Statistical Analysis Plan

Effects of a Supported Internet-Based Self-Help Mindfulness Intervention for Emotional Distress on Reflux Symptoms in Patients With Gastroesophageal Reflux Disease: A Multicenter Randomized Controlled Trial

SAP Version: 1.0

Date: June 13, 2026

Protocol version: 1.0

Sponsor/lead institution: Peking University

Principal investigator: Xinghua Liu

Statistician: Yinan Zhang

1. Purpose of the Statistical Analysis Plan

This Statistical Analysis Plan (SAP) specifies the planned statistical analyses for the multicenter randomized controlled trial evaluating the effects of a supported internet-based self-help Mindfulness Intervention for Emotional Distress (MIED) on reflux symptoms and reflux-related quality of life in patients with gastroesophageal reflux disease (GERD).

The SAP is intended to be finalized before database lock and before unblinding of group allocation for the primary analysis. Any deviations from this SAP after database lock will be documented, justified, dated, and reported in the final trial report.

2. Trial Overview

This study is a multicenter, parallel-group, randomized, waitlist-controlled clinical trial with blinded outcome assessment where feasible. Eligible participants with GERD will be randomized in a 1:1 ratio to either:

1. **MIED group:** supported internet-based self-help MIED plus treatment as usual; or
2. **Waitlist control group:** treatment as usual and routine health management during the 49-day observation period.

The MIED intervention is delivered through a standardized WeChat mini-program over 49 days. Participants complete approximately 30 minutes of daily learning and practice, including psychoeducational materials, guided mindfulness audio exercises, and daily check-ins. Assistants trained within the MIED system provide reminders, technical support, and standardized non-psychotherapeutic support.

The waitlist control group continues usual GERD care and routine health management during the active observation period. To avoid contamination of long-term follow-up outcomes, participants

in the waitlist control group may be offered the same MIED course after completion of the 24-month follow-up, according to the final approved protocol.

Assessments are scheduled at screening, baseline, week 3, week 5, week 8, and 3, 6, 12, and 24 months after baseline. Week 8 is the primary endpoint and represents the postintervention assessment window following completion of the 49-day course.

3. Study Objectives and Hypotheses

3.1 Primary Objective

The primary objective is to evaluate whether MIED plus treatment as usual improves reflux-related quality of life at week 8 compared with waitlist control plus treatment as usual.

3.2 Secondary Objectives

The secondary objectives are:

1. To evaluate the effects of MIED on GERD symptom severity, laryngopharyngeal reflux symptoms, anxiety, depression, insomnia, perceived stress, resilience, life satisfaction, and general health-related quality of life.
2. To evaluate adherence, acceptability, and safety of MIED in patients with GERD.
3. To evaluate whether MIED reduces health care resource use and disease-related costs and improves health utility and quality-adjusted life years.
4. To explore whether mindfulness, engaging in life, distress tolerance, discomfort intolerance, experiential avoidance, excessive emotional behaviors, and cognitive flexibility mediate intervention effects.
5. To explore whether baseline symptom severity, emotional distress, disease duration, sex, age, body mass index, proton pump inhibitor use, sleep disturbance, and study center moderate treatment outcomes.

3.3 Primary Hypothesis

Participants randomized to MIED plus treatment as usual will show a greater reduction in Gastroesophageal Reflux Disease-Health Related Quality of Life questionnaire (GERD-HRQL) total score from baseline to week 8 compared with participants randomized to waitlist control plus treatment as usual.

