participant outcomes are measured at baseline and at six months, 12 months, and 18 months after randomisation, see table below. Details on the scoring of outcomes can be found in Appendix A. The 'visit window' for each follow-up time point is defined as *time point minus two weeks* to *time point plus two months*.

Assessment	Screening	Baseline	End of intervention phase (6 months follow-up)	Covid-19 additional Follow-up (~10 months)	12 months follow-up	18 months follow-up
	All Patient Participants					
MANSA	**X**	**x**	**x**	**x**	**x**	**x**
Social Contacts Assessment	**x**	**x**	**x**	**x**	**x**	**x**
PANSS		**x**	**x**	**x**	**x**	**x**
Social situation		**x**	**x**	**x**	**x**	**x**
UCLA-8 Loneliness Scale		**x**	**x**	**x**	**x**	**x**
Time spent in social activities		**x**	**x**	**x**	**x**	**x**
EQ-5D-5L		**x**	**x**	**x**	**x**	**x**
Client service receipt inventory		**x**	**x**	**x**	**x**	**x**
Healthcare source use (NHS Digital)		**x**	**x**	**x**	**x**	**x**
	Intervention Participants only					
Semi-structured interviews			**x**	**x**		
	Clinician Participants					
Adherence schedule			**x**	**x**		
Semi-Structured Interviews			**x**			

3.2 Primary outcome

Subjective quality of life measured on the Manchester Short Assessment of Quality of Life (MANSA) at the end of the intervention (6 months after randomisation) (1).

- Range: 1-7, the MANSA score is the mean of twelve item responses which are each scored 1-7 (see Appendix A for further details), with higher scores indicating better quality of life [continuous].

3.3 Secondary outcomes

- Social situation using the Objective Social Outcomes Index (SIX) (2)
  - Range: 0 – 6. Higher scores indicate better outcome [ordinal].
- Psychopathological symptoms using the Positive And Negative Syndrome Scale (PANSS) (3,4)
  - Range: 30 – 210. The outcome measure is the overall score (sum of the three sub-scales positive, negative, general) consisting of 30 items in total, with lower values indicating better outcome. The three subscale scores will also be reported. [continuous].
- Feelings of loneliness (UCLA 8-item Loneliness Scale) (5)
  - Range: 8 – 32. The outcome measure is the sum of all eight items, with lower values indicate better outcome [continuous].
- Time spent in social activities (Time Use Survey) (6)
  - Range: 0 – 10,080. The outcome measure is the sum of time (i.e., minutes/week) spent in leisure/spare time activities AND sports activities during the previous week, with higher values indicating a better outcome. Time spend in an individual activity is capped at 1500 min/week [continuous].
- Number of face-to-face social contacts with non-first-degree relatives during the previous week outside of the home, work, and mental health services using the Social Contacts Assessment (SCA) (7)
  - Range: 0 – ∞. Meeting one person will count as one contact even if that person is met more than once during the previous week, with higher values indicating a better outcome [discrete, count].
- Subjective quality of life measured on the Manchester Short Assessment of Quality of Life (MANSA) at 12 and 18 months after randomisation.
- Health Related Quality of Life (HRQoL) using EQ-5D-5L utility (8)
  - The EQ-5D-5L questionnaire assesses participants' health-related quality of life (9). The EQ-5D-5L comprises 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression, each rated on a scale from 1 to 5, corresponding to no problems (1), slight problems (2), moderate problems (3), severe problems (4), and extreme problems (5). Overall QoL utility scores will be

derived for all contributing study participants using the UK National Institute for Health and Care Excellence (NICE) decision support unit EQ-5D scoring algorithm (15). The estimation algorithm will directly map from individual-specific, EQ-5D-5L, health states to individual-specific, EQ-5D-3L, utility scores, using age and sex as necessary covariates (10). The Stata command *eq5dmap* will be used to derive the utility scores. Estimated, individual-specific, EQ-5D-3L utility scores will be used as the secondary outcome during statistical analysis. The overall score of the EQ-5D-3L index ranges from -0.594 to 1.000 (i.e., higher scores correspond to a better quality of life).

3.4 Further outcomes added during the COVID-19 pandemic

- Time spent in online social activities
    - Range: 0 – 10,080. The outcome is the time (i.e., minutes/week) spent in online social activities during the previous, with higher values indicating a better outcome [continuous].

- Number of remote social contacts with non-first-degree relatives during the previous week outside of the home, work, and mental health services
    - Range: 0 - ∞. Interacting with one person remotely (i.e., via text message, email correspondence, telephone call, or video call) will count as one contact even if that person is met more than once during the previous week, with higher values indicating better outcome (i.e., increased social contacts) [discrete, count].

- Patient Health Questionnaire-9 (PHQ-9) for depression symptom severity (11)
    - Range: 0-27. The PHQ-9 is the sum of nine items, each of which are scored 0-3 by the participant. Higher scores correspond to greater depression severity [continuous].

- Generalised Anxiety Disorder-7 (GAD-7) score for anxiety (12)
    - Range: 0-21. The GAD-7 is the sum of seven items, each of which are scored 0-3 by the participant. Higher scores correspond to greater generalised anxiety [continuous].

Unless otherwise stated above, if >20% of questionnaire/scale items are missing, the total/overall score will be set to missing. If ≤20% of questionnaire /scale items are missing, missing values will be imputed using the mean value of the present items for this participant. This approach will be applied to a domain/dimension/subscale (instead of across all items) if the outcome has different domains/dimensions/subscales. The same approach shall be used for relevant mediator variables, as necessary.

4. Study methods

4.1 Sample size calculation

It is assumed that the proposed new intervention would be implemented and funded across the NHS only if it achieved at least a medium effect size. An effect size of 0.35 equates to improved satisfaction ratings on the MANSA of at least one scale point (on a seven-point scale) on four out of a total of 12 life domains. An improved quality of life in four life domains is usually regarded as a meaningful difference to patients' life (13).

For detecting the described effect size with 90% statistical power, assuming a conservative ICC of 0.07 for patients treated by the same clinician in the intervention arm, 229 patients in the intervention group and 229 in the control group will be required (total sample = 458). This sample size has been calculated using an iterative search algorithm. Initially, the required sample size for the pre-specified clinically relevant improvement and statistical power for a range of different pre-specified allocation ratios is calculated and the sample size in the intervention arm is then inflated to account for clustering due to participants being treated by the same clinician. Then, the minimal sample size resulting in equal group sizes is identified. This requires eight additional patients to be recruited compared with the absolute minimum sample size required (with slightly uneven groups). Assuming a drop-out rate (from the study) of 20% at six months follow-up, which is in line with recent trials of similar interventions with the same patient group (6,14), we will have to recruit a total sample of 572 patients, 286 in the intervention and 286 in the control group. The sample size calculation is based on 10 patients being treated and followed-up per clinician, on average. To account for drop-out, 12 patients need to be allocated to each clinician and therefore 24 clinician-coaches recruited to participate in the study. Based on recruiting 12 patients per clinician, the final total sample size is 576 (288 per arm).

Update:
Recruitment and intervention delivery to SCENE were paused during the Covid-19 pandemic. This made a study extension necessary. As part of the extension a sample size re-calculation was conducted. The dropout rate was inflated from 20% to 25% to reflect the actual loss to follow-up.

The mean cluster size in the intervention arm was reduced from ten patients to three patients per cluster as it was observed in practice that each coach was seeing approximately three patients, on average. The variability in the number of coaches per site (cluster) was also considered using the coefficient of variation (CV) of the cluster size, CV=0.37. Accounting for a 25% dropout rate, 504 participants are needed in total (252 participants per group).

4.2 Randomisation procedure (taken from protocol v3.0)

Patients will be randomised to either the intervention or control arm. There will be a 1:1 allocation ratio. Randomisation will be stratified by NHS Trust, ensuring a balanced number of patients randomised to each trial arm at each NHS Trust. Permuted block randomisation with block sizes of six (m=6), four (m=4), and 2 (m=2) will be used within each stratum. Patients will be allocated to clinicians based on locality and availability (i.e., not randomly).

Randomisation will be performed remotely by the Pragmatic Clinical Trials Unit at Queen Mary, University of London. One researcher per site will be given a login to the system to complete

randomisation at that site. Further details will be explained in the Data Management plan, which will be agreed and signed off between the trial study team and PCTU.

Deviation:
Due to issues with the SCENE randomisation system identified in April 2019, randomisation was moved to a manual system. Sites email an independent statistician at the PCTU who randomises the eligible participant according to a pre-determined randomisation list. The independent statistician then emails the unblinded researcher at the site with the allocation and participant ID.

4.3 Blinding (taken from protocol v3.0)

Due to the nature of the trial interventions, participants cannot be masked to treatment allocation. Researchers involved in assessing outcome measures will be blinded to participants' allocation. To minimise the risk of researchers becoming unblinded during the follow-up assessment, we will instruct participants to avoid revealing their allocation. To facilitate this, there will be one unblinded researcher per site (in addition to the principal investigator), who will organise assessments and remind participants to conceal their allocation. At the end of post-randomisation assessments, blinded researchers/outcome assessors will record their guesses as to whether participants are in the intervention arm or in the control arm. The trial co-lead (i.e., Priebe) will remain blinded to patients' allocation status and all post-randomisation information. The statistician analysing the trial will remain masked to patients' allocation until the Statistical Analysis Plan has been signed off and the trial database finalised and locked for statistical analysis.

