Statistical Analysis Plan

Translated and culturally adapted internet-delivered cognitive therapy for social anxiety disorder in Japanese clinical settings

Version 1.0

Protocol reference	Yoshinaga N, Thew GR, Hayashi Y, Tanoue H, Nakai M, Clark DM. Trials. 2024;25:492. doi: 10.1186/s13063-024-08303-6

Revision history

Version	Date	Summary of changes
1.0	10 July 2026	Initial SAP version before database lock.

Approvals

Role	Name	Signature	Date
Principal investigator	Naoki Yoshinaga		10 July, 2026
Trial statistician	Graham Thew		10 July, 2026

1. Purpose and scope

This Statistical Analysis Plan (SAP) specifies the planned analyses for the randomised comparison of the Japanese version of internet-delivered cognitive therapy for social anxiety disorder (iCT-SAD) combined with treatment as usual (TAU) versus TAU alone. The target condition is social anxiety disorder (SAD). The SAP is intended to be approved before database lock and before any comparative outcome analyses are undertaken. Any deviations from the final SAP will be documented and justified in the trial report.

2. Abbreviations

Abbreviation	Term
ADIS-5	Anxiety and Related Disorders Interview Schedule for Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
AE	Adverse event
CI	Confidence interval
FAS	Full Analysis Set
FU	Follow-up
GAD-7	Generalised Anxiety Disorder Questionnaire, 7-item version
iCT-SAD	Internet-delivered cognitive therapy for social anxiety disorder
LSAS	Liebowitz Social Anxiety Scale, self-report version
PHQ-9	Patient Health Questionnaire, 9-item version
PPS	Per Protocol Set
RCT	Randomised controlled trial
SAD	Social anxiety disorder
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SAQ	Social Attitudes Questionnaire
SBQ	Social Behaviour Questionnaire
SCQ	Social Cognitions Questionnaire
SPWSS	Social Phobia Weekly Summary Scale
TAU	Treatment as usual
WSAS	Work and Social Adjustment Scale

3. Trial objectives

Primary objective: To examine whether a Japanese version of iCT-SAD in combination with treatment as usual (iCT-SAD + TAU) is superior to TAU alone in reducing social anxiety symptoms at the post-treatment endpoint.

Secondary and exploratory objectives:

To examine whether the Japanese iCT-SAD programme is acceptable to clients; to evaluate social anxiety processes, general mood, functioning, response, remission, deterioration, loss of SAD diagnosis, and safety; to evaluate how the outcomes of the iCT-SAD + TAU intervention compare to previous studies of the iCT-SAD programme in the UK and Hong Kong; to examine whether baseline clinical and demographic characteristics are associated with clinical outcomes in the iCT-SAD + TAU arm; to examine candidate mediators of the relationship between treatment arm and clinical outcome.

4. Trial design and assessment schedule

This is a two-arm, parallel-group, superiority randomised controlled trial (RCT). Participants are randomised in a 1:1 ratio to iCT-SAD + TAU or TAU alone. Participants and therapists are not blinded to allocation. An independent assessor conducting the post-treatment diagnostic assessment is blinded to allocation.

Randomised phase: Assessments are conducted at baseline (week 0; Pre), mid-treatment/mid-observation (week 8; Mid), and post-treatment/post-observation (week 15; Post). The primary endpoint is week 15.

Follow-up and post-control intervention phase: Participants allocated to iCT-SAD + TAU have a 3-month follow-up (FU) assessment at week 27. After week 15, participants initially allocated to TAU alone may receive iCT-SAD. For these participants, post-control intervention assessments are defined relative to the start of iCT-SAD: Pre (post-control; randomised week 15), Mid (post-control; randomised week 23), Post (post-control; randomised week 30), and FU (post-control; randomised week 42). Post-control intervention analyses are not randomised between-group comparisons.

5. Analysis populations

The following analysis populations will be defined before database lock and before unblinded comparative analyses are undertaken.

- **Full Analysis Set (FAS):** The FAS will include all randomised participants with valid consent and baseline eligibility, analysed according to the group to which they were randomised. The FAS will be the primary population for efficacy analyses, including the primary randomised comparison.
- **Per Protocol Set (PPS):** The PPS will include participants in the FAS without major protocol deviations that are judged to affect the evaluation of the primary outcome. Major protocol deviations and any minimum treatment exposure or assessment requirements for inclusion in the PPS will be defined before database lock. PPS analyses will be used as sensitivity analyses and will not replace the FAS as the primary analysis population.
- **Safety Analysis Set:** The Safety Analysis Set will include all randomised participants who start their allocated post-randomisation management or for whom post-randomisation safety information is available. Safety analyses will be conducted according to treatment received where this differs from randomised allocation.