3.4 Secondary Hypotheses

1. Compared with the waitlist control group, the MIED group will show greater reductions in GERD-Q/GerdQ, Reflux Symptom Index (RSI), Overall Anxiety Severity and Impairment Scale (OASIS), Overall Depression Severity and Impairment Scale (ODSIS), Patient Health Questionnaire-9 (PHQ-9), Generalized Anxiety Disorder-7 scale (GAD-7), Insomnia Severity Index (ISI), and Chinese Perceived Stress Scale (CPSS) scores at week 8.
2. Compared with the waitlist control group, the MIED group will show greater improvements in Short Form-8 Health Survey (SF-8), Connor-Davidson Resilience Scale (CD-RISC), Satisfaction With Life Scale (SWLS), and EuroQol Five-Dimension Five-Level Questionnaire (EQ-5D-5L) outcomes.
3. Intervention effects observed at week 8 will be at least partially maintained at 3-, 6-, 12-, and 24-month follow-up.
4. Early changes in mechanistic variables from baseline to week 3 and/or week 5 will mediate the effects of MIED on GERD-HRQL and GERD-Q/GerdQ outcomes at week 8 and follow-up.
5. The incidence of serious adverse events will not be higher in the MIED group than in the waitlist control group.

4. Outcome Definitions

4.1 Primary Outcome

The primary outcome is the change in GERD-HRQL total score from baseline to week 8.

The GERD-HRQL is a disease-specific instrument assessing reflux symptom burden and reflux-related quality of life. Higher scores indicate greater symptom burden and poorer reflux-related quality of life. A reduction in score indicates improvement.

4.2 Key Secondary Outcomes

Key secondary outcomes include:

1. **GERD-Q/GerdQ:** GERD symptom frequency and severity.
2. **RSI:** laryngopharyngeal reflux and atypical reflux symptoms.
3. **OASIS:** overall anxiety severity and impairment.
4. **ODSIS:** overall depressive symptom severity and impairment.
5. **PHQ-9:** depressive symptom severity.
6. **GAD-7:** anxiety symptom severity.
7. **SF-8:** general health-related quality of life.
8. **ISI:** insomnia severity.
9. **CPSS:** perceived stress.
10. **CD-RISC:** resilience.
11. **SWLS:** life satisfaction.
12. **EQ-5D-5L:** health utility and quality-adjusted life years.
13. **TIC-P:** health care resource use, productivity loss, and cost-related outcomes.
14. **Adverse events and serious adverse events:** safety outcomes.

4.3 Mechanistic Outcomes

Mechanistic variables include:

1. **Five Facet Mindfulness Questionnaire (FFMQ):** mindfulness.
2. **Engaging in Life Scale (ELS):** engaging in life.
3. **Distress Tolerance Scale (DTS):** distress tolerance.
4. **Discomfort Intolerance Scale (DIS):** discomfort intolerance.
5. **Excessive Emotional Behavior Scale (EEBS):** excessive emotion-driven behaviors and avoidance behaviors.
6. **Brief Experiential Avoidance Questionnaire (BEAQ):** experiential avoidance.
7. **Cognitive Flexibility Inventory (CFI):** cognitive flexibility.

4.4 Safety Outcomes

Safety outcomes include:

1. Number and proportion of participants experiencing any adverse event.
2. Number and proportion of participants experiencing serious adverse events.
3. Number and proportion of participants with GERD symptom worsening.
4. Number and proportion of participants with alarm symptoms requiring medical referral.
5. Number and proportion of participants with psychological crisis, including suicidal ideation, self-harm behavior, or severe emotional deterioration.
6. Number and proportion of participants who discontinue or temporarily suspend the intervention due to adverse events.

Adverse events will be recorded regardless of presumed causal relationship with the intervention. Serious adverse events include death, life-threatening events, hospitalization or prolongation of hospitalization, significant disability, serious self-harm or suicidal behavior, and other events judged by investigators to be serious.

5. Timing of Assessments

The planned assessment time points are:

1. Screening.
2. Baseline.
3. Week 3.
4. Week 5.
5. Week 8.
6. 3-month follow-up.
7. 6-month follow-up.
8. 12-month follow-up.
9. 24-month follow-up.

Week 8 is the primary endpoint. Week 3 and week 5 are used primarily to capture early changes in mechanistic variables for longitudinal mediation analyses. The 3-, 6-, 12-, and 24-month follow-up assessments are used to evaluate maintenance of effects, long-term safety, and cost-effectiveness.

