GreenME
Mental health for all through nature

Date: 26 June 2026.

GreenME WP3 Task 3.3 Statistical Analysis Plan

Version 03 for statistical and scientific discussion

Authors and Contributers: Helen V. S. Cole, Paula de Prado Bert, Juan Vicente Luciano, Zachary Zeiger, Margarita Triguero-Mas, Richard Blecher, Sarah Knight, Penny Cook

Statistical review by: Ferran Torres / UAB Biostatistics Unit.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

UK Research and Innovation

OEI – Requirement No. 1

Grant Agreement nº	101084198
Project acronym	GreenME
Project title	Advancing Greencare in Europe: an integrated multi-scalar approach for the expansion of nature-based therapies to improve mental health equity (GreenME)
Call	Resilient, inclusive, healthy and green rural, coastal and urban communities (HORIZON-CL6-2022-COMMUNITIES-02-twostage)
Topic	HORIZON-CL6-2022-COMMUNITIES-02-02-two-stage - Developing nature-based therapy for health and well-being
Project Duration	01/09/2023 – 31/08/2027
Coordinator:	Universitat Autònoma de Barcelona (UAB)
Associated Beneficiaries	• Alma Mater Studiorum – Universita Di Bologna (UNIBO) • Sveriges Lantbruksuniversite (SLU) • ILS Research GGMBH (ILSR) • Szkola Glowna Gospodarstwa Wiejskiego (SGGW) • Old-Continent (OC) • Fundacja Neurolandscape (NL) • Gesellschaft Für Gartenbau und Therapie (GGUT) • Instytut Psychiatrii I Neurologii (IPIN) • Wetterholm Petra (SHINRIN-YOKU) • ETA BETA Cooperativa Social (ETA BETA) • Stadt HERNE (HERNE) • Asociación Española de Horticultura y Jardinería Social y Terapéutica (AEHJST) • Universitat Oberta de Catalunya (UOC)

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

UK Research and Innovation

Associated Partners	• University of Kent (UNIKENT) • The University of Salford (USAL) • Social Farms and Gardens (SF&G) • Mind in Bexley and East Kent LTD (MBEK) • Edge Hill University (EDGE-HILL) • Oregon Science and Health University (OSHU)

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

Project n. 101084198

Advancing Greencare in Europe: an integrated multi-scalar approach for the expansion of nature-based therapies to improve mental health equity (GreenME).

HISTORY CHART

Issue	Date	Changed page(s)	Cause of change	Implemented by
V1	01/02/2026	-	First complete draft	Helen V. S. Cole, Paula de Prado Bert, Zachary Zeiger and Juan Vicente Luciano
V2	23/06/2026	All	Revision by external statistician	Ferran Torres
V3	26/10/2026	All	Formatting	Helen V. S. Cole, Paula de Prado Bert, Zachary Zeiger and Juan Vicente Luciano

VALIDATION CHART

No.	Action	Beneficiary	Date
V1		UAB/UOC	19 feb 2026
V3		UAB/UOC	26 oct 2026

Disclaimer: The information in this document is subject to change without notice. Company or product names mentioned in this document may be trademarks or registered trademarks of their respective companies.

All rights reserved.

The document is proprietary of the GREENME consortium members. No copying or distributing, in any form or by any means, is allowed without the prior written agreement of the owner of the property rights.

This document reflects only the authors' view. The European Commission is not liable for any use that may be made of the information contained herein.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

UK Research and Innovation

Contents

Abbreviations 6
1. Introduction and purpose 6
2. Study objectives and hypotheses 6
3. Study design overview 7
4. Estimands 8
5. Analysis populations 9
6. Outcomes and derivations 10
6.1 Visit windows and timepoints 10
6.2 Primary endpoint 10
6.3 Secondary and exploratory endpoints 10
7. General statistical principles 11
8. Baseline and trial conduct summaries 11
9. Primary effectiveness analysis 12
9.1 Intervention-specific primary model 12
9.2 Meta-analysis of intervention-specific effects 13
10. Secondary and sensitivity effectiveness analyses 13
11. Missing data 15
12. Health-economic analysis 16
13. Mechanism, mediation and moderation analyses 17
14. Safety and harms 17
15. Multiplicity and interpretation 17
16. Software and reproducibility 18
17. References 19
Appendix A. Draft statistical table shells 20
A1. Primary endpoint table shell 20
A2. Sensitivity-analysis table shell 20
Appendix B. Proposed protocol-deviation classification 20

