Version Date: November 8, 2024
IRB Approval Date: November 20, 2024

wHOPE SAP Cover Page

Official Title: Implementation of a Pragmatic Trial of Whole Health Team vs. Primary Care Group Education to Promote Non-Pharmacological Strategies to Improve Pain, Functioning, and Quality of Life in Veterans

Acronym: wHOPE

Study Type: Interventional

Brief Title: Pragmatic Trial of WHT vs. PC-GE to Promote Non-Pharmacological Strategies to Treat Chronic Pain in Veterans

NCT Number: NCT04330365

Unique Protocol ID: UH3AT009765

DATA ANALYSIS PLAN FOR AIM 1

Implementation of a Pragmatic Trial of Whole Health Team vs. Primary Care Group Education to Promote Non-Pharmacological Strategies to Improve Pain, Functioning, and Quality of Life in Veterans

VERSION 1.0
November 8, 2024

FUNDING AGENCY: The National Center for Complementary and Integrative Health
AWARD NUMBER: UH3AT009765

Prepared by:
Denise Esserman, PhD (study statistician)
Eugenia Buta, PhD (statistician)

The Data Analysis Plan (DAP) is a living document and can be modified (as necessary) up to final data lock and before study unblinding. Only the blinded statistician will modify the DAP.

ClinicalTrials.gov Identifier: NCT04330365

Contents

1. Study Synopsis

Title	Implementation of a Pragmatic Trial of Whole Health Team vs. Primary Care Group Education to Promote Non-Pharmacological Strategies to Improve Pain, Functioning, and Quality of Life in Veterans: wHOPE trial
Study Design	Individually randomized group treatment (IRGT) hybrid type 1 effectiveness-implementation trial
Study Duration	5 years (4 years recruitment; 1 year follow-up)
Trial Sites	San Francisco VA Health Care System; VA Connecticut Healthcare System; VA Portland Health Care System; VA Tampa Health Care System; VA St. Louis Health Care System; VA Bedford Healthcare System
Objectives	Conduct a multi-site, pragmatic randomized clinical trial to compare two health care delivery approaches for chronic pain management (Whole Health Team (WHT) and Primary Care Group Education (PC-GE)) and both compared to Usual Primary Care (UPC) in veterans with moderate to severe chronic pain.
Number of Participants	Planned 765; actual 769
Main Inclusion Criteria	Assigned to VA primary care physician; report pain every day or nearly every day for ≥ 6 months using a phone eligibility screener; PEG score of ≥ 4.
Intervention(s)	The *more* intensive WHT model is delivered by an interdisciplinary pain team consisting of a primary care provider (PCP), integrative health clinician (e.g. mindfulness instructor). The WHT will use the VA Whole Health Personalized Health Planning and Coaching model to promote wellness in veterans with chronic pain. The *less resource-intensive*, PC-GE, uses an abbreviated Cognitive Behavioral Therapy for Chronic Pain (CBT-CP) manualized therapy adapted for a group format based in primary care. This model consists of 7 psychologist-led weekly group sessions emphasizing pain self-management skills and various multi-modal non-pharmacologic approaches for chronic pain management. Both interventions will be compared to Usual Primary Care in which enrolled patients with chronic pain will be randomized to receive usual primary care. Usual primary care incorporates the VA Stepped Care Model for Pain Care which encourages referrals from primary care for secondary and tertiary pain care services, as well as use of Whole Health and complementary and integrative health (CIH) services for pain.
Duration of Intervention and Follow-up	12-months
Blinding	Study statistician; PIs; assessors
Primary Outcome	Pain interference score measured using the Brief Pain Inventory (BPI) interference subscale at 12 months.
Primary Analysis	Hierarchical multi-level model, adjusting for baseline value of the outcome and stratification variables, i.e., site, gender, and current opioid therapy, accounting for repeated measures of individuals and individuals nested within treatment. Pairwise comparisons will be made.

Key Secondary Outcomes	Pain intensity; functioning; quality of life; use of higher-risk pain medications; engagement in non-pharmacological pain management activities.
Interim Analysis	No formal interim analyses for effectiveness or futility. Interim monitoring of the trial for safety and nuisance parameters (e.g., intracluster correlation coefficient) to determine changes to design (e.g., increases in sample size).