6. Sample Size

The study plans to recruit approximately 500 participants with GERD, randomized in a 1:1 ratio to the MIED group and the waitlist control group, with approximately 250 participants per group. Assuming an attrition rate of 15% to 25%, approximately 400 participants are expected to complete the study, with approximately 200 participants per group.

The sample size is based on the between-group difference in change in GERD-HRQL total score from baseline to week 8. Because previous mindfulness intervention studies in GERD have had small samples and unstable effect estimates, a conservative small-to-moderate standardized effect size assumption is used. Assuming a standardized effect size of approximately 0.30 to 0.35, a two-sided alpha level of 0.05, and power of 80% to 90%, approximately 175 to 235 participants per

group would be required. The planned sample size allows for attrition, multicenter implementation, long-term follow-up, and exploratory mediation and moderation analyses.

No interim efficacy analysis is planned.

7. Randomization and Blinding

Participants who meet eligibility criteria and complete baseline assessment will be randomly assigned in a 1:1 ratio to the MIED group or the waitlist control group. Randomization will be stratified by study center and implemented using a computer-generated randomization sequence prepared by an independent statistician.

Allocation concealment will be maintained using a central randomization system or a randomization list maintained by an independent data manager. Recruitment staff and baseline assessors will not know the allocation before enrollment and baseline assessment are completed.

Because of the nature of the intervention, participants and intervention-support staff cannot be blinded. Outcome assessors, data managers, and statistical analysts will remain blinded whenever possible. For the primary analysis, group allocation will be coded until the primary analysis dataset and analysis scripts have been finalized.

8. Analysis Populations

8.1 Intention-to-Treat Population

The intention-to-treat (ITT) population will include all randomized participants, analyzed according to their assigned group, regardless of intervention adherence, protocol deviations, withdrawal from the intervention, or missing follow-up data.

The ITT population will be the primary analysis population for efficacy analyses.

8.2 Per-Protocol Population

The per-protocol (PP) population will include randomized participants who:

1. Meet eligibility criteria.
2. Complete the primary endpoint assessment at week 8.
3. Have no major protocol violations.
4. Are assigned to the MIED group and complete at least 40 of the 49 daily sessions, or are assigned to the waitlist control group and do not receive any structured mindfulness or MIED-related intervention during the 49-day observation period.

The PP analysis will be supplementary and supportive. It will not replace the ITT analysis.

8.3 Safety Population

The safety population will include all randomized participants who complete at least one post-baseline safety assessment. Participants will be analyzed according to the intervention actually received for safety analyses. If a participant receives any MIED course content, they will be included in the MIED safety group for safety summaries.

8.4 Mediation Analysis Population

The mediation analysis population will include participants with baseline data, at least one early mechanistic assessment at week 3 or week 5, and at least one post-baseline outcome assessment. Mediation analyses will be exploratory and will use all available data under appropriate missing-data assumptions.

9. General Statistical Principles

All statistical tests will be two-sided unless otherwise specified. The nominal significance level for the primary outcome will be 0.05. Confidence intervals will be reported at the 95% level.

Continuous variables will be summarized using mean, standard deviation, median, interquartile range, minimum, maximum, and number of nonmissing observations. Categorical variables will be summarized using frequencies and percentages.

Baseline characteristics will be summarized by randomized group. No formal significance testing of baseline differences will be used to judge randomization success. Clinically important imbalances may be adjusted for in sensitivity analyses.

All analyses will be conducted using validated statistical software, such as R, SAS, Stata, or equivalent software. The software name and version will be reported in the final statistical report.

