SERENE

Supporting Evidence-based Responses to Emotional Needs in Emphysema

Statistical Analysis Plan

Principal Investigator	*Joanna Hart, MD, MSHP*
Study Name	*Supporting Evidence-based Responses to Emotional Needs in Emphysema (SERENE)*
Version Number	1.0
Date	9/5/2024

Palliative & Advanced Illness Research (PAIR) Center
University of Pennsylvania

Revision history

Date of change	Protocol version	Updated SAP version	Section number changed	Changes made -- Reasons for the Change

Contents

Background and rationale

Families are an important part of patients' chronic illness experiences, including for diseases like chronic obstructive pulmonary disease (COPD), and especially among patients who may experience structural barriers to health. This study will identify how an evidence-based Coping Skills Training (CST) program for families, in which patients and their caregivers partner to manage and overcome disease-related distress, reduces depression and anxiety among outpatients with COPD. The findings will provide a better understanding of how social relationships affect patients' health and will help health systems design and implement efficient strategies for family-centered chronic illness care, particularly for families experiencing health disparities.

Objectives and hypotheses

Our central hypothesis is that improving patients' coping skills and caregivers' ability to coach patients in use of these skills will improve patients' outcomes.

The primary objective of the proposed trial is to identify key mechanisms through which the family-focused CST program reduces patients' psychological distress, which can be applied to other disease models and interventions. The project will also assess heterogeneity in the efficacy of CST at reducing distress among patients who differ by key characteristics. Further, it will identify patient, caregiver, clinician, and contextual factors associated with program uptake and completion, emphasizing relevancy and modifiable structural barriers to participation.

Aim 1: Identify mechanisms through which CST reduces patient distress.

Aim 2: Elucidate causes of heterogeneity in the efficacy of the CST program.

Aim 3: Identify barriers and facilitators to uptake of the CST program.

Trial design

Eligibility

Patients will be recruited from primary care and specialty pulmonary clinics at the University of Pennsylvania (Penn) and Henry Ford (HFH) Health Systems. The selected clinics primarily serve Black, low-income, and rural communities and patient populations.

Inclusion criteria:

- All participants must be at least 18 years of age
- Patients must have a documented diagnosis of COPD and confirmation of obstructive lung disease by spirometry (American Thoracic Society/European Respiratory Society guidelines) or radiology (imaging report indicating emphysematous changes)
- Patients must score ≥8 during baseline screening using the PHQ-8
- Patients must identify an adult caregiver to participate with them
- Patients and caregivers must have the ability to access a telephone or videoconferencing call up to once weekly (for approximately 30 minutes) for 12 weeks
- Spoken proficiency in Spanish and/or English

Exclusion criteria:

- Patient or caregiver has significant dementia or other cognitive decline
- Documentation in the electronic health record that the COPD diagnosis has not yet been disclosed to the patient
- Patient is under the ongoing care of a licensed behavioral health clinician at the time of enrollment
- Patient requires immediate referral to specialized behavioral health management based on psychological risk assessment conducted by the study team

Arms and timepoints

We will conduct a 2-arm randomized trial of the CST program among 375 patient-caregiver dyads in which the patient is experiencing psychological distress. We will randomize dyads in a 2:1 ratio to participate in a CST program or minimal enhanced support. We will collect patient-reported outcomes at 5, 7, 12, 14, 24, and 26 weeks

following program initiation and clinical outcomes via the electronic health record up to 12 months after enrollment.

Arm 1: Coping Skills Training telehealth program. CST is a dyad-delivered intervention with content directed at both the patient and the caregiver. CST focuses on modifying the cognitive factors (i.e., expectations and attitudes) and behaviors (i.e., relationship functioning, physical activity, health promotion) that affect physical symptoms and disability in chronic illness. This CST program was specifically developed for use with patients with chronic lung disease and for the patient-caregiver dyad. The intervention focuses on improving the dyad's response to stress and imbalances in the relationship, communication between patients and caregivers, and caregivers' ability to coach patients to use positive coping strategies.

Arm 2: Minimal enhanced support (COPD Education program). Patients with COPD and their caregivers receive little to no support or management of psychological distress. Our minimal enhanced support augments usual care, providing an attention control for comparison. Dyads will receive 12 10-minute weekly calls during which trained research staff will provide scripted support without any specific psychoeducational techniques. Participants will receive a coping self-management booklet developed by The COPD Foundation, our study collaborator, given CST participants will receive mailed materials. This comparison arm is necessary to our aims interrogating the mechanisms through which CST works because patients with COPD often experience social isolation that may be affected by increased contact regardless of the content.