5. Analysis methods

5.1 Data cleaning process

Data cleaning will be aligned with the analysis timing described in 1.5. For the primary outcome analysis time point (i.e., six months after randomisation), once the research team have completed all data entry and checking, the data management team will part-freeze the database and make it available to the statistician responsible for the analysis via Excel spreadsheets. As described in the data management plan, the trial statistician will conduct data checks in addition to checks performed by the data management and research team. Data quality checks performed by the trial statistician will include complex range, logical, and consistency checks that may not be picked up by checks performed at the level of individual records. Discrepancies found during data checks will be reported to data management and the trial manager for updating the master database, as necessary. This data cleaning process will be repeated until the trial statistician and the data management team are satisfied that all identifiable errors have been corrected. Once all discrepancies have been identified and managed, the database will be soft-locked, so existing data cannot be edited or accessed, and used for analysis. If unforeseen queries are generated during the statistical analysis, they will be managed on a case-by-case basis. Changes to the data will be recorded in a Stata do-file and clearly documented in an appendix to the statistical analysis report to ensure transparency and reproducibility of the analysis. For all other analyses at the end of the trial, the same procedure will be followed as above but a full database freeze and lock, rather than a partial freeze, will be employed.

5.2 Baseline characteristics

Patient participant baseline characteristics will be summarised for each treatment arm by the mean and standard deviation, or median and interquartile range, for continuous or discrete variables as appropriate, and the number and percentage for categorical variables. No p-value for between-group differences will be presented. A draft table for the descriptive summary of baseline characteristics is given in Appendix D.

5.3 Information for CONSORT flow diagram

A dummy flow diagram, which will indicate the progression of participants through recruitment, enrolment, allocation, intervention provision, follow-up/attrition, and analysis of patients during the trial, is provided in Appendix C.

5.4 General analysis principles

Statistical analyses will be performed according to the intention-to-treat principle – all randomised participants with a recorded outcome will be included in the analysis and analysed according to the trial arm to which they were randomised. Participants who withdraw consent for their data to be included in the analysis will be excluded from all analyses.

For the analysis of the primary outcome and each secondary outcome, we will present the following information:

- The number of participants included in each analysis, by treatment arm
- A summary statistic of the outcome (e.g., number (%)), by treatment arm
- The estimated treatment effect θ
- A 95% confidence interval for the estimated treatment effect

- A p-value for the two-sided hypothesis test of H_0 θ=0

For all statistical analyses a significance level of 5% will be used.

5.5 Analysis of primary and secondary outcomes

The primary outcome analysis will compare MANSA scores between treatment groups at six-months after randomisation using a three-level, mixed-effects model. Trial site (level three) and clinician-coach (level two) will be fitted as random intercepts. MANSA score at baseline will be included as the only model covariate. A partially nested mixed-effects model with heteroskedastic error terms will be fitted with the Satterthwaite approximation for degrees of freedom to avoid upward bias of the type-I error rate (15-17). In the intervention arm only, a random intercept will be specified to allow for the clustering by clinician-coach. Participants in the control arm will be treated as independent.

Specifically, let *y* be the continuous outcome, *i* is the individual participant indicator, *j* is the clinician-coach indicator, *t* is the intervention indicator (0 = control, 1 = intervention), θ is the intervention effect, β_0 is an intercept term, *l* is the site indicator, and β_k represents further covariates. Then,

$$y_{ij} = \beta_0 + \theta t_{ij} + \beta_k + v_l + u_j t_{ij} + r_{ij}(1 - t_{ij}) + \epsilon_{ij} t_{ij}$$

where $u_j \sim N(0,\sigma^2_u)$ is a random-effects term representing between-cluster (coach) variation in the clustered intervention arm, $r_{ij} \sim N(0,\sigma^2_r)$ represents individual-level variation in the non-clustered control arm, $\epsilon_{ij} \sim N(0,\sigma^2_\epsilon)$ represents individual-level variation in the clustered intervention arm and $v_l \sim N(0,\sigma^2_w)$ is the random effect representing between-site variation.

No imputation at the outcome variable level will be used in the main analyses (i.e., complete case analysis). Mean imputation will be used for the baseline MANSA score. Blinded assessment of baseline data completeness suggests less than 1% of missing data for MANSA and other outcome variables.

Secondary outcomes at six-months after randomisation will be analysed using the same mixed-effects model as for the primary outcome, or an equivalent model appropriate for the outcome type where the secondary outcome is not continuous and/or normally distributed. The Social Contact Assessments outcome will be analysed using a negative binomial model. For the Time Use Survey, a previous study suggests that TUS outcome data is likely to be skewed (14). The distribution of TUS outcome data will be assessed graphically, and we will assess transforming TUS outcome data to a log-normal distribution. Separate analyses will be performed for outcomes at 12- and 18- months after randomisation.

The main analysis will use all observed outcome data (i.e., complete-case analysis), which is appropriate and efficient for estimating the ITT estimand (i.e., the average effect of randomisation/offering the intervention) when assuming missing outcome data are Missing At Random (MAR) given model covariates that have very little missingness (as is anticipated here) (18). Sensitivity analyses under Missing Not At Random (MNAR) assumptions will be conducted and are described in section 5.11. Consistent with other studies using the MANSA, we will allow up to two out of 12 satisfaction items to be missing without performing imputation (see Appendix A). The ICC estimated from the primary outcome analysis and its precision will be reported if possible. Cluster size distribution and the number of clusters present will also be reported.

Strategy for analysis of primary and secondary outcomes if model fails to converge.

If the analysis described above (i.e., row 0 in the below table) fails to converge for any outcome, the following sequential strategy will be employed for assessing between-group differences.

	Change from previous strategy	Example Stata code
0	Primary analysis	`mixed y_6m treat y_B || site: ||// therapist:treat, nocons reml /// residuals(independent, by(treat)) /// dfmethod(sat)`
1	Remove clustering of patients by clinician-coaches	`mixed y_6m treat y_B || site: ||// nocons reml /// residuals(independent, by(treat)) /// dfmethod(sat)`
2	Fit stratification factor, trial *site*, as fixed effect	`mixed y_6m treat y_B site ||// therapist:treat, nocons reml /// residuals(independent, by(treat)) /// dfmethod(sat)`
3	Remove clustering of patients by clinician-coaches and remove trial *site*	`regress y_6m treat y_B`
4	Remove covariate *baseline outcome* from the model (i.e., simple between group t-test)	`regress y_6m treat`

y= outcome, treat = intervention arm indicator, therapist = intervention group therapist indicator

5.6 Categorical variables

Trial site will be the only categorical variable included in analysis models. Each participating NHS Trust is a *site* apart from East London NHS Trust where East London and Luton are treated as individual site at randomisation and analysis. The total number of categories of the variable *site* is 14. As described in 5.5 *site* is included in analyses as a random-intercept effect.

5.7 Implementation fidelity

Implementation fidelity (i.e., the extent to which the intervention is delivered as intended) will be measured using the SCENE Adherence Scale. The SCENE Adherence Scale was developed by study investigator of the SCENE trial and assesses whether social coaches deliver the eight steps of the intervention. It comprises eight items that reflect the intended structure of the SCENE intervention. Each item is scored 0 (i.e., "Item not completed"), 1 (i.e., "Social Contacts Coach (SCC) and service user (SU) discuss some aspects, but encountered problems"), or 2 (i.e., "Item completed, and everything worked well"). The total score range is 0-16. 'Good compliance' refers to achieving '2' on all the individual 8 items across the intervention sessions. The scale is completed by the coach after delivering each intervention, as a record of what was achieved during each session. It is usually only done after the first few sessions. Once coaches have achieved a 2 on all eight items of the scale, they are no longer required to complete the form for subsequent sessions with that patient.

5.8 Interim analyses

The Data Monitoring and Ethics Committee (DMEC) will review recruitment, outcome, and safety data periodically during the trial. No interim analyses with formal stopping rules for either superiority or inferiority are in place.

5.9 Mediation analysis

A mediation analysis will be performed to identify whether the effect of random treatment allocation on the primary outcome (i.e., MANSA at six months after randomisation) is mediated through expanded social networks (SCA) at six months after randomisation, as hypothesised. A multi-level, structural equation model (SEM) will be constructed by fitting the explanatory, mediating, and outcome variables in a single analysis to estimate natural direct, natural indirect, and total intervention effects (19). Because the aim of this analysis is mediation, only a 'structural' model – a model with paths reflecting causal dependencies between endogenous and exogenous variables – with only observed variables, will be fitted (20). Random treatment allocation will be specified as an exogenous variable, and SCA and MANSA at six-month follow-up will comprise endogenous variables. Uncorrelated error terms will be indicated for both endogenous variables. A random intercept for trial site will also be fitted.

The mediator and outcome variable, and the amount of corresponding missing data, will be summarised using mean and standard deviation, or frequency and percentage, as appropriate. A negative binomial model equation will be specified when fitting SCA as the outcome (i.e., ultimately the mediator) and a linear model equation will be specified when fitting MANSA at 6-month follow-up as the outcome. Baseline and follow-up mediator and outcome variables will be standardised to baseline by subtracting the mean of the outcome variable at baseline and dividing by the standard deviation (SD) of the outcome variable at baseline (21). Thus, model coefficients will be interpreted in baseline SD units of the outcome both for direct and indirect/mediated effects. A single mediator model with a contemporaneous mediation (*b*) path – where the mediator and outcome are both measured at six months after randomisation – will be fitted (22). The mediator model will adjust for the baseline measure of the mediator and the baseline measure of the outcome in equations for both the mediator and the primary outcome (23).

The ''product of coefficients'' approach will be applied to calculate the indirect (mediated) effect by multiplying the intervention regression coefficient (a path) by the mediation regression coefficient (b path) (24,25). Percentile bootstrap 95% confidence intervals (CI) will be calculated for these effects, using 1000 repetitions. Full mediation and partial mediation will be considered based on a change in direct/intervention and indirect/mediated effect estimates from unadjusted to adjusted analysis (26). The gsem command and associated options in Stata 17.0 will be used to perform mediation analyses. If the proposed mediation analysis model fails to converge, the SEM will be simplified by removing the random intercept for clinician-coach (i.e., level two) from the model. A second mediation analysis, involving a lagged mediation (*b*) path, will evaluate the mediating effect of increases in SCA at 6-month follow-up on patients' MANSA score at 12-month follow-up using the same analysis method.