10. Data Preparation and Derived Variables

10.1 Scoring of Questionnaires

All questionnaire scores will be calculated according to their respective scoring manuals or published scoring rules. If item-level missingness occurs within a scale, the scoring rule recommended by the scale developer will be followed. If no formal rule is available, the following general rule will be applied:

1. If at least 80% of items in a scale are completed, the participant's mean item score will be imputed for missing items within that scale, and the total score will be prorated.
2. If more than 20% of items in a scale are missing, the total score will be treated as missing.

For multi-dimensional scales, the same rule will be applied within each subscale.

10.2 Change Scores

For descriptive purposes, change scores will be calculated as:

Change score = post-baseline score − baseline score.

For GERD-HRQL, GERD-Q/GerdQ, RSI, OASIS, ODSIS, PHQ-9, GAD-7, ISI, CPSS, DIS, EEBS, and BEAQ, negative change generally indicates improvement. For SF-8, CD-RISC, SWLS, FFMQ, ELS, DTS, and CFI, positive change generally indicates improvement, unless otherwise specified by scoring rules.

Primary inferential analyses will rely on mixed-effects models rather than simple change-score tests.

10.3 Center Variable

Study center will be included as a fixed effect in the primary model. If a center recruits very few participants, centers may be combined into an “other centers” category for adjusted analyses, or center may be treated as a random effect in sensitivity analyses.

10.4 Baseline Value

The baseline value of the outcome will be included as a covariate in primary and secondary outcome models where appropriate. If the model includes repeated baseline and follow-up observations, the role of baseline will be handled consistently by including baseline as part of the outcome vector and modeling time, or by adjusting for baseline value in a constrained longitudinal data analysis framework. The final modeling specification will be documented before database lock.

11. Primary Outcome Analysis

11.1 Primary Estimand

The primary estimand is the between-group difference in mean change in GERD-HRQL total score from baseline to week 8, comparing MIED plus treatment as usual with waitlist control plus treatment as usual in the ITT population.

The treatment condition contrast is MIED versus waitlist control. The primary time point is week 8. The primary intercurrent events include discontinuation of the intervention, poor adherence, or receipt of additional non-protocol psychological intervention. These events will be handled using a treatment-policy strategy in the ITT analysis; participants will be analyzed according to randomized group regardless of adherence or intercurrent events.

11.2 Primary Model

The primary analysis will use a linear mixed-effects model or mixed model for repeated measures (MMRM) including repeated GERD-HRQL observations.

The model will include:

1. Randomized group.
2. Time.
3. Group-by-time interaction.
4. Baseline GERD-HRQL score.
5. Study center.
6. Prespecified covariates if needed, such as age, sex, baseline GERD-Q score, baseline PHQ-9 or GAD-7 score, and proton pump inhibitor use.

Participant will be included as a random intercept. If model fit allows, a random slope for time may be explored in sensitivity analyses. An unstructured covariance matrix will be preferred for repeated measures. If the model fails to converge, alternative covariance structures, such as compound symmetry or autoregressive covariance, will be considered.

The primary treatment effect will be the estimated between-group difference in GERD-HRQL change from baseline to week 8. Estimated marginal means, between-group difference, standard error, 95% confidence interval, and P value will be reported.

11.3 Interpretation

A negative between-group difference favoring MIED will indicate that the MIED group experienced a greater reduction in GERD-HRQL score than the waitlist control group, corresponding to greater improvement in reflux-related quality of life.

12. Sensitivity Analyses for the Primary Outcome

The following sensitivity analyses will be conducted for the primary outcome:

1. **Per-protocol analysis:** The primary model will be repeated in the PP population.
2. **Multiple imputation:** Missing GERD-HRQL outcomes will be imputed under a missing-at-random assumption using multiple imputation by chained equations. The imputation model will include treatment group, time, baseline GERD-HRQL, study center, baseline demographic and clinical variables, adherence indicators, and available outcome values.
3. **Complete-case analysis:** Participants with observed GERD-HRQL at baseline and week 8 will be analyzed.
4. **Alternative covariance structure:** The mixed model will be repeated using alternative covariance structures if relevant.
5. **Center as random effect:** The model will be repeated treating center as a random effect, if data structure permits.
6. **Adjusted model:** The model will be adjusted for additional clinically important baseline covariates, including age, sex, body mass index, disease duration, baseline GERD-Q score, baseline emotional distress, and proton pump inhibitor use.
7. **Non-normality sensitivity:** If model residuals show substantial deviation from normality, robust standard errors, transformation, or generalized linear mixed models will be considered.
8. **Missing-not-at-random sensitivity:** Pattern-mixture models or delta-adjustment analyses will be used to explore departures from the missing-at-random assumption if missingness is substantial.