Background, Study Objectives and Specific Aims

2.1 Background

In Veterans, chronic pain and prescription opioid use are strongly associated with mental health problems. Veterans with chronic pain and comorbid mental health conditions experience greater pain severity and more pain catastrophizing that those with chronic pain alone. The higher prevalence of opioid use among patients with chronic pain and mental health problems may be driven, in part, by a desire to alleviate mental health symptoms. Prescribing opioids in the context of chronic pain and mental health conditions can exacerbate both mental health and chronic pain conditions, worsen psychosocial functioning, increase risk for serious adverse medical events, and culminate in break-downs in patient and provider communication and trust.[1–3]

Currently, there is little evidence that opioids are effective in the long-term treatment of chronic pain; furthermore, they can result in poorer quality of life and serious adverse outcomes. Evidence suggests Veterans may be disproportionately affected by opioid-related harms; VA patients have nearly twice the rate of accidental poisoning deaths as US adults overall and opioid analgesics are the drug class most commonly involved in these deaths.[4] In short, the expansion of opioid prescribing has triggered an epidemic of opioid-related deaths and addictive disorders without reducing the burden of pain or improving patients' functioning and quality of life.

Concurrently, the opioid crisis has led to a breakdown in primary care management of chronic pain that has negatively impacted both patients and providers. New national guidelines and polices intended to reduce harms from opioids have put pressure on health care systems to change their approach to pain management. The medical community has responded with attempts to rein in the use of opioids,[5] leaving many patients who have relied on opioids in need of effective alternatives. Unfortunately, numerous barriers limit effective pain management in primary care, including ineffective communication, patient and provider knowledge deficits, limited time, and insufficient access to information, assessment tools, and treatment pathways. Healthcare systems lack guidance and resources for implementing effective systems of chronic pain care with primary care. This study specifically addresses these health services gaps.

Newer research suggests that multi-modal and integrated pain care that incorporates non-pharmacological modalities such as exercise, physical therapy, behavioral health care, social supports, and self-care strategies can be more effective than medication-based therapies alone. Despite high-level recommendations for biopsychological pain management, little to no evidence is available for specific strategies to taper or replace opioid therapy and to promote a primary care-based, multi-modal, non-pharmacological approach to chronic pain care.[6] We posit that the VA's patient-centered Whole Health Model can help guide the implementation of multi-modal pain care for Veterans in primary care. The comparator intervention that will be tested in this trial – Primary Care-Group Education – is an adapted version of Cognitive Behavioral Therapy for Chronic Pain (CBT-CP) delivered in a group format in primary care. Both interventions will be compared to usual primary care.

2.2 Study Objectives

1. Conduct a multi-site, pragmatic randomized clinical trial to compare two health care delivery approaches for chronic pain management-Whole Health Team (WHT) and Primary Care Group Education

(PC-GE), and both will be compared to Usual Primary Care (UPC) in veterans with moderate to severe chronic pain.

2.3 Specific Aims

Aim 1: Conduct a multi-site pragmatic trial of Whole Health Team (WHT) vs. Primary Care Group Education (PC-GE) and both compared to Usual Primary Care in veterans with moderate to severe chronic pain. We hypothesize that WHT will be superior to PC-GE, and both will be superior to Usual Primary Care (UPC) with respect to the following outcomes:

> *H1.*a. Improved pain interference at 12 months (primary outcome) and pain intensity, functioning, and quality of life (secondary outcomes).
>
> *H1.b.* Decreased use of higher-risk pain medications, including opioids or high-risk combinations of pain medications (i.e., co-prescription of opioids and benzodiazepines) (secondary).
>
> *H1.c.* Engagement in a greater number of non-pharmacological pain management activities (secondary).
>
> *H1.d.* Improvement in mental health-related symptoms, including sleep problems and suicidality (secondary).

1. Randomization

The unit of randomization was the individual. Stratified block randomization with varying block sizes (24 or 48) was used. Participants were stratified by site, sex, and current opioid therapy and were randomized in an 11:11:2 (WHT:PC-GE:UPC) ratio. Allocation concealment was maintained through a module in Stata called "RALLOC".