5.10 Complier-average causal effect analysis

We will perform a complier-average causal effect (CACE) analysis to estimate the effect of the intervention on the primary outcome (i.e., MANSA at 6 months follow-up) with a latent approach using structural equation modelling. The CACE treatment effect will be defined as the difference, on average, between compliant participants who were randomly assigned to the intervention arm and

participants in the control arm who would have complied with treatment had they been randomised to the intervention arm (27). Using the *gsem* command in Stata, we will specify two regression paths within the structural equation model – a regression path for compliers and a regression path for non-compliers.

'Compliers' will only be observed in the intervention arm, where an indicator variable will identify whether the participant complied. Participants in the intervention arm will be classified as "compliant" if they underwent an initial 60-minute session with a clinician-coach AND at least one 20-minute review session AND a final session. Latent mixture modelling will be used to identify participants in the control group who would have complied with treatment if they had been randomly allocated to the intervention arm (28). The latent class variable, 'compliance', will be determine using relevant predictors. Specifically, compliance among control group participants will be estimated using (i) the observed compliance data available for the participants randomized to treatment, (ii) the missing compliance data for the participants randomized to the control group, and (iii) the distribution of the outcome variable in the sample (27). We will run two CACE analyses that (i) excludes covariate adjustment, and (ii) adjusts for MANSA score at baseline in the analysis model. The covariate-adjusted CACE will be considered the primary CACE analysis. Both models will adjust for clustering by clinician-coach to reflect the structure of the primary outcome analysis.

We will assume (i) monotonicity (i.e., there will be no "defiers" or "always-takers" in the study sample), (ii) stable unit treatment value (i.e., a participant's outcome depends only on their own group assignment and not on the group assignment of other participants), (iii) random assignment (i.e., exchangeability between intervention and control arms with respect to the trial outcome), and (iv) exclusion restriction (i.e., the treatment effect estimate will be fixed at zero for 'non-compliers' but freely estimated for 'compliers'). It has been argued that in the event of non-compliance in partially nested RCTs the CACE estimate is the unbiased ITT estimate (29).

5.11 Sensitivity analyses

a) An analysis of multiply imputed data will be performed on the primary outcome. A four-step strategy for analysing randomised trials with missing data, suggested by White & colleagues will be followed (30). Multiple imputation will be used to include participants with partially completed and not completed questionnaire data in the primary outcome analysis (31,32). Specifically, MANSA scores at baseline and at six-, 12-, and 18-month follow-up (i.e., after randomisation) will be imputed at score, not individual item, level when there are more than two missing MANSA items. Model variables will be included in the imputation model and will comprise MANSA scores at baseline, treatment site, facilitator (i.e., social coach), SCA number (included for the mediator analysis). Auxiliary variables – that is, baseline characteristic variables that are not included in the primary outcome analysis but that are moderately correlated (i.e., r>0.4) with (i) the observed values of the missing variable or (ii) its missingness – will also be included in the imputation model (33). We will assume missing data are Missing At Random (MAR). Multiple imputation by chained equations (MICE) with 50 imputations, and a pooling step combining estimated dataset-level results using Rubin's rules, will be employed. Imputation will be performed separately within each trial arm (i.e., arm-specific MAR).

b) A sensitivity analysis will be performed using imputed data that account for missing outcome data during the COVID-19 pandemic to assess the robustness of primary outcome analysis results. We will use a strategy that aims to facilitate the handling of missing outcome data in trials that overlap with the COVID-19 pandemic (34). First, because the SCENE trial (recruitment: March 2019 – April 2022) overlapped with the COVID-19 pandemic in the UK (March 2020 – February 2022), the primary treatment effect (i.e., MANSA at six months post-randomisation) partially occurred in the presence of the COVID-19 pandemic. Consequently, different types of events arising due to the COVID-19 pandemic are expected to have affected the presence and interpretation of the primary treatment effect estimate. The presence of MANSA scores at six months post-randomisation is expected to be *indirectly affected* by the COVID-19 pandemic due to disruption to the intervention's delivery (i.e., transitioning to online), Government-imposed social restrictions (i.e., nation-wide lockdowns limiting social interactions), and/or potential participant behaviour changes (e.g., pandemic-related anxiety) due to the COVID-19 pandemic.
Second, we will identify missing data that are required for the primary treatment effect estimate (e.g., MANSA score items at baseline, and at six, 12, and 18 months after randomisation). Participants' pre-, peri-, and post-pandemic data will be used during imputation analysis. We will define four time-periods related to the COVID-19 pandemic during which social restrictions would plausibly affect trial participants and the primary outcome:

Pre-pandemic: April 2019 – March 22nd, 2020
Peri-pandemic: March 23rd, 2020 (i.e., announcement of first Government-enforced national lockdown in the UK) – February 23rd, 2022
- 'Heavy' restrictions
- 'Some' restrictions

Post-pandemic: February 24th, 2022 (i.e., official removal of all COVID-19-related restrictions) – Trial end-date (i.e., follow-up assessment of final participant)

During the pandemic (i.e., the 'peri-pandemic' phase), we will define two sub-phases to represent the severity of social restrictions imposed during the COVID-19 pandemic (i.e., heavy restrictions and some restrictions). We will add one week to the beginning and end of the 'heavy' restriction time-period [in parentheses below], and to the final, 'some' restrictions time-period, to account for the fact that many of the outcome measurements concern experiences over the week prior to outcome measure completion.

'Heavy' restrictions (i.e., Government-enforced national lockdowns in the UK or Tier 4 ('Stay at Home') restrictions):
March 23rd, 2020 [March 16th, 2020] – May 31st, 2020 [June 7th, 2020]
November 5th, 2020 [October 29th, 2020] – December 1st, 2020 [December 8th, 2020]
January 5th, 2021 [December 29th, 2020] – March 28th, 2021 [April 4th, 2021]

December 20th, 2020 [December 13th, 2020] – January 4th, 2021 (Tier 4 restrictions were in place across London, Luton, Gloucestershire, Somerset, TEWVS, and Oxford until national the third national Government-enforced national lockdown in the UK was introduced).

'Some' restrictions (i.e., pandemic-related restrictions excluding Government-enforced national lockdowns in the UK or Tier 4 ('Stay at Home') restrictions):
June 1st, 2020 – November 4th, 2020
December 2nd, 2020 – January 4th, 2021
March 29th, 2021 – February 23rd, 2022 [March 2nd, 2022]

Next, an imputation analysis (using multiple imputation by chained equations) will be conducted under missing data assumptions that will be selected based on whether missing data occur in a pre- (i.e., not affected), peri- (i.e., indirectly affected), or post-pandemic (i.e., not affected) phase. Controlled multiple imputation (MI) with reference-based approaches will be applied, using within-trial information, to qualitatively specify the distribution of unobserved outcome data by referencing other groups of participants in the trial (35). For participants who are *indirectly affected* by the pandemic (e.g., due to 'heavy' restrictions or 'some' restrictions), unobserved outcome data will be assumed to be missing not at random (MNAR). For participants with missing outcome data during a period of 'some' restrictions, we will incorporate a "copy-increments-in-control-arm" method, which imputes data assuming that the participant behaviour follows the trajectory of earlier assessment timepoints in the control arm, after the participant's withdrawal from the trial. For participants with missing outcome data during a period of 'heavy' restrictions, we will incorporate a "jump-to-control arm" reference method, which imputes data assuming participant behaviour reflects that of participants in the control arm after the participant's withdrawal from the trial. Both reference methods retain pre-drop out information for participants from the intervention arm who deviate from the study protocol.

For participants with missing outcome data pre- (i.e., before March 23rd, 2020) or post-pandemic (i.e., from February 24th, 2022), we will assume data are MAR given recorded data that are associated with the outcome and outcome missingness. Pandemic time-period (i.e., pre-, peri-, or post-pandemic) will be included within the missing data assumption for the imputation analysis to ensure that missing data for unaffected participants (e.g., pre- or post-pandemic) are modelled based on the observed data of unaffected participants (and not based on observed values of indirectly affected participants).

c) A further analysis will be conducted to estimate the treatment effect accounting for the phases of the COVID-19 pandemic. Using the primary outcome analysis described in 5.5, the effect of the trial intervention on the MANSA score at six-months after randomisation will be compared to the control intervention in each phase. Four phases – pre-pandemic, peri-pandemic ('heavy' restrictions), peri-pandemic ('some' restrictions), and post-pandemic – will be used, as defined above, to perform a fixed-effect meta-analysis with inverse-variance weighting. The pandemic phase – pre-, peri-, or post-pandemic – in which trial participants undergo their outcome assessment will be the phase of analysis for each trial participant in

the phase-specific analyses described above. If any of the seven days prior to the outcome assessment falls within the previous pandemic phase, the participant will be assigned this pandemic phase.

d) For a subset of patients their intervention delivery was paused due to the COVID-19 pandemic. Intervention was later resumed and completed at approximately 10 months after randomisation. A sensitivity analysis will consider actual end-of-treatment (i.e., approximately 10-months post-randomisation) as the primary assessment time-point for these participants, and hence include the assessment of MANSA scores at 10 months after randomisation instead of the six months post-randomisation assessment.

e) A sensitivity analysis will investigate the mediating effect of the number of social contacts in the relationship between the intervention and MANSA at six months after randomisation when face-to-face and remote (i.e., online) social contacts are combined. An identical model will be constructed as employed for the mediation analysis described in section 5.9.

f) A sensitivity analysis will investigate the effect of the intervention on time spent in social activities when the total minutes per week of spare time/sports activities (as measured by the Time Use Survey) are summed with the total minutes per week of online social activities (as measured using the Online Time Use Survey).

g) The tenability of the exclusion restriction assumption (that the treatment effect is zero for non-compliers) in the CACE analysis will be assessed using a sensitivity analysis. Instead of restricting the treatment effect estimate amongst non-compliers to zero (as specified in the primary CACE model), we will allow the treatment effect amongst compliers AND non-compliers to be freely estimated. All other sensitivity CACE model components will be identical to the primary CACE model.

h) Two sensitivity analyses will be performed to investigate the effect of missing data on the relationship between the intervention and the PANSS at six months post-randomisation. First, a “last-observation-carried-forward” approach will be used to impute individual items of the PANSS using the value of that participant-specific item value at baseline. Second, a MAR assumption by treatment arm will be assumed and imputed as described in section 5.11 (a).

i) The primary outcome analysis will be repeated excluding participants who have been randomised and included in subsequent study procedures and follow-up despite not meeting one of the inclusion criteria. Frequencies and percentages or participants excluded in this analysis will be presented in Table 18 Protocol Deviations. The treatment effect estimate for the sensitivity analysis will be presented in Table 10.