The primary conclusion will be based on the prespecified ITT mixed-effects model. Sensitivity analyses will be used to evaluate robustness.

13. Secondary Outcome Analyses

13.1 Continuous Secondary Outcomes

Continuous secondary outcomes will be analyzed using linear mixed-effects models similar to the primary model. These outcomes include:

1. GERD-Q/GerdQ.
2. RSI.
3. OASIS.
4. ODSIS.
5. PHQ-9.
6. GAD-7.
7. SF-8.
8. ISI.
9. CPSS.
10. CD-RISC.
11. SWLS.
12. EQ-5D-5L utility score.

For each outcome, the model will include group, time, group-by-time interaction, baseline value of the outcome, and study center. Participant will be included as a random effect. The main secondary time point will be week 8. Long-term treatment effects will be estimated at 3, 6, 12, and 24 months.

Results will be reported as estimated marginal means, between-group differences, 95% confidence intervals, and P values.

13.2 Binary and Categorical Secondary Outcomes

Binary outcomes may include:

1. Clinically meaningful improvement in GERD-HRQL.
2. GERD symptom response.
3. GERD symptom remission.
4. Presence of any adverse event.
5. Presence of any serious adverse event.
6. Temporary suspension of intervention due to safety concerns.
7. Dropout before week 8.

Binary outcomes will be analyzed using logistic regression or generalized linear mixed models with logit link, including group, study center, and relevant baseline covariates. Results will be reported as odds ratios, risk differences, relative risks where appropriate, 95% confidence intervals, and P values.

Definitions of clinically meaningful improvement and symptom remission will be prespecified before database lock. If no validated minimal clinically important difference is available for a specific outcome, responder analyses will be treated as exploratory.

13.3 Long-Term Maintenance of Effects

Long-term maintenance will be evaluated using repeated-measures models including all follow-up time points. The key parameters will be group-by-time interactions at 3, 6, 12, and 24 months.

Because the waitlist group may receive MIED after completion of long-term follow-up according to the protocol, the primary long-term comparisons will remain uncontaminated if the course is offered only after 24 months. If any waitlist participants receive MIED or similar interventions before 24 months, this will be recorded and handled in sensitivity analyses.

---

14. Multiplicity Control

The primary outcome analysis will be tested at a two-sided alpha level of 0.05 without adjustment because there is a single primary outcome and a single primary endpoint.

Secondary outcomes will be interpreted hierarchically and cautiously. The following approach will be used:

1. Key secondary outcomes at week 8 will be considered supportive evidence.
2. P values for multiple secondary outcomes will be reported both unadjusted and adjusted using the Benjamini-Hochberg false discovery rate procedure within the family of key secondary outcomes.
3. Long-term follow-up analyses, mediation analyses, subgroup analyses, and dose-response analyses will be considered exploratory unless otherwise prespecified.
4. No confirmatory claims will be based solely on unadjusted secondary or exploratory analyses.

15. Missing Data

15.1 Description of Missing Data

The amount and pattern of missing data will be summarized by treatment group and time point. Reasons for missing data will be tabulated where available.

Baseline characteristics of participants with and without missing week-8 primary outcome data will be compared descriptively to assess potential missing-data mechanisms.

15.2 Primary Handling of Missing Data

The primary mixed-effects model will use all available outcome data and maximum likelihood estimation, assuming data are missing at random conditional on observed data.