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

UK Research and Innovation

Abbreviations

Abbreviation	Definition
ANCOVA	analysis of covariance
CI	confidence interval
CSRI	Client Service Receipt Inventory
EQ-5D-5L	EuroQol 5 Dimensions 5 Levels
ICER	incremental cost-effectiveness ratio
ITT	intention-to-treat
MAR	missing at random
MNAR	missing not at random
MMRM	mixed model for repeated measures
NBT	nature-based therapy
ONS-4	Office for National Statistics four personal wellbeing questions
PSS-10	10-item Perceived Stress Scale
QALY	quality-adjusted life year
SAP	statistical analysis plan
TAU	treatment/usual care as available locally
TIDieR	Template for Intervention Description and Replication

1. Introduction and purpose

GreenME is a Horizon Europe project evaluating whether existing nature-based therapies can be scaled up to improve adult mental health and wellbeing equity. WP3 Task 3.3 evaluates seven NBT interventions in Spain, Italy, the UK and Sweden. The interventions include therapeutic horticulture/social gardening, forest bathing/forest therapy, and surf/ocean-based therapy, delivered in real-world settings by trained facilitators.

This SAP specifies the statistical principles and analytical methods for the final quantitative analyses of GreenME WP3. It is intentionally more detailed than the protocol manuscript and the earlier WP3 analysis-plan draft. It should be read together with the final protocol manuscript, intervention-specific protocols, data dictionary, database-management plan, and health-economic analysis plan. The SAP is written to minimise analytic flexibility once outcome data have been collected and to make clear which analyses are confirmatory, secondary, sensitivity, exploratory, or descriptive.

This is not a medicinal-product clinical trial, but it is a randomised experimental study in humans. Therefore, the SAP adopts relevant clinical-trial statistical principles where appropriate, especially ICH E9, ICH E9(R1), and the JAMA SAP guidance by Gamble et al. This does not imply regulatory clinical-trial status; it is a quality framework for transparent analysis.

2. Study objectives and hypotheses

The overall objective is to estimate whether assignment to an NBT plus unrestricted usual care improves mental health outcomes compared with unrestricted usual care

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

UK Research and Innovation

alone among adults with moderate-to-high perceived stress who are eligible for the evaluated interventions.

Level	Objective	Endpoint(s)	Interventions
Primary	To estimate the effect of NBT plus usual care versus usual care on perceived stress immediately after the intervention.	PSS-10 post-intervention total score.	All seven interventions.
Secondary 1	To estimate effects on quality of life and subjective wellbeing immediately after the intervention.	EQ-5D-5L index and EQ-VAS; ONS-4 items/domains.	All seven interventions.
Secondary 2	To assess durability of effects at 1-month and 3-month follow-up.	PSS-10, EQ-5D-5L, EQ-VAS, ONS-4 at follow-up.	All interventions with available follow-up; UK follow-up after cross-over handled separately.
Secondary 3	To evaluate whether NBT is cost-effective compared with usual care.	Costs, QALYs, ICERs, net monetary benefit.	All non-UK interventions. UK interventions will be excluded from the primary health-economic evaluation because the planned post-FU1 crossover prevents a valid longer-term between-group comparison.
Secondary 4	To examine perceived stigma and self-esteem related effects.	SESQ stigma and self-esteem scales; stigma-by-treatment and stigma-by-outcome analyses.	All interventions collecting SESQ; note any deviations.
Exploratory	To examine mechanisms and heterogeneity of effect.	Nature connectedness and other candidate mediators/moderators; axes of inequality.	As data allow.