6. Other analyses, data summaries, and graphs

6.1 Other data summaries

- A dose-response relationship will be explored by assessing the relationship between the MANSA score at six months after randomisation with (i) participants' mean scores on the SCENE Adherence Scale, and (ii) the number of coaching sessions in the intervention arm. A within-group linear regression model will be fitted with MANSA score at six months after randomisation as the outcome, MANSA score at baseline, facilitator (i.e., social coach), and site (as random intercept) as model covariates, and (i) mean SCENE Adherence Scale score as the primary (continuous) independent variable, or (ii) number of coaching sessions held as the primary (ordinal) independent variable.

The number of social coaching sessions required for each participant to reach 'good compliance' (i.e., once coaches record a '2' on all eight item of the SCENE Adherence Scale reported in section 5.7) will be calculated and tabulated.

6.2. Subgroup analyses

Subgroup analyses will be used to explore whether the effectiveness of the intervention differs for participants with different characteristics. The analysis of the primary outcome will be repeated adding a fixed effect and an interaction term between treatment allocation and the subgroup of interest into the model. The following subgroups will be analysed:

- Gender – male vs female
- Age – under 35 vs over 34 years of age at baseline
- Severity – hospitalised in acute psychiatric wards in the six months before recruitment or not
- Location (at individual rather than site level) – urban/semi-rural/rural

Subgroup analysis will include all participants with complete outcome data and with complete data for the subgroup variable. The presence of an interaction will be tested using a likelihood ratio test comparing the subgroup analysis model, including the subgroup variable of interest by treatment interaction, and the model without the interaction term. For each subgroup analysis, we will report the numbers in each subgroup, summary statistics by subgroup, treatment effect estimates and 95% confidence intervals for each subgroup, and a p-value for the test of interaction.

6.3 Safety analyses

The total number (n) and percentage (%) of serious adverse events (SAEs) related to the SCENE intervention will be reported. Furthermore, the total number (n) and percentage (%) of SAEs, adverse events, adverse events leading to withdrawal, and the number of patients with at least one SAE will be reported by trial arm. Details on what constitutes a (serious) adverse event can be found in the study protocol.

6.4 Graphs

Line graphs for overall treatment effect

We will use box-and-whisker plots within violin plots to visualise between-arm differences in point estimates (i.e., group means), inter-quartile range limits, outliers, and continuous data distributions for the primary outcome measure at each post-randomisation assessment time-point.

Path diagrams for mediation analysis

A causal path diagram will be constructed that presents the natural direct, natural indirect/mediated, total effect estimates (and 95%CIs) of the intervention on the primary outcome.

Forest plot for subgroup analysis

We will construct a forest plot that presents the treatment effect estimate (and 95%CI) of both primary outcomes for each pre-specified subgroup.

Forest plot for four pandemic phases

A forest plot will be created from analysis 5.11 c).

6.5 Trial protocol modifications due to the COVID-19 pandemic

Unanticipated circumstances secondary to the COVID-19 pandemic, subsequent impacts on the trial, and required trial modifications will be transparently reported, according to the CONSERVE Guideline for Reporting Trial Protocols and Completed Trials Modified Due to the COVID-19 Pandemic (36).

Extenuating circumstances: The COVID-19 pandemic and its effect on trial data collection and data completeness.

Important modifications: In response to the COVID-19 pandemic, the UK Government instituted several social restrictions of varying severity from March 23rd, 2020 until approximately March, 2022. Due to SCENE's primary research questions relating to the relationship between the intervention, frequency of social contacts, and health-related quality of life, "important modifications" to the trial were warranted. This could also affect the trial's ability to perform the most appropriate statistical methods.

Impacts: Participant recruitment, intervention delivery, data collection, and outcome measures was/were altered.

Mitigation strategies: Throughout the COVID-19 pandemic, participant recruitment, intervention delivery, and data collection continued remotely via telecommunication or approved, web-based, teleconferencing. Additional training materials were prepared for coaches to enable the safe delivery of the trial intervention in accordance with social distancing measures. Additional outcome measures were added to monitor (i) participants' experience of COVID-19 and associated social distancing, (ii) subjective anxiety (GAD-7) and depression (PHQ-9) symptom severity, and (iii) remote social contacts and time spent socialising online. An additional assessment was introduced as a new time point for participants who were receiving the trial intervention during the initial COVID-19 lockdown (i.e., those randomised between the 17th of September 2019 and the 16th of March 2020). Sixty-six participants were excluded from the trial because, due to national lockdown measures and Tier 4 (i.e., 'Stay at home') restrictions during specific time-periods in the UK, it was not possible for these participants to meet more than three people in the week prior to their primary outcome measurement (a primary inclusion criterion in the SCENE trial was if participants' number of social contacts in the previous week is 3 or less). Changes were made to the statistical analysis plan to

incorporate analytical methods that account for the effect of the COVID-19 pandemic on missing data and the intervention effect estimate.

Modification timeline: The COVID-19 pandemic was identified as a potential extenuating circumstance in early February 2020 and started affecting the SCENE trial from March 23rd, 2020. In retrospect, chief investigators deem that the COVID-19 pandemic possessed the capacity to have affected the SCENE trial from March 23rd, 2020 (i.e., announcement of the first Government-enforced national lockdown in the UK) until February 24th, 2022 (i.e., official removal of all COVID-19-related restrictions). There were 280 participants who enrolled in SCENE before the COVID-19 pandemic (in March 2020) prompted modifications to the trial. After the COVID-19 pandemic necessitated modifications to SCENE, 296 participants enrolled in the trial.

7. References

1. Priebe S, Huxley P, Knight S, Evans S. Application and results of the Manchester Short Assessment of Quality of Life (MANSA). Int J Soc Psychiatry. 1999;45(1):7–12.
2. Priebe S, Watzke S, Hansson L, Burns T. Objective social outcomes index (SIX): a method to summarise objective indicators of social outcomes in mental health care. Acta Psychiatr Scand [Internet]. 2008 [cited 2023 Aug 21];118(1):57–63. Available from: https://onlinelibrary.wiley.com/doi/10.1111/j.1600-0447.2008.01217.x
3. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull. 1987;13(2):261–76.
4. Kay SR, Opler LA, Lindenmayer JP. The Positive and Negative Syndrome Scale (PANSS): rationale and standardisation. Br J Psychiatry Suppl. 1989 Nov;(7):59–67.
5. Hays RD, DiMatteo MR. A short-form measure of loneliness. J Pers Assess. 1987;51(1):69–81.
6. Priebe S, Pavlickova H, Eldridge S, Golden E, McCrone P, Ockenden N, et al. Effectiveness of one-to-one volunteer support for patients with psychosis: protocol of a randomised controlled trial. BMJ Open. 2016 Aug 3;6(8):e011582.
7. Giacco D, Palumbo C, Strappelli N, Catapano F, Priebe S. Social contacts and loneliness in people with psychotic and mood disorders. Comprehensive Psychiatry [Internet]. 2016 [cited 2023 Aug 21];66:59–66. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0010440X15300146
8. Herdman M, Gudex C, Lloyd A, Janssen Mf, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res [Internet]. 2011 [cited 2023 Aug 11];20(10):1727–36. Available from: http://link.springer.com/10.1007/s11136-011-9903-x
9. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy. 1990 Dec;16(3):199–208.
10. Hernández Alava M, Pudney S, Wailoo A. Estimating the Relationship Between EQ-5D-5L and EQ-5D-3L: Results from a UK Population Study. PharmacoEconomics [Internet]. 2023 [cited 2023 Aug 11];41(2):199–207. Available from: https://link.springer.com/10.1007/s40273-022-01218-7
11. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: Validity of a brief depression severity measure. J Gen Intern Med [Internet]. 2001 [cited 2023 Aug 11];16(9):606–13. Available from: http://link.springer.com/10.1046/j.1525-1497.2001.016009606.x
12. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med. 2006 May 22;166(10):1092–7.
13. Priebe S, Kelley L, Omer S, Golden E, Walsh S, Khanom H, et al. The Effectiveness of a Patient-Centred Assessment with a Solution-Focused Approach (DIALOG+) for Patients with Psychosis: A Pragmatic Cluster-Randomised Controlled Trial in Community Care. Psychother Psychosom. 2015;84(5):304–13.
14. Priebe S, Chevalier A, Hamborg T, Golden E, King M, Pistrang N. Effectiveness of a volunteer befriending programme for patients with schizophrenia: randomised controlled trial. Br J Psychiatry. 2020 Sep;217(3):477–83.
15. Roberts C, Roberts SA. Design and analysis of clinical trials with clustering effects due to treatment. Clin Trials. 2005;2(2):152–62.
16. Candlish J, Teare MD, Dimairo M, Flight L, Mandefield L, Walters SJ. Appropriate statistical methods for analysing partially nested randomised controlled trials with continuous outcomes: a simulation study. BMC Med Res Methodol. 2018 Oct 11;18(1):105.
17. Giesbrecht FG, Burns JC. Two-Stage Analysis Based on a Mixed Model: Large-Sample Asymptotic Theory and Small-Sample Simulation Results. Biometrics [Internet]. 1985 [cited