15.3 Multiple Imputation

If missingness for the primary outcome at week 8 exceeds 5%, multiple imputation will be used as a sensitivity analysis. The imputation model will include:

1. Randomized group.
2. Study center.
3. Baseline GERD-HRQL.
4. Available GERD-HRQL values at all time points.
5. Baseline GERD-Q/GerdQ.
6. Baseline anxiety and depression measures.
7. Age, sex, body mass index, disease duration, and medication use.
8. Adherence indicators, where appropriate.

The number of imputations will be at least equal to the percentage of incomplete cases and no fewer than 20.

15.4 Missing-Not-at-Random Sensitivity

If missingness is substantial or differential between groups, missing-not-at-random sensitivity analyses will be conducted using pattern-mixture or delta-adjusted models. These analyses will examine whether conclusions remain stable under plausible unfavorable assumptions for missing outcomes.

16. Adherence and Dose-Response Analyses

16.1 Adherence Measures

Adherence indicators in the MIED group will include:

1. Number of completed daily sessions.
2. Course completion rate.

3. Daily check-in rate.
4. Total course viewing time.
5. Total guided practice time, if available.
6. Assignment submission rate.
7. Login frequency.
8. Number of assistant contacts or reminders.

16.2 Course Completion

Course completion will be defined as completion of at least 40 of the 49 daily sessions, unless a different threshold is prespecified before database lock.

16.3 Dose-Response Analyses

Exploratory dose-response analyses will examine associations between adherence indicators and changes in primary and secondary outcomes within the MIED group. Linear regression, mixed-effects models, or generalized additive models may be used depending on the distribution of adherence variables and outcomes.

Because adherence is not randomized, dose-response analyses will be considered exploratory and interpreted cautiously. Models may adjust for baseline symptom severity, baseline emotional distress, age, sex, and study center.

17. Mediation Analyses

17.1 Conceptual Model

The hypothesized pathway is that MIED improves GERD symptoms and reflux-related quality of life through changes in psychological mechanisms, including mindfulness, engaging in life, distress tolerance, discomfort intolerance, experiential avoidance, excessive emotional behaviors, and cognitive flexibility.

The primary mediation framework will prioritize temporal ordering:

1. Treatment assignment: randomized group.
2. Early mediator change: baseline to week 3 and/or baseline to week 5.
3. Outcome change: baseline to week 8, and optionally long-term follow-up outcomes.

17.2 Candidate Mediators

Candidate mediators include:

1. Change in FFMQ.
2. Change in ELS.
3. Change in DTS.
4. Change in DIS.
5. Change in EEBS.
6. Change in BEAQ.
7. Change in CFI.
8. Change in OASIS and/or ODSIS as emotional distress mediators.

17.3 Outcomes for Mediation

The main mediation outcomes will be:

1. GERD-HRQL total score at week 8, adjusted for baseline GERD-HRQL.
2. GERD-Q/GerdQ at week 8, adjusted for baseline GERD-Q/GerdQ.

Exploratory mediation analyses may use 3-, 6-, 12-, and 24-month outcomes.

17.4 Statistical Approach

Longitudinal mediation models will be used where possible. The following approaches may be applied:

1. Product-of-coefficients mediation models.
2. Structural equation modeling.
3. Multilevel mediation models accounting for repeated measures and study center.
4. Bootstrap confidence intervals for indirect effects.

Models will adjust for baseline levels of mediator and outcome, randomized group, study center, and clinically important baseline covariates. Indirect effects will be reported with 95% bootstrap confidence intervals.

Mediation analyses will be considered exploratory and hypothesis-generating. Causal interpretations will be made cautiously.

18. Subgroup and Moderation Analyses

Exploratory moderation analyses will examine whether treatment effects differ by:

1. Baseline GERD symptom severity.
2. Baseline GERD-HRQL severity.
3. Baseline anxiety severity.
4. Baseline depressive symptom severity.
5. Baseline perceived stress.
6. Disease duration.
7. Sex.
8. Age.
9. Body mass index.