Participants originally allocated to TAU alone who subsequently receive iCT-SAD after the week-15 assessment will form a post-control intervention subgroup. This subgroup will not be treated as a randomised comparison population; analyses of these data will be descriptive and exploratory.

6. Outcomes

6.1 Primary outcome

The primary outcome is the total score on the self-report version of the Liebowitz Social Anxiety Scale (LSAS) at week 15.

6.2 Continuous secondary outcomes

- **Social anxiety symptoms and processes:** Social Phobia Weekly Summary Scale (SPWSS) total score, Social Cognitions Questionnaire (SCQ) total score, Social Behaviour Questionnaire (SBQ) total score, and Social Attitudes Questionnaire (SAQ) total score.
- **General mood and functioning:** Patient Health Questionnaire, 9-item version (PHQ-9) total score; Generalised Anxiety Disorder Questionnaire, 7-item version (GAD-7) total score; and Work and Social Adjustment Scale (WSAS) total score.
- **Additional descriptive outcome:** monthly income, where collected at relevant assessment points.

6.3 Categorical secondary outcomes

- **Reliable response:** improvement of more than 31% in LSAS score from Pre to Post. The same criterion will be applied to Pre-to-FU comparisons in iCT-SAD recipients.
- **Remission:** a decrease of at least 12 LSAS points from Pre to Post and a Post LSAS score of 38 or lower. The same criterion will be applied to Pre-to-FU comparisons in iCT-SAD recipients.
- **Reliable deterioration:** an increase of at least 12 LSAS points from Pre to Post. The same criterion will be applied to Pre-to-FU comparisons in iCT-SAD recipients.

- **Loss of SAD diagnosis:** no longer meeting Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition criteria for SAD on the Anxiety and Related Disorders Interview Schedule for Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (ADIS-5) at Post and, for iCT-SAD recipients, at FU where assessed.

6.4 Acceptability, safety, engagement, and intervention delivery

Acceptability and safety: Dropout, adverse events (AEs), and serious adverse events (SAEs) will be summarised.

Engagement and intervention delivery among iCT-SAD recipients: Engagement, delivery, and participant feedback measures will be summarised among participants receiving iCT-SAD, separately for the randomised iCT-SAD + TAU arm and the post-control intervention subgroup where applicable. These variables may include module completion, behavioural experiments completed, programme use, and therapist contact, as available from the treatment platform and trial records. Treatment satisfaction ratings will also be summarised.

7. General analysis principles

- The primary randomised comparison will be conducted using the FAS.
- The PPS will be used for sensitivity analyses of the primary outcome where appropriate.
- All statistical tests will be two-sided. The nominal significance level for the primary outcome will be 0.05.
- Secondary and exploratory analyses will be interpreted cautiously. No formal multiplicity adjustment is planned.
- For continuous outcomes, adjusted mean differences, standard errors, 95% confidence intervals (CIs), p-values, and effect sizes will be reported where appropriate.
- For categorical outcomes, counts, percentages, risk differences, odds ratios, 95% CIs where estimable, and Fisher exact test p-values will be reported.
- For all scales, the direction of scoring will be stated in the results tables.

8. Baseline characteristics and concomitant care

Prespecified baseline demographic and clinical characteristics, including demographic variables, SAD history, comorbidity, medication use, previous pharmacological treatment, and treatment-resistant status, will be summarised by randomised arm using mean and standard deviation or median and interquartile range for continuous variables, and frequency and percentage for categorical variables. No significance testing of baseline differences is planned. Standardised mean differences may also be presented descriptively.

Concomitant care from week 0 to week 15, including visits to the primary psychiatrist, psychotropic medication use and dose equivalents, and medication regimen changes at Mid and Post, will be summarised by randomised arm.

Therapist characteristics, including primary profession, gender, age, post-qualification clinical experience, and experience in delivering cognitive and behavioural therapies, will be summarised descriptively for therapists delivering iCT-SAD.

9. Primary efficacy analysis

The primary analysis will compare LSAS scores between iCT-SAD + TAU and TAU alone across week 8 and week 15 using a linear mixed-effects model. The dependent variable will be LSAS total score at week 8 and week 15. Fixed effects will include treatment arm, time, treatment-by-time interaction, ~~and~~ baseline LSAS score, and the randomisation stratification variable(s), where applicable. Each participant will be included as a random intercept. The model will be fitted using restricted maximum-likelihood (REML), and an unstructured covariance structure will be used where model convergence permits.