2023 Aug 11];41(2):477. Available from: https://www.jstor.org/stable/2530872?origin=crossref
18. Sullivan TR, White IR, Salter AB, Ryan P, Lee KJ. Should multiple imputation be the method of choice for handling missing data in randomized trials? Stat Methods Med Res. 2018 Sep;27(9):2610–26.
19. Krull JL, MacKinnon DP. Multilevel Modeling of Individual and Group Level Mediated Effects. Multivariate Behav Res. 2001 Apr 1;36(2):249–77.
20. MacKinnon DP. Introduction to Statistical Mediation Analysis [Internet]. 1st ed. Routledge; 2012 [cited 2023 Aug 11]. Available from: https://www.taylorfrancis.com/books/9781136676147
21. Goldsmith K, Hudson JL, Chalder T, Dennison L, Moss-Morris R. How and for whom does supportive adjustment to multiple sclerosis cognitive-behavioural therapy work? A mediated moderation analysis. Behav Res Ther. 2020 May;128:103594.
22. Goldsmith KA, MacKinnon DP, Chalder T, White PD, Sharpe M, Pickles A. Tutorial: The practical application of longitudinal structural equation mediation models in clinical trials. Psychol Methods. 2018 Jun;23(2):191–207.
23. Landau S, Emsley R, Dunn G. Beyond total treatment effects in randomised controlled trials: Baseline measurement of intermediate outcomes needed to reduce confounding in mediation investigations. Clinical Trials [Internet]. 2018 [cited 2023 Aug 11];15(3):247–56. Available from: http://journals.sagepub.com/doi/10.1177/1740774518760300
24. MacKinnon DP. Mediating Variable. In: International Encyclopedia of the Social & Behavioral Sciences [Internet]. Elsevier; 2001 [cited 2023 Aug 11]. p. 9503–7. Available from: https://linkinghub.elsevier.com/retrieve/pii/B0080430767007324
25. MacKinnon DP, Dwyer JH. Estimating Mediated Effects in Prevention Studies. Eval Rev [Internet]. 1993 [cited 2023 Aug 11];17(2):144–58. Available from: http://journals.sagepub.com/doi/10.1177/0193841X9301700202
26. Windgassen S, Goldsmith K, Moss-Morris R, Chalder T. Establishing how psychological therapies work: the importance of mediation analysis. J Ment Health. 2016;25(2):93–9.
27. Peugh JL, Strotman D, McGrady M, Rausch J, Kashikar-Zuck S. Beyond intent to treat (ITT): A complier average causal effect (CACE) estimation primer. J Sch Psychol. 2017 Feb;60:7–24.
28. Troncoso P, Morales-Gómez A. Estimating the complier average causal effect via a latent class approach using gsem. The Stata Journal [Internet]. 2022 [cited 2023 Aug 11];22(2):404–15. Available from: http://journals.sagepub.com/doi/10.1177/1536867X221106416
29. Schweig JD, Pane JF. Intention-to-treat analysis in partially nested randomized controlled trials with real-world complexity. International Journal of Research & Method in Education [Internet]. 2016 Jul 2 [cited 2023 Aug 21];39(3):268–86. Available from: https://www.tandfonline.com/doi/full/10.1080/1743727X.2016.1170800
30. White IR, Horton NJ, Carpenter J, Statistics RIMAS, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. BMJ [Internet]. 2011 Feb 7 [cited 2023 Aug 21];342(feb07 1):d40–d40. Available from: https://www.bmj.com/lookup/doi/10.1136/bmj.d40
31. Austin PC, White IR, Lee DS, van Buuren S. Missing Data in Clinical Research: A Tutorial on Multiple Imputation. Can J Cardiol. 2021 Sep;37(9):1322–31.
32. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. BMC Med Res Methodol. 2017 Dec 6;17(1):162.
33. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? Int J Methods Psychiatr Res. 2011 Mar;20(1):40–9.

34. Cro S, Morris TP, Kahan BC, Cornelius VR, Carpenter JR. A four-step strategy for handling missing outcome data in randomised trials affected by a pandemic. BMC Med Res Methodol. 2020 Aug 12;20(1):208.
35. Cro S, Morris TP, Kenward MG, Carpenter JR. Sensitivity analysis for clinical trials with missing continuous outcome data using controlled multiple imputation: A practical guide. Statistics in Medicine [Internet]. 2020 Sep 20 [cited 2023 Aug 11];39(21):2815–42. Available from: https://onlinelibrary.wiley.com/doi/10.1002/sim.8569
36. Orkin AM, Gill PJ, Ghersi D, Campbell L, Sugarman J, Emsley R, et al. Guidelines for Reporting Trial Protocols and Completed Trials Modified Due to the COVID-19 Pandemic and Other Extenuating Circumstances: The CONSERVE 2021 Statement. JAMA. 2021 Jul 20;326(3):257–65.
37. Sullivan O, Gershuny J. UKTUS; UK Time Diary StudyUnited Kingdom Time Use SurveyUnited Kingdom Time Use Survey, 2014-2015 [Internet]. UK Data Service; 2023 [cited 2023 Aug 22]. Available from: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=8128#1
38. Hodgekins J, French P, Birchwood M, Mugford M, Christopher R, Marshall M, et al. Comparing time use in individuals at different stages of psychosis and a non-clinical comparison group. Schizophrenia Research [Internet]. 2015 [cited 2023 Aug 22];161(2–3):188–93. Available from: https://linkinghub.elsevier.com/retrieve/pii/S092099641400721X

8. APPENDICES

Appendix A: Derived outcomes

a) MANchester Short Assessment of quality of life (MANSA)

MANSA comprises 16 questionnaire items, of which 12 are strictly subjective and scored 1 (worst outcome) to 7 (best outcome) (1). Items 4, 5, 9 and 10 are termed objective, have a binary response (i.e., "yes"/"no"), and are not used in the calculation of the overall MANSA score. The overall MANSA score is calculated as the mean of the 12 items score (i.e., 1 [i.e., "couldn't be worse"] – 7 [i.e., "couldn't be better"]). If up to 2 individual items are missing the MANSA score is the mean of the non-missing items. If more than 2 items are missing the score is set to missing.

Appendix table1. MANchester Short Assessment of quality of life (MANSA) items

Item number	Item question	Score
1	How satisfied are you with your life as a whole today?	Satisfaction scale (1-7)
2	How satisfied are you with your job (or sheltered employment, or training/education as your main occupation)? Or if unemployed or retired, how satisfied are you with being unemployed/retired?	Satisfaction scale (1-7)
3	How satisfied are you with your financial situation?	Satisfaction scale (1-7)
4	Do you have anyone who you would call a "close friend"?	Yes/No
5	In the last week, have you seen a friend (i.e., visited a friend, been visited by a friend, or met a friend outside both your home and work)?	Yes/No
6	How satisfied are you with the number and quality of your friendships?	Satisfaction scale (1-7)
7	How satisfied are you with your leisure activities?	Satisfaction scale (1-7)
8	How satisfied are you with your accommodation?	Satisfaction scale (1-7)
9	In the past year have you been accused of a crime?	Yes/No
10	In the past year have you been a victim of physical violence?	Yes/No
11	How satisfied are you with your personal safety?	Satisfaction scale (1-7)
12	How satisfied are you with the people that you live with? Or if you live alone, how satisfied are you with living alone?	Satisfaction scale (1-7)
13	How satisfied are you with your sex life?	Satisfaction scale (1-7)
14	How satisfied are your relationship with your family?	Satisfaction scale (1-7)
15	How satisfied are you with your health?	Satisfaction scale (1-7)
16	How satisfied are you with your mental health?	Satisfaction scale (1-7)

Appendix table 2. MANchester Short Assessment of quality of life (MANSA) satisfaction scale

Response description	Scale rating
Couldn't be worse	1
Displeased	2
Mostly dissatisfied	3
Mixed	4
Mostly satisfied	5
Pleased	6
Couldn't be better	7

b) Positive And Negative Symptom Scale (PANSS)

The Positive and Negative Syndrome Scale (PANSS) is a 30-item, seven-point, clinician-administered, symptom scale that is administered by a subject matter expert to identify, grade, and monitor symptoms in schizophrenia (3,4). It quantifies positive symptoms, which refer to an excess or distortion of normal functions (e.g., hallucinations and delusions), and negative symptoms, which represent a diminution or loss of normal functions. The PANSS comprises 3 subscales – Positive Scale, Negative Scale, and General Psychopathology Scale. Each subscale is rated with 1 (i.e., "absent") to 7 (i.e., "extreme") points. The range for the Positive and Negative Scales is 7-49, and the range for the General Psychopathology Scale is 16-112. Higher scores correspond to more severe illness. The total PANSS score is the sum of the sub-scales. In addition to these measures, a Composite Scale is scored by subtracting the negative score from the positive score. This yields a bipolar index that ranges from –42 to +42, which is essentially a difference score reflecting the degree of predominance of one syndrome relative to the other.

Appendix table 3. Positive And Negative Symptom Scale (PANSS) items

Positive Scale		**General Psychopathology Scale**	
P1	Delusions	G1	Somatic concern
P2	Conceptual disorganization	G2	Anxiety
P3	Hallucinatory behaviour	G3	Guilt feelings
P4	Excitement	G4	Tension
P5	Grandiosity	G5	Mannerisms & posturing
P6	Suspiciousness/persecution	G6	Depression
P7	Hostility	G7	Motor retardation
		G8	Uncooperativeness
Negative Scale		G9	Unusual thought content
N1	Blunted affect	G10	Disorientation
N2	Emotional withdrawal	G11	Poor attention
N3	Poor rapport	G12	Lack of judgment & insight
N4	Passive/apathetic social withdrawal	G13	Disturbance of volition
N5	Difficulty in abstract thinking	G14	Poor impulse control
N6	Lack of spontaneity & flow of conversation	G15	Preoccupation
N7	Stereotyped thinking	G16	Active social avoidance

Appendix table 4. Positive And Negative Symptom Scale (PANSS) rating scale

Symptom severity	**Symptom rating**
Absent	1
Minimal	2
Mild	3
Moderate	4
Moderate - severe	5
Severe	6
Extreme	7

There are eight "observational" items on the PANSS included in the Positive and Negative subscales. These items are considered observational because they are based on the clinician's observation of the patient's behaviour and presentation, rather than on the patient's self-report. These items are as follows:

Positive subscale

- Delusions
- Conceptual disorganization
- Hallucinations
- Excitement

Negative subscale

- Blunted affect
- Emotional withdrawal
- Poor rapport
- Passive/apathetic social withdrawal

When greater than 20% of scale items have missing values OR all observational items of the PANSS have missing values, the overall PANSS score will be set to missing. When ≤20% of scale items have missing values, missing values will be imputed using the mean value of the present item for this participant.

c) Social inclusion (as measured using the Objective Social Outcomes Index (SIX)

The Objective Social Outcomes Index (SIX) is a brief and easily administered ordinal measure of objective indicators of social outcomes in mental health (2). These objectives indicators are capture aspects of an individual's social situation and is reliably assessed by an independent rater. Objective indicators of social outcomes in mental health care capture aspects of the social situation that can be assessed by an independent observer. SIX contains four domains: (i) employment, (ii) accommodation, (iii) partnership/family, and (iv) friendship. An overall score is calculated by summing the four domains, resulting in a scale ranging from 0 to 6. A greater score reflects a more positive social outcome. SIX does not possess floor or ceiling effects. A difference of point on SIX is considered meaningful and represents a significant increase (i.e., more positive) or decrease (i.e., more negative) in social outcome.