10. Proton pump inhibitor use.
11. Insomnia severity.
12. Study center.

Moderation will be tested by adding group-by-time-by-moderator interaction terms to the mixed-effects model. Continuous moderators may be analyzed continuously and, if clinically useful, categorized for descriptive presentation.

All subgroup and moderation analyses will be exploratory. No definitive claims of differential treatment effect will be made without strong statistical and clinical evidence.

19. Cost-Effectiveness Analysis

19.1 Perspective

Economic evaluation will be conducted from both the health care system perspective and the societal perspective, if data permit.

19.2 Cost Measures

Costs will be derived from TIC-P and supplementary health care resource use data. Cost categories may include:

1. Outpatient visits.
2. Emergency visits.
3. Hospitalizations.
4. Diagnostic examinations.
5. Medication costs.
6. Psychological intervention delivery costs.
7. Transportation costs.

8. Productivity loss.

9. Informal caregiving costs.

19.3 Health Utility and QALYs

Health utility will be derived from EQ-5D-5L using an appropriate value set. Quality-adjusted life years will be estimated using the area-under-the-curve method across assessment time points.

19.4 Incremental Cost-Effectiveness

Incremental cost-effectiveness ratios will be calculated as:

ICER = (Cost_MIED − Cost_Control) / (Effect_MIED − Effect_Control).

Effect measures may include:

1. QALYs.

2. Change in GERD-HRQL.

3. Clinically meaningful improvement in GERD-HRQL.

Uncertainty will be assessed using nonparametric bootstrapping. Cost-effectiveness acceptability curves will present the probability that MIED is cost-effective across different willingness-to-pay thresholds.

Cost-effectiveness analyses will be exploratory unless otherwise specified in the final registration.

20. Safety Analyses

Safety analyses will be conducted in the safety population. Adverse events will be summarized by treatment group.

The following will be reported:

1. Number and percentage of participants with any adverse event.

2. Number and percentage of participants with serious adverse events.
3. Number and percentage of participants with GERD symptom worsening.
4. Number and percentage of participants with alarm symptoms.
5. Number and percentage of participants with psychological crisis.
6. Number and percentage of participants requiring referral.
7. Number and percentage of participants temporarily suspending or discontinuing the intervention due to adverse events.

Adverse events will be categorized by type, severity, relatedness to intervention, action taken, and outcome. No formal hypothesis testing is planned for safety outcomes unless event rates allow meaningful comparison. Between-group comparisons of adverse event incidence may be presented as risk differences, relative risks, or odds ratios with 95% confidence intervals.

21. Protocol Deviations

Protocol deviations will be classified before database lock as major or minor.

21.1 Major Protocol Deviations

Major deviations may include:

1. Randomization of an ineligible participant.
2. Incorrect group assignment or allocation error.
3. Receipt of a structured mindfulness or MIED-related intervention by a waitlist participant during the 49-day observation period.
4. Major violation of assessment timing for the primary endpoint.
5. Missing primary endpoint due to avoidable procedural error.
6. Unblinding of outcome assessor before primary endpoint assessment.

7. Serious data integrity issue.

21.2 Minor Protocol Deviations

Minor deviations may include:

1. Delayed questionnaire completion within an acceptable assessment window.
2. Minor missing items in questionnaires.
3. Missed assistant reminder.
4. Temporary technical issues with the mini-program not affecting intervention completion.

All deviations will be summarized by group. Major deviations will determine inclusion in the PP population but will not affect inclusion in the ITT population.

22. Assessment Windows

The following assessment windows will be applied unless otherwise specified:

1. Baseline: before randomization.
2. Week 3: target day 21, acceptable window ±5 days.
3. Week 5: target day 35, acceptable window ±5 days.
4. Week 8: target day 56, acceptable window ±10 days.
5. 3 months: target 3 months after baseline, acceptable window ±21 days.
6. 6 months: target 6 months after baseline, acceptable window ±30 days.
7. 12 months: target 12 months after baseline, acceptable window ±45 days.
8. 24 months: target 24 months after baseline, acceptable window ±60 days.