2. Blinding

It is not possible to blind participants or those delivering the intervention to treatment assignment. However, we have a blinded the study statistician, the PIs of the study, and those assessing patient reported outcomes will be blinded.

3. Outcomes

5.1 Primary Outcomes

The primary outcome is pain interference at 12-months measured using the BPI interference sub-scale. The BPI was be measured at baseline, 3-,6-, 9-, and 12-months. The BPI interference subscale includes seven numeric ratings of pain interference with general activity, mood, walking, work, relations with other people, sleep and enjoyment of life and is scored as the average of the individual item scores.[7,8]

5.2 Key Secondary Outcomes

Secondary outcomes include the BPI pain intensity, functioning/quality of life (Veterans Rand 12-Item Health Survey), use of pain medications (including opioids if applicable) from VA administrative data (corroborated by self-report), and number of non-pharmacological pain self-management activities. Past-year of use of complementary integrative health (CIH) therapies and self-management practices were measured using the Nonpharmacological and Self-Care Approaches developed by the PMC (NSCAP).[9]

5.3 Additional Secondary Outcomes

Additional secondary outcomes include total BPI and co-morbid mental health-related symptoms. The mental health symptoms were measured using validated brief measures: the Patient Health Questionnaire (PHQ-9) and the General Anxiety Disorders questionnaire (GAD-7) measured depression and anxiety, respectively;[10] the PHQ-9 was used to screen for suicidality; the primary care PTSD screen assessed PTSD-related symptoms;[11] Tobacco, Alcohol, Prescription medication and other Substance (TAPS) Tool;[12] and AUDIT-C[13] measured alcohol and substance abuse; and the PROMIS – Sleep Disturbance[14] measure assessed sleep and fatigue. The Patient Global Impression of Change Scale (PGIC) was used to assess change in overall status during the study.

5.4 Safety Outcomes

An **adverse event (AE)** is generally defined as any unfavorable and unintended diagnosis, symptom, sign (including an abnormal laboratory finding or passive suicidal ideation), syndrome, or disease that either occurs during the study, having been absent at baseline, or if present at baseline, appears to worsen. Adverse events are recorded regardless of their relationship to the study intervention.

When an adverse event occurs, study coordinators will document information regarding the event in the study application event log. Coordinators will also complete an AE form in RedCap and save the downloaded PDF to their site's regulatory binder. Study coordinators will alert the overall trial project manager to all AEs that occur during the trial. AEs that are unanticipated problems involving risks to subjects or others will be reported promptly to the VA Central IRB (CIRB) within five business days. AEs that do not meet criteria for reporting within five days in accordance with VHA Handbook 1058.01 will be reported to the CIRB annually at Continuing Review.

A **serious adverse event (SAE)** is generally defined as any untoward medical occurrence that results in death, is life-threatening, requires inpatient hospitalization or prolongation of existing hospitalization, or results in persistent or significant disability/incapacity. In the case of the wHOPE study, this could include an opioid overdose, for example.

When an SAE occurs, study coordinators will document information regarding the event in the study application event log. Coordinators will also complete an SAE form in RedCap and alert the overall trial project manager, Site PI, and a lead PI (Drs. Seal or Dr. Becker) of the event. After reviewing the SAE form, the site PI and lead PI will sign the form and the coordinator will save the signed PDF to their site's regulatory binder. All SAEs are reported biannually to the DSMB. Serious unanticipated problems involving risks to subjects or others, and unanticipated SAEs will be reported promptly to the CIRB within five business days. SAEs that do not meet criteria for reporting within five days in accordance with VHA Handbook 1058.01 will be reported to the CIRB annually at Continuing Review.