Appendix table 5. Objective Social Outcome Index (SIX)

SIX domains	Domain categories	Scoring from CRF
Employment (CRF 2.4)	None (0) Voluntary/protected/sheltered work (1) Regular employment (2)	0 if "5", "8" 1 if "3", "4" 2 if "1", "2", "6", "7"
Accommodation (CRF 2.2a)	Homeless or 24-hour supervision (0) Sheltered or supported accommodation (1) Independent accommodation (2)	0 if "3" 1 if "2" 2 If "1"
Partnership/family (CRF 2.3?)	Living alone (0) Living with a partner or family (1)	0 if "1" 1 if "2", "3", "4"
Friendship (MANSA Q05)	Not meeting a friend within the last week (0) Meeting at least one friend in the last week (1)	0 if MANSA Q05 answer is No 1 if answer is Yes

CRF, case report form

d) UCLA-8 Loneliness Scale

The UCLA-8 Loneliness Scale is a short-form, 8-item version of the original 20-item scale that assesses feelings of loneliness and social isolation in adults. It comprises a 4-point rating scale (i.e., never [1], rarely [2], sometimes [3], always [4]), where higher scores correspond to a greater sense

of loneliness. Items three and six (i.e., positively worded items) are reverse-coded so that higher values (i.e., never = 4; always = 1) correspond to more loneliness. To score the scale, responses to the 8 items are summed. The total score for each participant ranges from 8 to 32, with higher scores indicating greater feelings of loneliness and social isolation.

Appendix table 6. UCLA-8 Loneliness Scale

Number	Item	Rating Scale
1	I lack companionship	Never (1), Rarely (2), Sometimes (3), Always (4)
2	There is no one I can turn to	Never (1), Rarely (2), Sometimes (3), Always (4)
3	I am an outgoing person*	**Never (4), Rarely (3), Sometimes (2), Always (1)**
4	I feel left out	Never (1), Rarely (2), Sometimes (3), Always (4)
5	I feel isolation from others	Never (1), Rarely (2), Sometimes (3), Always (4)
6	I can find companionship when I want it*	**Never (4), Rarely (3), Sometimes (2), Always (1)**
7	I am unhappy being so withdrawn	Never (1), Rarely (2), Sometimes (3), Always (4)
8	People are around me but not with me	Never (1), Rarely (2), Sometimes (3), Always (4)

*Items 3 and 6 are reverse-coded

e) Time Spent in Social Activities (using the Time Use Survey)

The Time Use Survey (TUS) questionnaire is a standardized survey instrument used to collect detailed information about how individuals allocate their time during a defined period (37). The TUS covers a wide range of activities including work, household chores, leisure, and sports activities. The focus of the SCENE trial concerns survey items reflecting time spent in social activities, which includes 'spare time' and 'sports activities' domains. The data collected from the TUS can be used to analyse patterns and trends in time use, as well as factors that influence time spent in social activities, and provides a useful way to assess social disability (38). Lists of activities are provided for each category (e.g., leisure activities include going to the cinema, pub, eating out, etc.). Participants are asked how many times they had engaged in each activity over the past week, for how long, whether they did this with someone and, if so, what type of relationship they have with this person.

The TUS score of time (in minutes) spent in social activities is calculated by summing the stated weekly totals of all leisure/spare time and sports activities categories (ignoring the number of times an activity has been carried out). Activities without values are counted as 'zero' rather than missing.

In May 2020, the online social activities were added the TUS, which captures online social activities in trial participants who were not interviewed face-to-face in follow-up assessments.

f) Social Contacts Assessment (SCA)

The Social Contacts Assessment (SCA) questionnaire is a patient-reported measure used to assess the quality and quantity of an individual's social contacts (7). It is designed to assess an individual's social network by capturing information on the frequency and types (e.g., face-to-face, telecommunication, etc.) of social interactions with friends and acquaintances. Participants list (without providing names) the people with whom they have been in contact (i.e., a chat that involves more than just a greeting) during the previous week (i.e., in the last seven days). First-degree relative (i.e., parents, siblings, children), co-habitants, healthcare professionals, and work colleagues (unless the contact took place outside of work) are excluded because these contacts are not always considered social relations. The quantitative value of interest comprises the total number of individuals with whom the participant has had at least one face-to-face contact in the last seven days. The value of SCA is the response to the question "**How many people have you had face-to-face contact with during the past week?**". During data cleaning the sum of the number of rows

completed in the questionnaire is calculated and compared against the SCA value. Discrepancies will be assessed on an individual basis and documented in the data cleaning appendix.

Appendix table 7. Social Contacts Assessment questionnaire

Domains	Response (categories, where applicable)
List of contacts	Initials of contacts
Type of relationship	Friend (1), Partner (2), Acquaintance (3), Other, specify (4)
On how many days, in the last week, have you been in face-to-face contact with him/her?	
Was the meeting one to one or in a group?	(a) One to one (b) Group (c) Both
If both, on how many days did you have one to one meeting(s)?	
On how many days, in the last week, have you been in contact by voice or video call (using phone, skype or facetime, etc.)?	
On how many days, in the last week, have you been in contact by social networking, e-mail or text message?	
Can you talk to him/her about your personal feelings and worries?	(a) Yes (b) No
Did you do something for him/her in the last week? If yes, what?	(c) Yes (d) No
Did he/she do something for you in the last week? If yes, what?	(a) Yes (b) No
How many people have you had face-to-face contact with during the past week?	
On how many days in the previous week have you had a face-to-face contact?	

In May 2020, the Remote SCA questionnaire was added, which captures remote social contacts in trial participants who were not interviewed face-to-face in follow-up assessments. Remote contacts identified using the Remote SCA questionnaire were accepted as eligible social contacts from this point in the study. Participants list (without providing names) the people with whom they have had a remote social contact (i.e., a two-way interaction involving a conversation or message exchange with a person they can name) during the previous week (i.e., in the last seven days). First-degree relative (i.e., parents, siblings, children), co-habitants, healthcare professionals, and work colleagues (unless the contact took place outside of work) are excluded because these contacts are not always considered social relations. The value of SCA is the response to the question "**How many people have you had remote contact with during the past week?**".

Appendix table 8. Remote Social Contacts Assessment questionnaire

Domains	Response (categories, where applicable)
List of contacts	Initials of contacts
Type of relationship	Friend (1), Partner (2), Acquaintance (3), Other, specify (4)
On how many days, in the last week, have you been in contact by messaging or email? If you have, which platforms?	
On how many days, in the last week, have you been in contact by video call? If you have, which platforms?	
On how many days, in the last week, have you been in contact by telephone call?	
Were these interactions one-to-one or in a group?	(a) One-to-one (b) Group (c) Both
Who initiated contact first?	(a) Me (b) Them (c) Both of us
Can you talk to him/her about your personal feelings and worries?	(a) Yes (b) No
In the last week, have you met this person face-to-face?	(a) Yes (b) No
How many people have you had a remote contact with during the past week?	
On how many days in the previous week have you had a remote contact?	

g) PHQ-9 score for depression

The Patient Health Questionnaire-9 (PHQ-9) is scored from a 9-item questionnaire assessing depression symptom severity (11). The total score is the sum of nine items and ranges from 0-27. For each item, responses range from 0-3 (i.e., not at all [0]; several days [1]; more than half the days [2]; nearly every day [3]). A score of 0-4 indicates no depression, 5-9 indicates mild depression, 10-14 indicates moderate depression, 15-19 indicates moderately severe depression, and 20-27 indicates severe depression.

h) GAD-7 score for anxiety

The Generalised Anxiety Disorder-7 (GAD-7) score is derived by summing the coded responses to 7 questions assessing symptoms and behaviours characterising generalised anxiety (12). For each item, responses range from 0-3 (i.e., not at all [0]; several days [1]; more than half the days [2]; nearly every day [3]). The score can range from 0-21. A score of 0-4 indicates no anxiety, 5-9 indicates mild anxiety, 10-14 indicates moderate anxiety, and 15-21 indicates severe anxiety.

i) **Health Related Quality of Life (HRQoL) using the EQ-5D-5L questionnaire**

The EQ-5D-5L questionnaire assesses participants' health-related quality of life (8). The EQ-5D-5L comprises 5 dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression, each rated on a scale from 1 to 5, corresponding to no problems (1), slight problems (2), moderate problems (3), severe problems (4), and extreme problems (5). Overall QoL utility scores will be derived for all contributing study participants using the UK National Institute for Health and Care Excellence (NICE) decision support unit EQ-5D scoring algorithm (15). Briefly, this estimation algorithm (i.e., the eq5dmap Stata command) will directly map from individual-specific, EQ-5D-5L, health states to individual-specific, EQ-5D-3L, utility scores, using age and sex as necessary covariates (10). Estimated, individual-specific, EQ-5D-3L utility scores will be used as the secondary outcome during statistical analysis.