If multiple assessments occur within a window, the assessment closest to the target date will be used for the primary analysis. Alternative rules will be documented before database lock if required.

23. Data Monitoring, Database Lock, and Blinding of Analysis

Before database lock, the following steps will be completed:

1. Verification of participant eligibility.
2. Verification of randomization records.
3. Completion of data cleaning.
4. Resolution of data queries.
5. Review and classification of protocol deviations.
6. Verification of outcome scoring.
7. Finalization of analysis populations.
8. Finalization of the statistical analysis scripts whenever possible.

The primary analysis will be performed with blinded group labels when feasible. Group labels will be unblinded only after analysis datasets and primary analysis code have been finalized.

24. Presentation of Results

The final statistical report will include:

1. CONSORT flow diagram.
2. Recruitment and retention summary.
3. Baseline characteristics by group.
4. Intervention adherence and exposure.
5. Primary outcome results.

6. Secondary outcome results.
7. Long-term follow-up results.
8. Mediation and moderation analyses.
9. Safety results.
10. Cost-effectiveness results.
11. Sensitivity analyses.
12. Protocol deviations.
13. Missing-data summaries.

Tables and figures will include estimated effects, confidence intervals, and number of participants contributing data at each time point.

25. Planned Tables and Figures

Table 1. Baseline Characteristics

Baseline demographic and clinical characteristics by randomized group.

Table 2. Intervention Adherence

Course completion, check-in rate, viewing time, practice time, assignment submission, and assistant contact frequency.

Table 3. Primary Outcome

GERD-HRQL observed means, model-estimated marginal means, between-group differences, 95% confidence intervals, and P values.

Table 4. Secondary Outcomes at Week 8

Model-estimated group differences for GERD-Q/GerdQ, RSI, OASIS, ODSIS, PHQ-9, GAD-7, SF-8, ISI, CPSS, CD-RISC, SWLS, and EQ-5D-5L.

Table 5. Long-Term Outcomes

Between-group differences at 3, 6, 12, and 24 months.

Table 6. Adverse Events

Adverse events and serious adverse events by group.

Table 7. Mediation Analyses

Indirect effects through mechanistic variables.

Table 8. Cost-Effectiveness Results

Costs, QALYs, ICERs, and uncertainty intervals.

Figure 1. CONSORT Flow Diagram

Participant flow from screening to analysis.

Figure 2. Trajectory of GERD-HRQL Over Time

Estimated marginal means by group and time point.

Figure 3. Trajectories of Key Secondary Outcomes

Selected symptom and psychological outcomes over time.

Figure 4. Mediation Model

Path diagram showing intervention, early mediator change, and outcome change.

Figure 5. Cost-Effectiveness Acceptability Curve

Probability of cost-effectiveness across willingness-to-pay thresholds.

26. References

Barlow DH. Unified Protocol for Transdiagnostic Treatment of Emotional Disorders. Oxford University Press; 2011.

Bellg AJ, Borrelli B, Resnick B, et al. Enhancing treatment fidelity in health behavior change studies. Health Psychology. 2004.

Chan AW, et al. SPIRIT 2025 statement: updated guideline for protocols of randomized trials. 2025.

Gupta SK. Intention-to-treat concept: A review. Perspectives in Clinical Research. 2011.

Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D. Quality of Life Research. 2011.

Laird NM, Ware JH. Random-effects models for longitudinal data. Biometrics. 1982.

Little RJA, Rubin DB. Statistical Analysis with Missing Data. Wiley; 2019.

MacKinnon DP. Introduction to Statistical Mediation Analysis. Routledge; 2012.

Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behavior Research Methods. 2008.

Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. BMJ. 2010.