3. Study design overview

The study is a pragmatic, multicentre, parallel-group experimental evaluation of seven existing interventions: CSMaresme and PSSJD in Spain, Salus and Battirame in Italy, Northern Roots and TONIC in the UK, and ShinrinYoku in Sweden. Participants are randomised within intervention in a 1:1 ratio to NBT plus unrestricted usual care or unrestricted usual care alone. The comparator differs by site and reflects local usual practice; it may include active usual-care activities in some sites and waiting-list/usual care in others.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

UK Research and Innovation

Intervention	Country	NBT type	Exposure schedule	Comparator
CSMaresme	Spain	Social/therapeutic horticulture	12 sessions: 2 x 90 minutes/week for 6 weeks	Active usual-care activities at the day hospital
PSSJD	Spain	Forest bathing / forest therapy	12 weekly sessions, generally 120 minutes	Grief therapy group / usual care
Salus	Italy	Social/therapeutic horticulture	12 weekly sessions, generally 120 minutes	Community health/social usual care; no study structured activity
Battirame	Italy	Social/therapeutic horticulture	12 weekly sessions, generally 120 minutes	Community health/social usual care; no study structured activity
Northern Roots	UK	Nature connection / social therapeutic gardening	10 weekly sessions	Waiting-list/usual care until post-intervention; then offered NBT
TONIC	UK	Surf/ocean therapy	10 weekly sessions	Waiting-list/usual care until post-intervention; then offered NBT
ShinrinYoku	Sweden	Forest bathing / eco-forest therapy	12 weekly sessions, generally 120 minutes	Usual care; no study structured activity

Randomisation is by computer-generated block randomisation within each intervention. Allocation is implemented using sealed envelopes by personnel not involved in recruitment/data collection. Participants and data collectors cannot be blinded; analysis should be performed using masked group codes until the SAP and analysis code are approved and the analysis dataset is locked.

4. Estimands

The estimand framework follows ICH E9(R1). The primary question is the effect of assignment to NBT plus usual care compared with usual care, not the effect among only adherent participants.

Attribute	Primary estimand specification
Treatment condition	Assignment to NBT plus unrestricted usual care versus assignment to unrestricted usual care / waiting-list control.
Population	Randomised eligible adults with moderate-to-high perceived stress in each intervention-specific target population.
Variable	PSS-10 total score at the immediate post-intervention assessment. Lower values indicate less perceived stress.
Intercurrent events	Non-adherence, missed sessions, discontinuation of NBT, changes in usual care, and use of additional services are handled using a treatment-policy strategy in the primary estimand. Death or inability to provide outcome data will be handled as missing, with sensitivity analyses if non-negligible. UK safety/logistical reassignment after randomisation will be excluded from the primary ITT only if the participant is formally flagged

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

UK Research and Innovation

	as non-randomised before exposure; otherwise it will be retained by randomised allocation and examined in sensitivity analyses. For the UK interventions, the protocol-planned crossover of control participants to NBT after FU1 is also considered an intercurrent event. This does not affect the primary estimand, which is defined at the post-intervention assessment (FU1), but precludes a valid between-group treatment-policy comparison at subsequent follow-up timepoints. Consequently, pooled comparative analyses beyond FU1 will exclude the UK interventions, while UK follow-up data may be analysed separately as supportive or exploratory analyses.
Population-level summary	Adjusted mean difference in post-intervention PSS-10 score: NBT minus usual care, with 95% CI. Negative values favour NBT.

A secondary per-protocol/complier estimand will estimate the effect under sufficient receipt of the assigned NBT, defined before analysis as attendance at at least 75% of scheduled sessions among those allocated to NBT and no pre-FU1 NBT exposure among controls. This analysis is supportive and potentially biased because adherence is post-randomisation.