Randomization

We will randomize patient-caregiver dyads, stratified by recruitment site to one of two arms using random number generation. Two-thirds will be randomized to participate in the CST program and one-third to the COPD Education program (minimal enhanced support). Because many patients with COPD in the United States experience health disparities and may lack the technological resources to engage in videoconferencing, we will offer all study programs as telephone (i.e., the modality of prior testing) or videoconferencing calls.

Primary outcome and analytic method

Primary outcome

The primary efficacy outcome is patient psychological distress, as measured by the Patient Health Questionnaire-9 (PHQ-9).

Setting this as the primary outcome allows us to test mechanisms leading to its improvement. The PHQ-9 is a 9-item, self-report questionnaire with a 4-point rating for the burden of each symptom over the prior 2 weeks, with higher scores indicating more depressive symptoms. The PHQ-9 is highly sensitive and specific for major depression, using a score cut-off of 10 (out of 27 possible). The PHQ-9 is sensitive to change. We will measure the PHQ-9 at baseline, after 7 weeks, after 14 weeks, and after 26 weeks (i.e., approximately 6 months). The primary outcome is at 14 weeks.

Primary analytic sample

Our primary analytic approach will be to conduct intention-to-treat (ITT) analyses, including all patients randomized in the trial, even if they do not participate in the offered program.

Primary analysis

Our ITT analyses will use truncated regression, due to the implementation of the Patient Health Questionnaire-9 (PHQ-9) with a cut-off point at 10 out of 27. This method is specifically chosen to address the distributional characteristics imposed by the PHQ-9 scoring system, where responses are bounded and thus not all potential values are observed in the sample. Truncated regression will allow for more accurate estimation of effects by effectively dealing with the left-censored nature of the data, ensuring that the analyses accurately reflect the impact of interventions on respondents scoring above the established threshold. With patients' baseline PHQ-9 score included in the model, this approach will facilitate a comparison of the efficacy of the intervention in achieving reduced distress among all randomized participants. Those effect sizes will be diluted, as some patients will not use the offered CST telehealth program. We will model the health

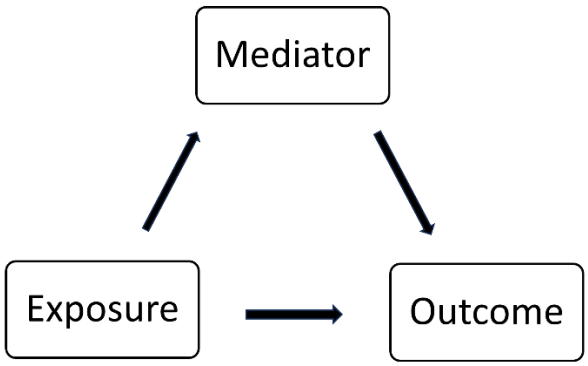
system from which the patient was recruited as a fixed effect, thereby accounting for confounding by health system and adjusting variance estimates for clustering of participants within health systems. We will also adjust for covariates that are listed below.

Mediation analysis

We will evaluate whether differences in five outcomes mediate the observed differences in PHQ-9 scores based on treatment arm, drawing on existing evidence and our conceptual model of CSTs theoretical mechanisms of action. We will measure these constructs at baseline, after 5 weeks, after 12 weeks (i.e., immediately after the final program session), and after 24 weeks to facilitate causal testing.

Causal mediation methods provide a powerful means of identifying mechanistic pathways through which interventions or exposures affect outcomes[1]. We will estimate the causal pathway from a single exposure (i.e., treatment arm), to a temporally subsequent mediator (i.e., the five mediator outcome variables listed below), and ultimately to an outcome (i.e., psychological distress). We will model receipt of the CST program as a binary variable and mediating and psychological distress outcome scores as count variables. The 12-week measurement of the mediation variable corresponding to the 14-week PHQ-9 score will be entered into the primary model. These models will allow us to differentiate and test the CST program's direct and indirect (i.e., mediated) effect on patients' psychological distress. We will account for covariance among the mediators. In a final model, we will retain only those mediators significantly associated with randomization to the CST program or improvement in patients' PHQ-9 scores. Under the standard assumption of ignorability of the treatment assignment (i.e., the treatment is effectively randomized given baseline covariates), plus the additional assumption of sequential ignorability (i.e., no unmeasured confounding between the mediator and the outcome, given the treatment and covariates), we can interpret the mediated effect as indicative of a causal pathway[2].

Our five potential mediators we will test include (a) family relationship quality, as reported by patients and caregivers using the Family Emotional Involvement and Criticism Scale (FEICS); (b) patient and (c) caregiver self-efficacy, measured with the General Self-efficacy Scale (GSE); (d) patient loneliness, measured by the UCLA Loneliness Scale; and (e) caregiver psychological distress, measured by the PHQ-9 and Generalized Anxiety Disorder-7 (GAD-7).