4. Sample Size

The primary outcome is change in BPI interference from baseline to 12-months. Sample size estimates were based on repeated-measures ANOVA assuming two-repeated measures, using the F-test with a focus on the pairwise comparisons. The overall family-wise type I error rate was set at 0.05 with a split of 0.03 for the WHT vs. PC-GE comparison, 0.01 for the WHT vs. UPC comparison and 0.01 for the PC-GE vs. UPC comparison. We assumed a power of 90% for all comparisons. For the comparison between WHT and PC-GE we assumed 1:1 randomization and an effect size of 0.15, which was informed by our prior OPTI study where the correlation between repeated measures was 0.66 and the variance (assuming constant over time) was 5.9. With these assumptions, we determine we would need a sample

size of 273 per arm in the WHT and PC-GE arms. Given the sample size of 273, and assuming an effect size of 0.3, we determine the sample size (using an unequal allocation) needed for the UPC arm to be 50. Assuming a 20% attrition rate, the final sample size was 765 (341 in each of the WHT and PC-GE arms and 83 in the UPC arm).

5. Interim Monitoring

Interim monitoring will focus on patient accrual, baseline comparability of treatment groups, protocol adherence, loss to follow-up, data completeness, quality, and safety. No formal interim look for effectiveness or futility are planned.

6. Data Collection and Data Freeze

Data will be collected using REDCap through primary data collection of patient-reported outcomes. We will use the VHA electronic health record (EHR) to capture selected outcome measures from VA clinical visits (e.g. medications).

Data will be locked one month prior to meeting with the DSMB so that reports can be prepared and distributed to members. The final data set will be locked before conducting any of the proposed analyses.

7. Analytic Plan

7.1 General Considerations

The analysis of the primary and secondary outcomes will be according to the principle of intent-to-treat, i.e., participants will be analyzed according to their original treatment assignment regardless of adherence to the protocol. All analyses will account for the IRGT design and stratification by site, sex, and opioid use. The unit of analysis will be the participant. Three pairwise comparisons will be conducted: WHT vs. PC-GE; WHT vs UPC; PC-GE vs. UPC. SAS version 9.4 and the latest version of R will be used for analyses. Overall, family-wise type I error rates (two-sided) of 0.05 will be considered for both the primary and secondary outcomes split such that 0.03 will be given to the comparison between WHT and PC-GE and 0.01 will be given to each of the comparisons of PC-GE and WHT to UPC.

Comparability of treatment groups will be assessed by comparing the distribution of baseline characteristics in the three groups using appropriate graphical procedures and summary statistics (e.g., means and standard deviations for continuous data; frequencies and proportions for categorical data). The randomization is designed to produce balance on important covariates. Any baseline covariates that appear to be imbalanced between the treatment groups will only be adjusted for in sensitivity analyses.

7.2 Primary Outcome

The primary outcome in this study is pain interference score of the BPI at 12 months. We will use all available data for a participant measured across baseline, 3-, 6-, 9- and 12-months. Hierarchical multi-level models (i.e. linear mixed-effects model), including treatment (WHT, PC-GE, UPC), time (categorical: 3-, 6-, 9- and 12-months), and the treatment by time interaction, and adjusting for the baseline value of the outcome and the stratification variables-site, gender, and current opioid therapy (yes/no), will be used. We will include a random intercept for participant and variance-covariance matrices to account for correlations within a participant and will use random effects to account for individuals nested within treatment (the IRGT nature of the trial). Pairwise comparisons will be made (the comparison WHT vs PC-GE at alpha=0.03, WHT vs UPC at alpha=0.01 and PC-GE vs UPC at alpha=0.01), and we will estimate the

difference between arms at 12 months using a contrast. Least squares mean differences along with confidence intervals (97% for WHT vs. PC-GE; 99% for WHT vs. UPC and PC-GE vs. UPC) will be reported.

7.3 Secondary Outcomes

To control for false discovery for the secondary outcomes, we will use the Benjamini and Hochberg method[15] for each of the pairwise comparisons with family-wise alpha=0.03 for the comparison WHT vs PC-GE and family-wise alpha=0.01 for the comparisons vs UPC . Secondary outcomes are pain intensity, total BPI, functioning, quality of life, use of high-risk pain medications (including opioids if applicable), number of nonpharmacological pain self-management activities, and co-morbid mental health-related symptoms (sleep, suicidality, depression, PTSD, anxiety, alcohol and substance use), and global impression of change. Like the primary outcome, we will use hierarchical multi-level models that include time, treatment, and the treatment by time interaction, and adjust for baseline value of the outcome, site, gender, and current opioid therapy. Random effects and variance-covariance structures will be used to account for repeated measurements within participants and the nesting of individuals within treatment group. Pairwise comparisons will be made using a contrast statement at the 12-month time point. For continuous outcomes, least squares mean differences and the appropriate confidence interval will be reported. For binary or count outcomes, generalized multi-level models will be used, and odds ratios and incident risk ratios will be reported with the appropriate confidence interval.