5. Analysis populations

Population	Definition	Use
Screened population	All individuals assessed for eligibility.	CONSORT flow, screening failures, representativeness.
Randomised population	All individuals randomised, regardless of whether they start the assigned condition.	Denominator for trial conduct summaries.
Full analysis set / ITT	All randomised participants with at least baseline/screening PSS-10 and at least one valid post-baseline assessment, analysed according to randomised assignment. Participants without any post-baseline outcome contribute to mixed-model analyses only if the model can use available data; otherwise they are included in imputation-based sensitivity analyses.	Primary and secondary effectiveness analyses.
Per-protocol set	Participants without major eligibility/randomisation deviations and with adequate adherence. Adequate NBT adherence is at least 75% scheduled sessions; adequate control adherence requires no NBT exposure before FU1. Additional site-specific criteria to be finalised before lock.	Supportive sensitivity analysis.
Safety/harms set	All participants who start any assigned study-related activity, analysed according to activity actually received.	Adverse events and harms summaries.
Economic set	Participants with sufficient EQ-5D-5L and cost data for QALY and cost estimation; imputation set to include all eligible participants with baseline data.	Health-economic analyses.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

UK Research and Innovation

6. Outcomes and derivations

6.1 Visit windows and timepoints

Timepoint	Label	Definition/notes
Screening / baseline	-t1/t0	Eligibility, demographics/covariates, baseline outcomes. The baseline adjustment variable for the primary analysis will be the pre-intervention PSS-10 score, defined as the last valid PSS-10 total score collected before randomisation or, if randomisation occurs immediately before the intervention, before initiation of the allocated condition. Depending on the intervention-specific data-collection procedures, this may correspond to screening (-t1) or baseline/Questionnaire A (t0). The source timepoint will be documented by intervention in the data dictionary and checked before database lock.
Post-intervention	FU1 / t2	Primary endpoint. End of exposure/usual-care period: approximately 6 weeks for CSMaresme, 10 weeks for UK interventions, 12 weeks for others.
1-month follow-up	FU2 / t3	Approximately 4 weeks after intervention end. UK controls may have crossed over to NBT after FU1 and will not be pooled for this timepoint.
3-month follow-up	FU3 / t4	Approximately 12 weeks after intervention end. UK controls may have crossed over to NBT after FU1 and will not be pooled for this timepoint.

Visit windows should be finalised in the data-management plan. Suggested analysis windows are target date +/-14 days for FU1-FU3. If multiple valid questionnaires occur in a window, use the assessment closest to the target date; if tied, use the earliest non-missing completed questionnaire.

6.2 Primary endpoint

The primary endpoint is PSS-10 total score at post-intervention. PSS-10 items are scored 0 to 4, with positively worded items reverse-coded according to the instrument manual. The total score ranges from 0 to 40, with higher scores indicating higher perceived stress. A valid PSS-10 total score requires at least 8 of 10 items completed. If 8 or 9 items are complete, the total score will be prorated as mean completed item score x 10; otherwise the score is missing. The same scoring rule will be used across all timepoints and languages.

6.3 Secondary and exploratory endpoints

Measure	Endpoint	Derivation/analysis note
EQ-5D-5L	Index utility and EQ-VAS	Use country-specific value sets where available; if not, pre-specify a single fallback value set before analysis. EQ-VAS ranges 0-100.
ONS-4	Life satisfaction, worthwhile, happiness, anxiety	Analyse each item separately. Anxiety item direction should be clearly labelled; for a total wellbeing score,

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

UK Research and Innovation

		reverse-code anxiety and pre-specify derivation.
SESQ	Stigma and self-esteem subscales	Use instrument-specific scoring. Higher score direction must be documented and checked before analysis.
Nature connectedness	Pathway/moderator measure	Analyse as potential mediator or effect modifier; not confirmatory.
GAD-7, PSQI-2, DERS-16, RSES, PHQ-9, K10, ICG, WEMWBS	Exploratory mental health and wellbeing endpoints	Analysed only for interventions collecting the scale. Clearly label as exploratory and avoid confirmatory claims.
Costs and QALYs	Cost-utility and cost-effectiveness endpoints	See Section 12.