The overall score of the EQ-5D-3L index ranges from -0.594 to 1.000 (i.e., Higher scores correspond to a better quality of life). A score of -0.594 represents the worst possible health status while a score of 1.000 represents the best possible health status. A score of 0.000 indicates a health status that is considered as bad as being dead (in terms of quality of life). The absolute minimum score of -0.594 indicates that an individual's health status is worse than being dead because an individual of such health status is not only experiencing significant health problems but is are also experiencing a lower quality of life compared to someone who is deceased. Due to the mapping from 5L to 3L the boundary values cannot be reached and the actual range of possible values is slightly smaller.

The EQ-VAS is a patient-reported measure of perceived overall health. It is a continuous measure that ranges from 0-100, with 100 indicating "the best health imaginable" and 0 indicating "the worst health imaginable." This score requires no further derivation.

Appendix B: Stata code for primary outcome, mediation, and CACE analyses.

```
* ************************ *
* primary outcome analysis *
* ************************ *

** analysis strategy (i - iv) if any outcome analysis fails to
converge when assessing between-group differences

* (i) primary analysis

mixed y_6m treat y_B || site: || ///
      therapist:treat, nocons reml ///
      residuals(independent, by(treat)) ///
      dfmethod(sat)


* (ii) remove clustering of patients by clinician-coaches

mixed y_6m treat y_B || site ///
      , nocons reml ///
      residuals(independent, by(treat)) ///
      dfmethod(sat)


* (iii) fit stratification factor - trial site - as fixed effect

regress y_6m treat y_B


* (iv) remove covariate "site" from the model

regress y_6m treat
```

```
* ****************** *
* mediation analysis *
* ****************** *

** analysis strategy

* (i) multi-level structural-only sem

gsem (MANSA_6m <- SCAtot_6m treat M1[site]) ///
     (SCAtot_6m <- treat M2[site]) ///
     cov(M1[site]*M2[site]@0)

gsem, coeflegend // to observe coefficient labels

nlcom _b[MANSAm_6m:SCAtot_6m]*_b[SCAtot_6m:treat] // to obtain
indirect effect coefficient, SE, & 95%CI

nlcom
_b[MANSAm_6m:treat]+_b[MANSAm_6m:SCAtot_6m]*_b[SCAtot_6m:treat] //
to obtain total effect coefficient, SE, & 95%CI


* (ii) one-level structural-only sem

gsem (MANSAm_6m <- SCAtot_6m treat) (SCAtot_6m <- treat)

gsem, coeflegend // to observe coefficient labels

nlcom _b[MANSAm_6m:SCAtot_6m]*_b[SCAtot_6m:treat] // to obtain
indirect effect coefficient, SE, & 95%CI

nlcom
_b[MANSAm_6m:treat]+_b[MANSAm_6m:SCAtot_6m]*_b[SCAtot_6m:treat] //
to obtain total effect coefficient, SE, & 95%CI


* alternative one-level coding approach

sem (MANSAm_6m <- SCAtot_6m treat) (SCAtot_6m <- treat)

estat teffects // to obtain direct, indirect, & total effects
```

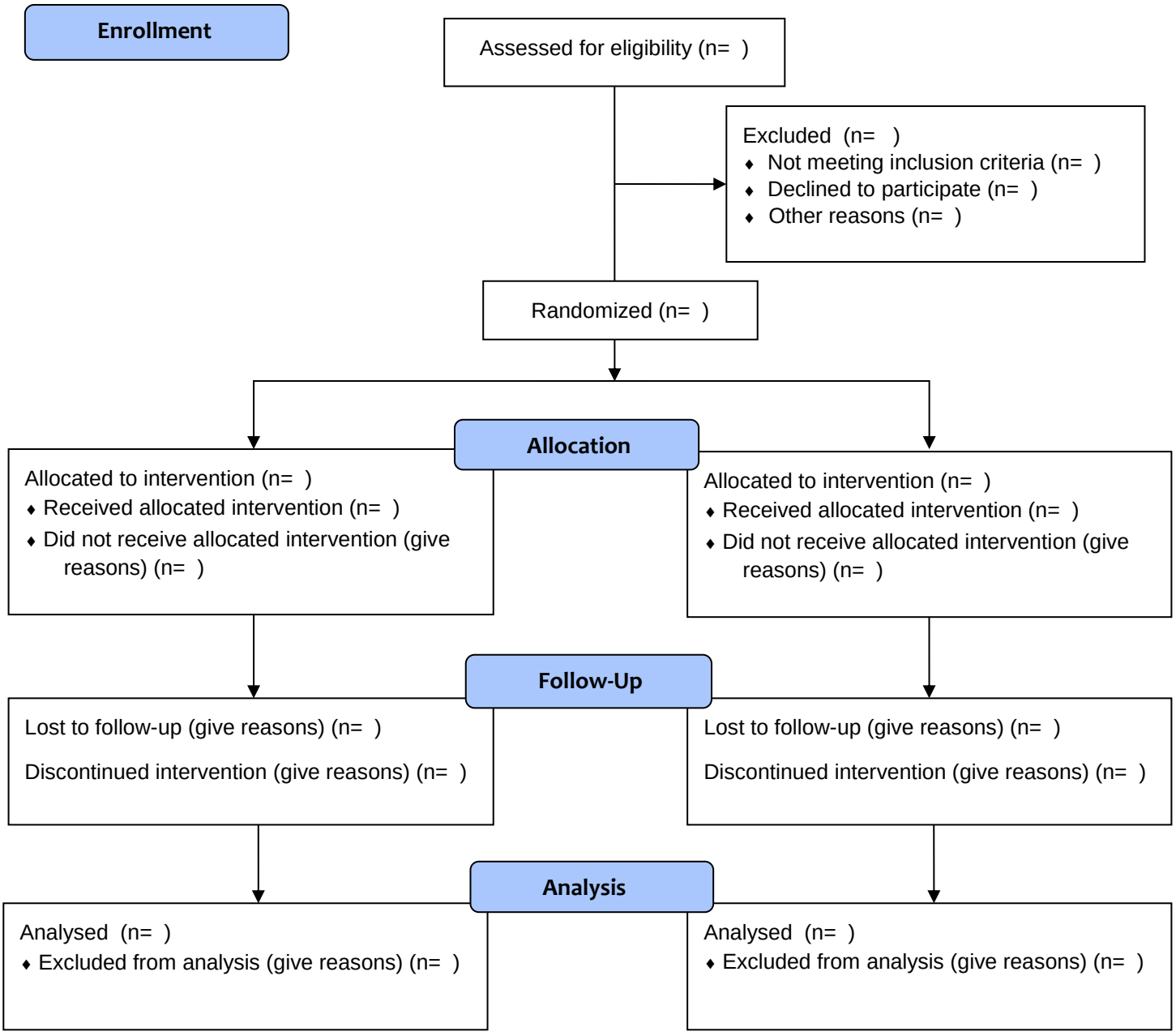
```
* ********************************************** *
* complier average causal effects (CACE) analysis *
* ********************************************** *

* generate compliance indicator variable

gen comp = c if treat==1 // compliance data is missing in control
group


** specify CACE model
// latent class regression model with specific constraints necessary
for CACE estimation

* convert standard latent class model into CACE model using the
following steps:

* step 1: extend regression model for MANSA at six months into two
paths
* step 2: fix the effect of the intervention in the non-compliers
class to zero (ie, specify the exclusion restriction assumption)
* step 3: extend the latent class model for compliance into two
paths to treat observed compliance in the treatment arms as known

gsem (1.C: MANSAm_6m <- i.treat@0 /// to specify regresison path for
non-compliers, with treatment effect fixed to zero
     MANSAm_base@C1 site@C2, /// to constrain the effects of
covariates in the regresison equations to be equal across classes
     vce(cluster site) /// to estimate robust standard errors
     (2.C: MANSAm_6m <- i.treat /// to specify regression path for
compliers, with treatment effect estimated freely (ie, CACE
estimate)
     MANSAm_base@C1 site@C2, /// to constrain the effects of
covariates in the regression equations to be equal across classes
     vce(cluster site)) /// to estimate robust standard errors
     (1.C: comp <- _cons@-15, logit) /// to specify the path for
non-compliers (comp=0) in the intervention arm
     (2.C: comp <- _cons@15, logit) /// to specify the path for
compliers (comp=1) in the intervention arm
     (C <- x1 x2 x3 x4 x5 x6), /// predicting the latent class of
compliance for individuals in the control arm (ie, predicting
"would-be" compliers in the control arm)
     lclass(C 2) /// to assign the name of the latent class
predicted (ie, C) and the number of classes (ie, 2 for [i] non-
compliers & [ii] compliers)
     nolog



* obtain summary of model fit for comparing competing models:
```

```
estat ic //


* obtain predicted values of MANSA at six months amongst non-
compliers (class 1) at the quartiles of MANSA values at baseline:

margins, at((p25) MANSAm_base) at((p50) MANSAm_base) at((p75)
MANSAm_base)

predict(outcome(MANSAm_6m) class(1))


* visualise predicted values in the noncompliers class (class 1):

marginsplot, title("Noncompliers (overall)") ///
      xtitle("Predicted MANSA score at six months") ///
      ytitle("MANSA score at baseline") ///
      recast(scatter) ///
      ylabel(1 "" 2 "" 3 "") /// INSERT CATEGORY NUMBERS FOR Y-AXIS
      xlabel(-1 (0.2) 0.5) /// INSERT MIN & MAX LABELS FOR X-AXIS
      plotopts(msymbol(Oh)) ///
      horizontal xline(0, lpattern(dash)) ///
      scheme(sj)


* obtain predicted values of MANSA at six months amongst compliers
(class 2) at the quartiles of MANSA values at baseline:

margins treat, at((p25) MANSAm_base) at((p50) MANSAm_base) at((p75)
MANSAm_base)

predict(outcome(MANSAm_6m) class(2))


* sensitivity analysis relaxing exclusion restriction assumption

gsem (1.C: MANSAm_6m <- i.treat /// to specify regression path for
non-compliers, with treatment effect estimated freely
      MANSAm_base@C1 site@C2, /// to constrain the effects of
covariates in the regression equations to be equal across classes
      vce(cluster site) /// to estimate robust standard errors
      (2.C: MANSAm_6m <- i.treat /// to specify regression path for
compliers, with treatment effect estimated freely (ie, CACE
estimate)
      MANSAm_base@C1 site@C2, /// to constrain the effects of
covariates in the regression equations to be equal across classes
      vce(cluster "site")) /// to estimate robust standard errors
      (1.C: comp <- _cons@-15, logit) /// to specify the path for
non-compliers (comp=0) in the intervention arm
```

```
     (2.C: comp <- _cons@15, logit) /// to specify the path for
compliers (comp=1) in the intervention arm
     (C <-x1 x2 x3 x4 x5 x6), /// predicting the latent class of
compliance for individuals in the control arm (ie, predicting
"would-be" compliers in the control arm)
     lclass(C 2) /// to assign the name of the latent class
predicted (ie, C) and the number of classes (ie, 2 for [i] non-
compliers & [ii] compliers)
     nolog
```

Appendix C: CONSORT diagram

CONSORT 2010 Flow Diagram



Appendix D: Draft tables

Table1 – Baseline characteristics

Summaries are mean (SD) unless stated otherwise.