7.4 Exploratory Analyses/ Per Protocol Analyses

Responder analysis (exploratory analysis). Response will be defined as at least a 30% improvement in BPI interference score from baseline. Reduction in pain/interference of this size has been found to be clinically meaningful and IMMPACT recommended that the percentage of participants with this reduction be reported in chronic pain clinical trials.[16] In addition, they recommended the percentages of patients with at least 50% improvement (substantial improvement) also be reported. We will also report the number who achieved a reduction in pain interference of at least 1 point (minimal clinically important difference for BPI interference).

With the COVID-19 pandemic, treatment implementation may have varied during the study period, and there may have been incomplete adherence to the interventions and outcome assessments. We collected additional time-dependent data to attempt to capture these fluctuations (variables such as mode of treatment delivery, availability of pain management services, and COVID infection status). Our strategies are to: (1) adjust for time-dependent variables that reflect changes in the implementation of treatments, and (2) perform analyses using interactions to see whether the treatment effect differs by time periods, sites, and site characteristics such as degree of availability of pain management services. Inherently these will have limited power, and no further adjustments will be made for multiple testing given the exploratory nature of these analyses.

Additional analyses will consider a per protocol population and repeat the primary analysis in this population. In the per protocol population, we will include any subject who minimally or completely adhered to treatment assignment indicated by the protocol, but will exclude the protocol violators who did not meet the following minimal threshold criteria: if randomized to the WHT arm, participants attended at least 6/8 coaching sessions, the WHT initial visit, and at least 2 of 3 follow-up WHT visits; if randomized to the PC-GE arm, participants attended the orientation and at least 4 of 5 core modules. In addition, participants will be excluded from the per-protocol analysis if they received prohibited concomitant interventions or if they missed all outcome measurements.

We will follow recommendations by Hernan and Robins[17] and also adjust analyses for post-randomization prognostic factors of adherence. We will examine if the post-randomization prognostic factors are affected by prior treatment. If so, we will use the inverse-probability weighting method. This entails estimating a sequential propensity score based on post-randomization time-dependent factors and reweighting subjects using the propensity score estimated earlier. This method helps us minimize post-randomization confounding and selection bias.

9.5 Safety Analysis

Analysis of the safety data will involve tabulating the occurrence of serious adverse events, adverse events, and unanticipated problems between the two groups (see 5.4 Safety Outcomes above).

9.6 Missing data

Several strategies will be imposed to accommodate the likelihood that missing data will occur during the study. Prevention is the most obvious and effective manner to control bias and loss of power from missing data. We will follow the intent-to-treat principle, requiring follow-up of all participants enrolled regardless of the treatment received. Timely data entry combined with quarterly missing data reports will trigger protocols for tracking and obtaining missing data. Despite these efforts, it is reasonable to assume missing data will occur. Our primary analysis is valid under the assumption that missing data are missing at random (MAR). We will evaluate the plausibility of this assumption by determining the extent of missing data and using logistic regression to identify factors associated with missing data. If we find baseline factors associated with missingness, then we will conduct sensitivity analyses to investigate the effect of including those factors as covariates in the models evaluating the difference between groups. As a second sensitivity analysis under the MAR assumption, we will use multiple imputation based on chained equations (which can accommodate both continuous and categorical variables) to impute missing variables. We will conduct sensitivity analysis using pattern-mixture models and multiple imputation under missing not at random (MNAR) assumptions to examine the robustness of conclusions of the primary analysis to missing data. The multiple imputation under MNAR will assume that those who drop out have worse outcomes than those who stay in the study.[18] We will also consider inverse probability weighting to account for missing data, if necessary.