7. General statistical principles

- All statistical tests will be two-sided. The primary analysis will use a 5% significance level and 95% confidence intervals.
- The primary analysis will be conducted separately within each intervention, followed by meta-analysis of intervention-specific adjusted effects.
- Effect estimates and confidence intervals are more important than p values; all p values will be interpreted in the context of multiplicity and the pragmatic nature of the study.
- Baseline characteristics will be summarised by randomised group but will not be tested using p values. Imbalances will be described clinically and statistically.
- Continuous outcomes will be summarised by mean, SD, median, IQR, minimum and maximum; categorical outcomes by counts and percentages.
- Model assumptions will be assessed using residual plots and influence diagnostics. Transformations will not be used for the primary endpoint unless specified before unblinding; robust or non-parametric supportive analyses may be added if severe violations occur.
- All analysis code should be version-controlled and run on a frozen, locked analysis dataset. Deviations from this SAP will be documented in the statistical report.

8. Baseline and trial conduct summaries

A CONSORT-style flow diagram will be prepared overall and by intervention: screened, excluded with reasons, randomised, allocated to each group, received allocated condition, discontinued NBT/usual-care follow-up, completed FU1/FU2/FU3, included in each analysis population.

Baseline tables will be produced by intervention and randomised group, and overall. Variables include age, gender, sex, ethnicity/race, education, perceived income, sexual orientation where collected, discrimination/marginalisation measures, stigma, baseline PSS-10, baseline EQ-5D-5L, baseline ONS-4, medication intake, exposure to green care, nature connectedness, recruitment wave/season, and intervention-specific target-population descriptors.

Adherence will be summarised as number and percentage of sessions attended, total minutes attended where available, attendance category (<25%, 25-49%, 50-74%,

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

UK Research and Innovation

>=75%), missed-session reasons, and comparator contamination (control participants receiving NBT or similar exposure before FU1). Protocol deviations will be listed and classified as major or minor before database lock.

9. Primary effectiveness analysis

The primary endpoint is the PSS-10 total score at the immediate post-intervention assessment (FU1). Although repeated PSS-10 measurements are available throughout follow-up, the primary estimand concerns the treatment effect at FU1, immediately after completion of the intervention period. Therefore, the primary analysis will use an intervention-specific mixed model for repeated measures (MMRM), incorporating all available post-baseline PSS-10 assessments under a Missing At Random (MAR) assumption, while estimating the primary treatment contrast at FU1. This approach is consistent with the intention-to-treat principle and avoids restricting the primary analysis to participants with an observed FU1 assessment only.

The primary analysis will include all seven interventions, including the UK interventions, for which the randomised comparison remains valid up to FU1. Because participants allocated to the control group in the UK interventions are offered NBT after FU1 by design, pooled comparative analyses beyond FU1 will exclude the UK interventions, as these follow-up comparisons no longer preserve the original randomised contrast. As a supportive sensitivity analysis, the primary pooled treatment effect at FU1 will also be estimated after excluding the UK interventions to assess the robustness of the overall conclusions.

9.1 Intervention-specific primary model

For each intervention, the primary treatment effect will be estimated using a Mixed Model for Repeated Measures (MMRM), including all available PSS-10 assessments collected from the pre-intervention assessment through the scheduled follow-up visits. The model will include fixed effects for randomised treatment group, assessment timepoint, treatment-by-timepoint interaction, pre-intervention PSS-10 score, and any pre-specified design variables (e.g. recruitment wave or season, if applicable). The primary estimand will be the adjusted mean difference between treatment groups at the immediate post-intervention assessment (FU1), estimated from the treatment-by-timepoint interaction. Negative values indicate lower perceived stress and therefore favour the Nature-Based Therapy (NBT) intervention.

Repeated measurements within participants will be modelled using an appropriate within-participant covariance structure. An unstructured (UN) covariance matrix will be used whenever model convergence and estimation are satisfactory. If the UN covariance structure fails to converge or produces unstable parameter estimates, a first-order autoregressive [AR(1)] covariance structure will be fitted. Compound symmetry (CS) may be used as a final fallback option if required. Estimation will be performed using Restricted Maximum Likelihood (REML). The selected covariance structure will be applied consistently across interventions whenever feasible.