Sensitivity analysis

Since in reality the mediator is not randomized, we plan to perform a sensitivity analysis by avoiding sequential ignorability assumption. After computing the average causal mediation effects (ACME), the sensitivity parameter p (correlation between the residuals of mediator and outcome) will be varied over a range to examine how ACME changes. The analysis will be performed using 'mediation' package in R.

Adjustment for covariates

We will conduct all analyses with adjustment for prespecified, patient-level covariates, including the caregiver's relationship to patient and the recruiting health system.

Factors for adjustment	Variable coding or definition
Caregiver's relationship to patient	Categorical
Recruiting health system	Categorical

Secondary analytic sample

We will repeat the primary analysis using a per-protocol analysis that includes patients who participate in ≥50% of the CST program sessions. Doing this will provide insight into the efficacy of the CST program under optimal adherence conditions, allowing us to assess the potential benefits of the intervention for individuals who engage with the program at a significant level. This analysis aims to complement the intention-to-treat analysis by offering a clearer understanding of the intervention's effect among adherent participants, thereby offering a more nuanced perspective on its practical effectiveness.

Secondary analyses

For treatment heterogeneity, we will test statistical interactions between pre-selected patient characteristics and the two interventions in truncated regression, as we are interested in differential effects among those who use the program. Patients who complete at least 50% (6 sessions) of the CST program will be included. Our outcomes will be the same as those for our primary analysis. We will evaluate each interaction term separately in a model adjusted for all main effects and with health system entered as a fixed effect to estimate stratum specific effects of each characteristic on the interventions' efficacy. We will also report results from a fully adjusted model that retains all significant interaction terms. We will limit our search for heterogeneity of treatment effects to variables selected based on a priori hypotheses and will finalize the list with the Community Advisory Board (CAB) prior to launch. These variables include (a) race, (b) income, (c) gender, and (d) rurality of residence.

Potential effect modifiers	Variable coding or definition
Race	Categorical
Income	Continuous
Gender	Binary
Rurality of residence	Categorical

We will conduct another secondary analysis by applying linear regression to the primary outcome variables, substituting the truncated regression method used in the primary analysis. This method variation is intended to explore the robustness of our findings across different statistical model.

Additionally, we plan to explore the potential dosing effect of the CST by recoding the intervention variable into a fraction that represents the proportion of CST sessions each patient attended. This approach allows the hypothesis that the impact of the CST program may vary in a dose-dependent manner, with the number of sessions attended potentially influencing the magnitude of the intervention's effect. By quantifying session attendance as a continuous variable, this analysis aims to assess how varying levels of engagement with the CST program affect the primary outcome measures, providing valuable insights into the importance of session adherence for maximizing therapeutic benefits.

Secondary outcomes

Secondary outcome measures

Outcome measure	Variable coding or definition
Patient psychological distress (PHQ-9)	Continuous, measured at 7 and 26 weeks (primary outcome at 14 weeks)
Patient anxiety (GAD-7)	Continuous, measured at 7, 14 and 26 weeks
Health-related quality of life (SGRQ)	Continuous, measured at 7, 14 and 26 weeks
Health-related quality of life (mMRC Dyspnoea Scale)	Ordinal, measured at 7, 14 and 26 weeks
Health-related quality of life (COPD Assessment Test™)	Continuous, measured at 7, 14 and 26 weeks
Time to next COPD-related hospitalization following enrollment	Continuous
All-cause mortality following enrollment	Binary, measured over 12 months
Caregiver burden (Zarit Caregiver Burden Inventory)	Continuous, measured at 7, 14 and 26 weeks
Family relationship quality (Family Emotional Involvement and Criticism Scale)	Continuous, measured at 5, 12 and 24 weeks
Self-efficacy (General Self-efficacy Scale)	Continuous, measured at 5, 12 and 24 weeks
Loneliness (UCLA Loneliness Scale)	Continuous, measured at 5, 12 and 24 weeks
Caregiver psychological distress (PHQ-9)	Continuous, measured at 5, 12 and 24 weeks
Caregiver anxiety (GAD-7)	Continuous, measured at 5, 12 and 24 weeks

Secondary outcome analysis

Patient- and caregiver-reported secondary outcomes will be similarly analyzed using hierarchical linear regression. Analyses for the time to COPD-related hospitalization outcome will use time to event survival analysis, while analyses using all-cause mortality as an outcome will be analyzed using logistic regression. Ordinal outcomes will be analyzed using ordinal logistic regression. If the assumption of normality is violated for any outcome, we will use another appropriate model (e.g., quantile regression).