References

1. Sullivan MD, Edlund MJ, Zhang L, Unützer J, Wells KB. Association between mental health disorders, problem drug use, and regular prescription opioid use. *Arch Intern Med*. 2006;166(19):2087-2093. doi:10.1001/archinte.166.19.2087

2. Sullivan MD, Leigh J, Gaster B. Brief report: Training internists in shared decision making about chronic opioid treatment for noncancer pain. *J Gen Intern Med*. 2006;21(4):360-362. doi:10.1111/j.1525-1497.2006.00352.x

3. Harris KM, Edlund MJ. Self-medication of mental health problems: new evidence from a national survey. *Health Serv Res*. 2005;40(1):117-134. doi:10.1111/j.1475-6773.2005.00345.x

4. Bohnert ASB, Ilgen MA, Galea S, McCarthy JF, Blow FC. Accidental poisoning mortality among patients in the Department of Veterans Affairs Health System. *Med Care*. 2011;49(4):393-396. doi:10.1097/MLR.0b013e318202aa27

5. Dowell D, Haegerich TM, Chou R. CDC Guideline for Prescribing Opioids for Chronic Pain--United States, 2016. *JAMA*. 2016;315(15):1624-1645. doi:10.1001/jama.2016.1464

6. Chou R, Turner JA, Devine EB, et al. The effectiveness and risks of long-term opioid therapy for chronic pain: a systematic review for a National Institutes of Health Pathways to Prevention Workshop. *Ann Intern Med*. 2015;162(4):276-286. doi:10.7326/M14-2559

7. Keller S, Bann CM, Dodd SL, Schein J, Mendoza TR, Cleeland CS. Validity of the brief pain inventory for use in documenting the outcomes of patients with noncancer pain. *Clin J Pain*. 2004;20(5):309-318. doi:10.1097/00002508-200409000-00005

8. Tan G, Jensen MP, Thornby JI, Shanti BF. Validation of the Brief Pain Inventory for chronic nonmalignant pain. *J Pain*. 2004;5(2):133-137. doi:10.1016/j.jpain.2003.12.005

9. Non-pharmacological and self-care approaches from PMC Survey. Published online August 2019.

10. Seal KH, Borsari B, Tighe J, et al. Optimizing Pain Treatment Interventions (OPTI): A Pilot Randomized Controlled Trial of Collaborative Care to Improve Chronic Pain Management and Opioid Safety--Rationale, Methods, and Lessons Learned. *Contemp Clin Trials*. 2018;77:76. doi:10.1016/j.cct.2018.12.006

11. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ*. 2015;350:h2147. doi:10.1136/bmj.h2147

12. McNeely J, Wu LT, Subramaniam G, et al. Performance of the Tobacco, Alcohol, Prescription Medication, and Other Substance Use (TAPS) Tool for Substance Use Screening in Primary Care Patients. *Ann Intern Med*. 2016;165(10):690-699. doi:10.7326/M16-0317

13. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality

Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. *Arch Intern Med*. 1998;158(16):1789-1795. doi:10.1001/archinte.158.16.1789

14. Yu L, Buysse DJ, Germain A, et al. Development of short forms from the PROMIS™ sleep disturbance and Sleep-Related Impairment item banks. *Behav Sleep Med*. 2011;10(1):6-24. doi:10.1080/15402002.2012.636266

15. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B Methodol*. 1995;57(1):289-300.

16. Dworkin RH, Turk DC, Wyrwich KW, et al. Interpreting the Clinical Importance of Treatment Outcomes in Chronic Pain Clinical Trials: IMMPACT Recommendations. *J Pain*. 2008;9(2):105-121. doi:10.1016/j.jpain.2007.09.005

17. Hernán MA, Robins JM. Per-Protocol Analyses of Pragmatic Trials. *N Engl J Med*. 2017;377(14):1391-1398. doi:10.1056/NEJMsm1605385

18. Yuan Y. Sensitivity Analysis in Multiple Imputation for Missing Data. In Proceedings of the SAS Global Forum 2014 Conference:[http://support. sas. com/resources/papers/proceedings14/SAS270-2014. pdf].