	Summary measure		Complete data (No. (%))	
	Intervention	Control	Intervention (N =)	Control (N =)
Baseline demographics				
Age (years)				
Gender - n(%)				
Female				
Male				
Non-binary				
Prefer not to say				
Ethnicity - n(%)				
Asian/Asian British				
Black/African/Caribbean/Black				
White British				
Mixed/Multiple Ethnic Groups				
Any other White background				
Other				
Recruited from (%)				
Primary care				
Secondary care				
Number of hospitalisations				
Main Diagnostic group (%)				
F20-29				
Other eligible disorders				
Time since first contact with MH service (years)				
Informal care (% yes)				
Level of Education – n(%)				
Primary				
Secondary				
Tertiary				
Other general				
Accommodation n(%)				
Independent				
Supported (staffed)				
Supported (unstaffed)				
Homeless				
Other				
Living situation n(%)				
Living alone				
Living with partner/family				
Living with friends				
Living shared				
Employment				

Paid full-time Paid part-time Volunteering Sheltered employment Unemployed Student Housewife/husband Retired Other		
Receive state benefits (% yes)		

Table 2 – Results for analysis of primary and secondary outcomes at 6 months after randomisation

	Number included in analysis		**Summary measure**				
	Intervention n (%)	Usual Care n (%)	Intervention Mean (SD)	Usual Care Mean (SD)	Treatment effect	(95% CI)	p-value
MANSA							
Social Contacts Assessment *							
PANSS							
Social situation (SIX)							
UCLA-8 Loneliness Scale							
Time spent in social activities							
EQ-5D-5L Utility							
EQ-5D-5L VAS							

Table 3 – Results for analysis of primary and secondary outcomes at 12 months after randomisation

	Number included in analysis		**Summary measure**				
	Intervention n (%)	Usual Care n (%)	Intervention Mean (SD)	Usual Care Mean (SD)	Treatment effect	(95% CI)	p-value
MANSA							
Social Contacts Assessment *							
PANSS							
Social situation							
UCLA-8 Loneliness Scale							
Time spent in social activities							
EQ-5D-5L Utility							
EQ-5D-5L VAS							

Table 4 - Main results for analysis of primary and secondary outcomes at 18 months after randomisation

	Number included in analysis		Summary measure				
	Intervention n (%)	Usual Care n (%)	Intervention Mean (SD)	Usual Care Mean (SD)	Treatment effect	(95% CI)	p-value
MANSA							
Social Contacts Assessment *							
PANSS							
Social situation							
UCLA-8 Loneliness Scale							
Time spent in social activities							
EQ-5D-5L Utility							
EQ-5D-5L VAS							

Table 5. Results of mediation analysis investigating whether the effect of the intervention of MANSA at 6 and 12 after randomisation is mediated through expanded social networks at six months after randomisation

Effect type	IV → DV	Effect estimate	Std. error	95%CI	P-value
MANSA at 6 months after randomisation					
Direct effect	Intervention → MANSA				
Indirect effect	Intervention → SCA6 → MANSA				
Total effect	Intervention → MANSA				
MANSA at 12 months after randomisation					
Direct effect	Intervention → MANSA				
Indirect effect	Intervention → SCA6 → MANSA				
Total effect	Intervention → MANSA				

SCA6, Social Contact Assessment at six months after randomisation; *MANSA,* Manchester Short Assessment of quality of life

Table 6 – Results of Complier-Average Causal Effect analysis investigating the effect of the intervention on MANSA at six months after randomisation amongst compliers*

Estimator	N	Effect estimate (β)	Std error	95%CI	p-value
Unadjusted					
ITT					
CACE					
Adjusted					
ITT					
CACE					

95%CI, 95% confidence interval; ITT, intention-to-treat; CACE, complier-average causal effect; Adjusted, adjusted for MANSA at baseline

*Participants in the intervention arm classified as "compliant" if they underwent an initial 60-minute session with a clinician AND at least one 20-minute review AND a final session.

Table 7 – Results for subgroup analysis of primary outcome

	Number included in analysis		MANSA at 6 months post-randomisation				
	Intervention n (%)	Usual Care no. (%)	Intervention mean (SD)	Usual Care mean (SD)	Treatment effect	(95% CI)	p-value for interaction
Gender							
Male							n/a
Female							n/a
Age at baseline							
<35 years							n/a
≥35 years							n/a
Severity*							
Hospitalised							n/a
Not hospitalised							n/a
Location							
Urban							n/a
Semi-rural							n/a
Rural							

*severity six months before recruitment defined as hospitalised (or not) in acute psychiatric ward

Table 8. Frequency (n) of adverse events and serious adverse events

	Intervention arm (n=)	Control arm (n=)
AEs		
Participants experiencing AEs		
SAEs		
Total number of SAEs		
Number of participants experiencing one or more SAEs		
Unexpected SAEs that are related to the intervention		
Type of SAEs		
Death		
Life-threatening complication		
Admission or prolongation of hospitalisation (for mental health)		
Admission or prolongation of hospitalisation (for other condition)		
Significant disability or incapacity		
"Other" important medical event		

*column-wise percentages

AEs, adverse events; *SAEs*, Serious Adverse Events

Table 9. Results of analysis investigating dose-response relationship between the MANSA score at six months after randomisation with (i) participants' mean scores on the SCENE Adherence Scale, and (ii) the number of coaching sessions in the intervention arm*

Explanatory variable	Regression coefficient (β)	95%CI	P-value
SCENE Adherence Scale			
Number of coaching sessions			

*adjusting for MANSA score at baseline, facilitator (i.e., social coach), and site (as random intercept)

Table 10. Sensitivity analysis investigating different approaches, assumptions, and inclusion criteria on the effect of intervention on MANSA six months after randomisation

Assumption	Treatment effect	95%CI	P-value
Complete case analysis			
Complete case actual end of treat *			
Randomised in error excluded			
Not accounting for COVID-19 pandemic (MAR by trial arm)			
Accounting for COVID-19 pandemic (controlled MI)			

95%CI, 95% confidence interval; *MAR,* missing at random; *MI*, multiple imputation

*replaced 6m value with end-of-treatment due to COVID-related trial disruption (approx. 10m)

Table 11. Results of sensitivity analysis investigating the effect of the intervention multiple imputation (i) not accounting, and (ii) accounting for the COVID-19 pandemic

Pandemic phase	Number analysed N*	Regression coefficient (β)	Std error	95%CI	P-value
Pre-pandemic					
Peri-pandemic – heavy restrictions					
Peri-pandemic – some restrictions					
Post-pandemic					
Pooled, aggregate effect					

Table 12. Sensitivity analysis investigating the mediating effect of the number of social contacts at six months after randomisation in the relationship between the intervention and MANSA at six months after randomisation when face-to-face and remote (i.e., online) social contacts are combined

Social contact type	Effect type	IV → DV	Effect estimate	95%CI	P-value
Face-to-face social contacts only	Direct effect	Intervention → MANSA			
	Indirect effect	Intervention → SCA6 → MANSA			
	Total effect	Intervention → MANSA			
Face-to-face AND remote social contacts combined	Direct effect	Intervention → MANSA			
	Indirect effect	Intervention → SCA6 → MANSA			
	Total effect	Intervention → MANSA			

95%CI, 95% confidence interval; *SCA6*, social contacts assessment at six months post-randomisation

Table 13. Sensitivity analysis investigating the effect of the intervention on time spent in social activities and online social activities combined

	Treatment effect	95%CI	P-value
Main analysis			
Sensitivity analysis			

95%CI, 95% Confidence Interval

Table 14. Sensitivity analysis investigating the robustness of CACE analysis* results to the exclusion restriction assumption (that the treatment effect is zero for non-compliers)

	Treatment effect	95%CI	P-value
Main analysis			
Sensitivity analysis			

**effect of intervention amongst 'compliers' when exclusion criterion does/does not apply*

95%CI, 95% Confidence Interval

Table 15. Sensitivity analysis investigating the effect of the intervention on PANSS scores six months after randomisation when missing individual items are imputed

	Treatment effect	95%CI	P-value
Main analysis			
Sensitivity analysis			

95%CI, 95% Confidence Interval; *CCA*, complete case analysis; *LOCF*, last observation carried forward; *MAR*, missing at random

Table 16: Protocol deviations

	Intervention arm	Control arm
Total number of protocol deviations		
Number of participants with at least one protocol deviation		
Randomised despite inclusion criterion MANSA ≤ 5.0 not met		
Other inclusion/exclusion criteria violation		
Any other protocol deviation		