The primary analysis will be performed separately within each intervention. Intervention-specific adjusted treatment effects at FU1 and their corresponding

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

UK Research and Innovation

standard errors will subsequently be synthesised using the pre-specified two-stage individual participant data (IPD) meta-analysis described in Section 9.2.
If the MMRM cannot be reliably fitted for a particular intervention because of limited sample size or persistent convergence problems, a baseline-adjusted ANCOVA at FU1 will be used as a predefined fallback analysis for that intervention. Any such deviation from the primary analysis strategy will be documented and justified in the Statistical Report.

9.2 Meta-analysis of intervention-specific effects

The intervention-specific adjusted mean differences at FU1, estimated from the intervention-specific MMRM models, together with their corresponding standard errors, will be synthesised using a random-effects two-stage individual participant data (IPD) meta-analysis. Restricted maximum likelihood (REML) will be used as the primary estimator of the between-intervention variance (τ^2). Hartung-Knapp confidence intervals will be explored as a supportive sensitivity analysis, acknowledging the relatively small number of interventions. A fixed-effect meta-analysis will also be performed as a sensitivity analysis but will not be considered the primary pooled analysis because important clinical and operational heterogeneity across the GreenME interventions is expected.
Results will be presented as forest plots showing intervention-specific treatment effects together with the pooled estimate, corresponding 95% confidence intervals, τ^2 and I^2 statistics. The primary interpretation will focus on both the magnitude of the pooled effect and the consistency of treatment effects across interventions. Accordingly, the pooled estimate should be interpreted as the average treatment effect across the evaluated GreenME nature-based interventions rather than as the effect of a single homogeneous intervention.
As a supportive analysis, a one-stage IPD mixed-effects model including intervention and treatment-by-intervention interaction terms will be fitted to evaluate the robustness of the primary findings and to formally assess potential heterogeneity of treatment effects across interventions. Results from this model will be interpreted as supportive and will not replace the primary two-stage analysis.

10. Secondary and sensitivity effectiveness analyses

Sensitivity analyses are intended to assess the robustness of the primary findings to key analytical assumptions. They are considered supportive analyses and will not modify the interpretation of the primary confirmatory analysis. Unless otherwise specified, statistical significance from sensitivity analyses will be interpreted as supportive rather than confirmatory.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

UK Research and Innovation

Analysis	Method	Purpose
Quality of life and wellbeing at FU1	Intervention-specific MMRM using the same modelling strategy as the primary analysis; random-effects two-stage IPD meta-analysis where common endpoints are available.	EQ-5D-5L index, EQ-VAS and ONS-4 outcomes.
Durability of treatment effect (FU1-FU3)	Intervention-specific MMRM including all available post-baseline assessments.	Evaluate maintenance of treatment effects over time. UK interventions will be excluded from pooled comparative analyses beyond FU1 because of the planned crossover; UK follow-up results will be presented descriptively or analysed separately as supportive analyses.
Baseline-adjusted ANCOVA at FU1	ANCOVA including treatment group and pre-intervention PSS-10, restricted to participants with observed FU1 assessments.	Supportive sensitivity analysis comparing results with the primary MMRM analysis.
Primary analysis excluding UK interventions	Repeat the primary MMRM and two-stage IPD meta-analysis excluding UK interventions.	Evaluate the robustness of the pooled treatment effect to the UK-specific crossover design.
Per-protocol analysis	Repeat the primary analysis in the Per-Protocol population.	Assess consistency with the primary intention-to-treat analysis.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

UK Research and Innovation

Stigma and self-esteem	MMRM using the same modelling framework as the primary analysis. For stigma as a treatment-effect modifier, include the baseline stigma-by-treatment interaction.	Exploratory analyses unless otherwise pre-specified.
Responder analyses	Pre-defined responder threshold analysed using risk difference, risk ratio or logistic regression adjusted for pre-intervention PSS-10.	Supportive only.
One-stage IPD model	Mixed-effects model including intervention and treatment-by-intervention interaction terms.	Supportive analysis to evaluate the robustness of the primary two-stage IPD meta-analysis and formally assess treatment-effect heterogeneity across interventions.