Sample size and statistical power calculations

Calculating power to detect the indirect, or mediated, proportion of a total effect requires simulation. In an R statistical simulation using 250 iterations, Dr. Michael Harhay (co-investigator) confirmed that a sample size of 300 dyads would result in at least 80% power to detect mediators that account for a PHQ-9 score difference of at least 2.75 points and a direct effect of the intervention on PHQ-9 scores of at least 3.25 points. Combined, these differences exceed the clinically significant difference for PHQ-9. Enrolling 375 dyads (i.e., approximately 250 in the CST arm and 125 in the enhanced usual care arm) equally across the two health systems allows for 20% attrition (i.e., a final sample size of 300).

Additional analyses

We will summarize quantitative data of the use of the CST program using descriptive statistics. We will implement generalized linear regression models to compare rates of uptake based on pre-specified patient characteristics and delivery modality, with health system modeled using a fixed effect.

Qualitative data will be used to conduct an analysis of barriers and facilitators to uptake of the CST program. All interviews will be digitally recorded and transcribed verbatim

by a professional transcription company. We will use the qualitative data analysis software ATLAS.ti to establish a database for the qualitative study. PI Hart is experienced in qualitative methods, and Dr. Carmen Alvarez (co-investigator) has expertise in implementation science, particularly among individuals and groups experiencing health disparities. Trained research staff members, with oversight from Drs. Hart and Alvarez, and input from the CAB, will identify emerging themes and develop a codebook that includes two pre-specified Consolidated Framework for Implementation Research (CFIR) codes: (1) acceptability and feasibility and (2) fidelity. We will catalog key concepts within the transcripts and staff field notes using the codebook, with ≥2 coders reviewing each to ensure reliability. The coding team will meet to resolve uncertainties and iteratively refine the codebook. This process will continue until the codebook is static and all transcripts and field notes are coded. We will write analytic memos summarizing emergent patterns or concepts. In collaboration with the CAB, we will then use an abductive approach to generate explanations for observed trends not explained by existing literature, evaluate these explanations, and iteratively revise the proposed explanations until we arrive at a final theory. Abduction is a creative inferential analysis process with which Dr. Hart is experienced.

Fidelity

Fidelity will be independently assessed for both the CST and COPD Education arm using methods based on the NIH Behavior Change Consortium. We will record a random sample of 20% of sessions. Random selection will be stratified according to site of enrollment, intervention arm, and time since study launch to capture “drift” over time. At least two independent, trained assessors will rate each session’s fidelity based on the session protocol and key elements pre-defined for each session. Assessors will engage in monthly supervision by Dr. Virginia O’Hayer (co-investigator) to ensure consistency in ratings over time. Study staff administering the sessions will also document each session’s duration in minutes and their field notes of discussions during clinical oversight meetings, including the details of and rationale for adaptations made for participants.

Approach to missing data

We have chosen the primary endpoint at 14 weeks after enrollment and designed the study based on prior retention success to minimize missingness of our key outcomes.

However, missing data for the primary or mediator outcomes can occur due to non-response or patient death, as these outcomes are patient reported. We will fully describe all data that are missing at each follow-up time point (i.e., at 5, 7, 12, 14, 24, and 26 weeks). Our secondary analyses will use multiple imputation methods that assume data are missing at random by including all baseline characteristics, treatment assignment, and reasons for missingness in the imputation model. This approach allows us to use baseline data and any available follow-up data (e.g., at 12 and 14 weeks) to inform the generation of values for subsequent missing data points (e.g., at 24 and 26 weeks), and properly adjusts for the uncertainty in the resulting imputed values. We will implement the multiple imputation using chain regression method. We plan to use R package 'mice' to carry out this analysis.

Final analytic timeline

Timing of final analysis

Final analysis will occur following completion of the trial in June 2027.

Person performing analysis

Dr. Harhay will prepare reports describing all data using appropriate summary statistics, with estimates of variance and graphical representations of the distributions.

Dr. Lu and Mr. Whitman will be responsible for developing interim and final analysis plans, performing quantitative analysis, and interpreting results for all aims under the direction of Drs. Hart and Harhay.

Interim analyses and safety considerations

Interim analyses and timing

We will not conduct interim data analysis and, therefore, do not specify trial stopping rules.

Statistical software

Analyses will be performed using R and Stata.

References

1. Whittle, R., Mansell, G., Jellema, P. & van der Windt, D. Applying causal mediation methods to clinical trial data: What can we learn about why our interventions (don't) work? *Eur. J. Pain Lond. Engl.* **21**, 614–622 (2017).
2. Lynch, K. G., Cary, M., Gallop, R. & Ten Have, T. R. Causal Mediation Analyses for Randomized Trials. *Health Serv. Outcomes Res. Methodol.* **8**, 57–76 (2008).