The primary treatment effect will be the adjusted between-group difference at week 15, calculated as adjusted mean LSAS in the TAU-alone arm minus adjusted mean LSAS in the iCT-SAD + TAU arm. A positive difference therefore favours iCT-SAD + TAU. The adjusted difference, 95% CI, p-value, and standardised between-group effect size will be reported.

Model assumptions: Residual normality, homoscedasticity, and influential observations will be examined descriptively. If model convergence or distributional problems occur, the trial statistician will prespecify and document a parsimonious alternative model before unblinded interpretation of results.

10. Continuous secondary outcome analyses

Each continuous secondary outcome will be analysed using a linear mixed-effects model analogous to the primary model. The baseline score of the outcome being analysed will be included as a fixed covariate, together with treatment arm, time, treatment-by-time interaction, and the randomisation stratification variable where applicable. Results will be reported as adjusted between-group differences at week 15 with 95% CIs and p-values.

11. Effect size definitions

Effect sizes will be clearly separated into controlled between-group effect sizes and within-group effect sizes.

- **Between-group effect sizes for randomised comparisons:** For continuous outcomes, Cohen's d will be calculated by dividing the adjusted between-group difference by the observed baseline standard deviation across both randomised treatment groups for the relevant outcome (i.e. using all randomised participants with available baseline data). Where appropriate, supplementary effect sizes may be reported to facilitate comparison with previous iCT-SAD RCTs.
- **Within-group effect sizes:** Within-group effect sizes will not be used as confirmatory evidence for the randomised comparison. They will be calculated for descriptive interpretation within iCT-SAD recipients and for post-control intervention analyses. These estimates will be obtained from separate mixed-effects models that include baseline as a timepoint rather than as a covariate, where feasible.
- **Direction:** For LSAS, SPWSS, SCQ, SBQ, SAQ, PHQ-9, GAD-7, and WSAS, positive effect sizes will be coded to indicate greater improvement in favour of iCT-SAD where possible.

12. Categorical secondary outcome analyses

Reliable response, remission, reliable deterioration, and loss of SAD diagnosis will be summarised by randomised arm at week 15 in the FAS. Between-arm comparisons will use Fisher exact tests. Risk differences and odds ratios will be reported with 95% CIs where estimable.

Participants with missing or unevaluable data for a categorical endpoint will remain in the denominator. The numerator will include only participants confirmed to meet the endpoint definition; all other participants will be classified as not having confirmed endpoint achievement for the purpose of categorical outcome analyses.

13. Independent assessor blinding

The success of independent assessor blinding for the post-treatment diagnostic assessment will be examined descriptively. Assessors' allocation guesses at Post will be summarised by actual randomised arm, and Bang's blinding index will be calculated where data permit. Ancillary forced-choice guesses following "do not know" responses will be reported separately or used in ancillary analyses.

14. Missing data

Data completeness will be reported by arm and assessment timepoint. The primary mixed-effects models for continuous outcomes will use all available outcome data under a missing-at-random assumption and will not impute missing observations. No formal sensitivity analyses under missing-not-at-random assumptions are planned because of the expected low level of missing data and multiple imputation is not planned for the primary analysis. For categorical endpoints, missing or unevaluable data will remain in the denominator as specified above. Reasons for missing data and withdrawals will be summarised where available.

15. Post-control intervention analyses

Post-control intervention analyses will include the post-control intervention subgroup, defined as participants originally allocated to TAU alone who receive iCT-SAD after the week-15 assessment. These analyses will be descriptive and exploratory because they do not preserve the original randomised between-group contrast.

Within this population, changes in continuous outcomes across Pre (post-control), Mid (post-control), Post (post-control), and FU (post-control) will be summarised using observed means and standard deviations and, where appropriate, within-group mixed-effects models with time as a categorical fixed effect. Pre-to-Post and Pre-to-FU within-group effect sizes will be calculated where data permit. These analyses may include p-values for within-person change, but they will not be interpreted as confirmatory efficacy tests.

Categorical response, remission, deterioration, loss of diagnosis, AEs, SAEs, engagement, delivery, and feedback measures will be summarised descriptively using the same definitions as for the randomised intervention arm.