11. Missing data

The primary analysis will follow the intention-to-treat principle and will use all available outcome data through the intervention-specific mixed model for repeated measures (MMRM). Under the likelihood-based estimation used by the MMRM, valid inference is obtained under the Missing At Random (MAR) assumption without explicit imputation of missing outcome values.

The extent and pattern of missing data will be summarised by intervention, treatment group and assessment timepoint. Reasons for missing data, where available, will also be summarised descriptively.

Component	Method	Purpose
Primary analysis	MMRM using all available repeated PSS-	Primary confirmatory analysis.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

UK Research and Innovation

	10 assessments under the MAR assumption.	
Sensitivity analysis	Baseline-adjusted ANCOVA restricted to participants with observed FU1 assessments.	Evaluate consistency with the primary MMRM analysis.
MNAR sensitivity	If the extent or pattern of missing data raises concerns regarding the MAR assumption, additional sensitivity analyses under Missing Not At Random (MNAR) assumptions (e.g. delta-adjusted analyses or tipping-point analyses) may be performed.	Exploratory assessment of robustness to departures from MAR.

Item-level missingness will be handled according to the scoring rules for each instrument described in Section 6. If unexpected levels of item-level missingness are observed, additional exploratory analyses comparing prorated and complete-case scoring methods may be performed.

12. Health-economic analysis

The health-economic analysis should be finalised by the economics lead in a dedicated HEAP, but the SAP should define the core principles. The main evaluation will be conducted from both healthcare/service and societal perspectives, where data allow. Costs will include intervention delivery costs, health and social care resource use from CSRI or equivalent, medication use where available, and productivity or informal-care costs if collected and defensible.

QALYs will be estimated from EQ-5D-5L utility values using area under the curve, adjusting for baseline utility. The primary economic horizon will be baseline to 3-month follow-up for interventions with valid follow-up. QALYs will be estimated from EQ-5D-5L utility values using area under the curve, adjusting for baseline utility. The primary economic horizon will be baseline to 3-month follow-up for interventions with valid follow-up. The primary health-economic evaluation will include only interventions for which a valid between-group comparison can be maintained throughout the planned evaluation period. Consequently, UK interventions will not contribute to the primary pooled health-economic analysis because the planned crossover after FU1 prevents a valid longer-term between-group comparison. UK data may be summarised descriptively or analysed separately over the baseline-to-FU1 period, provided that cost and utility data allow a valid short-term comparison.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

UK Research and Innovation

- Report mean costs, mean QALYs, adjusted incremental costs, adjusted incremental QALYs, ICERs, and net monetary benefit.
- Use seemingly unrelated regression or appropriate regression models adjusting for baseline utility/costs, treatment group, and intervention/wave as relevant.
- Use non-parametric bootstrap for uncertainty and cost-effectiveness acceptability curves over pre-specified willingness-to-pay thresholds.
- Use multiple imputation for missing cost and utility data, ensuring skewed cost distributions are handled appropriately.
- Present cost-effectiveness planes and scenario analyses for intervention cost assumptions.

13. Mechanism, mediation and moderation analyses

Mechanistic analyses are exploratory and should not be used to make confirmatory causal claims unless assumptions are explicitly stated. The main candidate mechanism is change in nature connectedness. A pragmatic approach is to estimate whether treatment affects nature connectedness and whether change in nature connectedness is associated with change in outcomes, adjusting for baseline values. Formal mediation may be conducted using counterfactual or structural equation methods if timing supports mediator-before-outcome ordering; otherwise analyses should be described as pathway analyses.