16. Benchmarking against previous iCT-SAD studies

Benchmarking will be performed through descriptive comparisons of the present iCT-SAD + TAU arm with iCT-SAD arms from previous RCTs in the United Kingdom (Clark et al., 2023) and Hong Kong (Thew et al., 2022). Because these are cross-trial comparisons and the trial contexts differ, benchmarking will not be interpreted as a formal efficacy comparison.

The benchmarking summary will include the following items where available. Definitions and assessment timings will be aligned as closely as possible across studies, and any important differences will be noted in the trial report:

- Country or setting;
- iCT-SAD adaptation status;
- Study design, including the number randomised to iCT-SAD and comparator arm(s);
- LSAS outcomes in the iCT-SAD arm, including means and standard deviations at baseline, post-treatment, and follow-up, and Pre-to-Post and Pre-to-FU within-group effect sizes;
- Controlled between-group effect size versus comparator at post-treatment;
- Response, remission, deterioration, and loss-of-diagnosis rates at post-treatment and follow-up;
- Dropout rates during the treatment period and follow-up retention.

17. Exploratory analyses

17.1 Exploratory moderation analyses

Baseline clinical and demographic characteristics will be examined as potential moderators of treatment outcomes by running linear mixed-effect models with each candidate moderator and its interaction with treatment and time (treatment × moderator × time interaction) added.

17.2 Exploratory mediation analyses

Mediation models will examine candidate mediators of the relationship between randomisation (treatment arm) and post-treatment LSAS scores. The candidate mediators will be negative social cognitions (SCQ), safety behaviours (SBQ), self-focused attention (the general self-focused attention item of the SPWSS), rumination (the rumination item of the SPWSS), and depressed mood (PHQ). Mid-timepoint scores are used as the mediator, and post-treatment LSAS scores as the outcome. All models include baseline LSAS and baseline mediator scores as covariates. Mediation analyses will be conducted using regression-based mediation models, with indirect effects estimated using non-parametric bootstrap confidence intervals. Mediation analyses will be conducted following the procedure described in Freeman et al. (2017), using linear mixed effects models within the Baron & Kenny (1986) framework.

17.3 Additional exploratory analyses

Further exploratory analyses may be considered for additional secondary analysis of the trial data. This includes, but is not necessarily limited to, the following:

- Patient feedback on iCT-SAD
- Psychiatrist feedback on iCT-SAD
- Efficacy of individual treatment modules and components
- Examination of module sequencing
- Longitudinal mediation

18. Statistical software and reproducibility

Analyses will be conducted using validated statistical software such as R or Stata. The software, package names, and versions used for the final analyses will be recorded. Analysis scripts will be archived with the final analysis dataset.

19. References

Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology. 1986;51(6):1173-1182. doi: 10.1037/0022-3514.51.6.1173.

Clark DM, Wild J, Warnock-Parkes E, Stott R, Grey N, Thew G, Ehlers A. More than doubling the clinical benefit of each hour of therapist time: a randomised controlled trial of internet cognitive therapy for social anxiety disorder. Psychological Medicine. 2023;53(11):5022-5032. doi: 10.1017/S0033291722002008.

Freeman D, Sheaves B, Goodwin GM, Yu LM, Nickless A, Harrison PJ, Emsley R, Luik AI, Foster RG, Wadekar V, Hinds C, Gumley A, Jones R, Lightman S, Jones S, Bentall R, Kinderman P, Rowse G, Brugha T, Blagrove M, Gregory AM, Fleming L, Walklet E, Glazebrook C, Davies EB, Hollis C, Haddock G, John B, Coulson M, Fowler D, Pugh K, Cape J, Moseley P, Brown G, Hughes C, Obonsawin M, Coker S, Watkins E, Schwannauer M, MacMahon K, Siriwardena AN, Espie CA. The effects of improving sleep on mental health (OASIS): a randomised controlled trial with mediation analysis. Lancet Psychiatry. 2017;4(10):749-758. doi: 10.1016/S2215-0366(17)30328-0.

Thew GR, Kwok APL, Lissillour Chan MH, Powell CLYM, Wild J, Leung PWL, Clark DM. Internet-delivered cognitive therapy for social anxiety disorder in Hong Kong: A randomized controlled trial. Internet Interventions. 2022;28:100539. doi: 10.1016/j.invent.2022.100539..

Yoshinaga N, Thew GR, Hayashi Y, Tanoue H, Nakai M, Clark DM. Translated and culturally adapted internet-delivered cognitive therapy for social anxiety disorder in Japanese clinical settings: study protocol for a randomised controlled trial. Trials. 2024;25(1):492. doi: 10.1186/s13063-024-08303-6.