Effect modification will be explored using treatment-by-subgroup interaction terms within intervention where possible and/or in one-stage IPD models. Candidate moderators should be pre-specified and limited: baseline PSS-10 severity, gender, sex, age, perceived income/education, baseline nature connectedness, baseline stigma, and intervention type. Results will be presented as interaction estimates with 95% CIs; subgroup-specific p values without interaction tests will be avoided.

14. Safety and harms

Although the interventions are expected to be low risk, adverse events, serious adverse events, distress episodes, injuries, safeguarding events, and withdrawals due to worsening mental health will be summarised by intervention and actual activity received. Narratives should be provided for serious or clinically important events. No formal hypothesis testing of harms is planned.

15. Multiplicity and interpretation

The primary confirmatory hypothesis concerns the treatment effect on PSS-10 at FU1. No formal multiplicity adjustment will be applied to secondary endpoints, supportive analyses or sensitivity analyses, which will be interpreted as exploratory and hypothesis-supporting. Confidence intervals and effect sizes will be emphasised over statistical significance. Any nominal p-values reported for secondary analyses will be interpreted descriptively and will not constitute confirmatory evidence.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

UK Research and Innovation

16. Software and reproducibility

All statistical analyses will be performed using validated statistical software. The primary analyses are expected to be conducted in R (R Foundation for Statistical Computing, Vienna, Austria; version 4.4 or later) using validated and publicly available packages. All statistical code will be version-controlled and archived to ensure reproducibility of the analyses.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

UK Research and Innovation

17. References

Gamble C, Krishan A, Stocken D, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. JAMA. 2017;318(23):2337-2343. doi:10.1001/jama.2017.18556.

Hiemstra B, Keus F, Wetterslev J, Gluud C, van der Horst ICC. DEBATE-statistical analysis plans for observational studies. BMC Med Res Methodol. 2019;19:233. doi:10.1186/s12874-019-0879-5.

International Council for Harmonisation. ICH E9: Statistical Principles for Clinical Trials. 1998.

International Council for Harmonisation. ICH E9(R1): Addendum on Estimands and Sensitivity Analysis in Clinical Trials. 2020.

Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav. 1983;24(4):385-396.

Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. BMJ. 2010;340:c221.

Gupta SK. Intention-to-treat concept: A review. Perspect Clin Res. 2011;2(3):109-112.

McCoy CE. Understanding the intention-to-treat principle in randomized controlled trials. West J Emerg Med. 2017;18(6):1075-1078.

Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: Template for Intervention Description and Replication (TIDieR) checklist and guide. BMJ. 2014;348:g1687.

Busk H, Sidenius U, Kongstad LP, et al. Economic evaluation of nature-based therapy interventions - a scoping review. Challenges. 2022;13(1):23.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

UK Research and Innovation

Appendix A. Draft statistical table shells

A1. Primary endpoint table shell

Intervention	NBT n	Control n	NBT adjusted mean	Control adjusted mean	Adjusted MD (NBT-control)	95% CI / p-value
CSMaresme						
PSSJD						
Salus						
Battirame						
Northern Roots						
TONIC						
ShinrinYoku						
Random-effects pooled						

A2. Sensitivity-analysis table shell

Analysis	No. interventions	Pooled MD	95% CI	Interpretation
Primary ANCOVA				
Complete-case only				
MMRM				
Multiple imputation MAR				
Per-protocol				
Fixed-effect meta-analysis				
Excluding non-randomised/reassigned UK participants				
MNAR/tipping-point				

Appendix B. Proposed protocol-deviation classification

Class	Examples
Major	Ineligible participant randomised; randomisation envelope/order violation; incorrect allocation not due to documented safety/logistical reassignment; no baseline PSS-10; control receives NBT before FU1; NBT participant receives <25% sessions with no post-baseline outcome.
Minor	Assessment outside visit window but usable; missed individual items with valid prorated scale score; temporary session interruption; session make-up within protocol allowance.
To classify case-by-case	Change in medication/usual care; hospitalisation; severe adverse event; participant withdrawal from intervention but not follow-up; cross-over after FU1 in UK.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101084198

UK Research and Innovation