VA Maryland Healthcare System

Cooperative Study

Can Service Dogs Improve Activity and Quality of Life in Veterans with PTSD?

Statistical Analysis Plan

Version Number 2
July 15, 2019

Privileged & Confidential

This document is a confidential communication of the Veteran's Administration. Acceptance of this document constitutes agreement by the recipient that no unpublished information contained herein will be published or disclosed without the Veteran's Administration prior written approval, except that this document may be disclosed to appropriate Institutional Review Boards under the condition that they are requested to keep it confidential.

STATISTICAL ANALYSIS PLAN APPROVAL

Version 2.0

Responsibility	Name	Signature	Date
PPCSPCC Study Biostatistician	Eileen M. Stock, PhD		
Study Chairman	Joan T. Richerson, DVM		
PPCSPCC Director	Kousick Biswas, PhD		

Table of Contents

1 INTRODUCTION

1.1 Preface

The number of Veterans with post-traumatic stress disorder (PTSD) within the VA population has increased dramatically in the past years, largely due to Operation Enduring Freedom and Operation Iraqi Freedom (OEF/OIF), but also due to increasing numbers of Veterans from all eras seeking treatment and disability claims. PTSD is associated with increased disability and decreased functioning (Gellis et al, 2010). Most treatments are designed to reduce symptoms, with the expectation that improvements in functioning and decrease in disabling health condition will naturally result; however, for many Veterans, PTSD is chronic, and symptom control is the best hope. As with many illnesses, which cannot be cured, strategies to decrease limitations on activity and improve quality of life are important.

Case studies demonstrate that Service Dogs trained for mental health disabilities may fit that need for people with PTSD. As part of the Department of Defense Bill in 2010, Congress enacted a law that stated 'a pilot study would be completed' to determine the benefits of Service Dogs in helping individuals with mental and physical disabilities. Results of the proposed study will be used to shape future policy for use of Service Dogs for Veterans with mental health diagnoses, specifically PTSD.

1.2 Goal of the analyses

The goal is to compare how the provision of dogs (Service Dogs and Emotional Support Dogs) impact function, quality of life, mental health, healthcare utilization and healthcare costs for Veterans diagnosed with PTSD.

2 STUDY OBJECTIVES AND ENDPOINTS

2.1 Study Objectives and Hypotheses

Primary Objectives:

Objective 1: To examine how limitations on activity and quality of life in Veterans with PTSD are impacted by the provision of a Service Dog versus an Emotional Support Dog.

Limitation on activity is defined as the inability to fully engage in important life domains, such as cognition, mobility, self-care, and participation in society. For all the hypotheses listed below we will examine change over the 18-month intervention period between the two groups of participants with PTSD: those who are paired with a Service Dog and those who are paired with an Emotional Support Dog.

Hypothesis 1a: Compared to Veterans who receive an Emotional Support Dog, Veterans who receive Service Dogs will have improved ability to fully engage in important life domains over time as measured by the WHO-DAS 2.0 domain scores and the WHO-DAS 2.0 total score.

Hypothesis 1b: Compared to Veterans who receive an Emotional Support Dog, Veterans who receive Service Dogs will have improved quality of life, as measured by the global mental and physical health component scores of the VR-12.

Secondary Objectives:

Objective 2: To examine how mental health is impacted by the provision of a Service Dog versus an Emotional Support Dog.

Hypothesis 2a: Compared to Veterans who receive an Emotional Support Dog, Veterans who receive a Service Dog will have reduced PTSD symptom severity, as assessed by the PCL-5
NOTE: Although we are collecting the CAPS periodically, its use will be to confirm diagnosis of PTSD.

Hypothesis 2b: As compared to Veterans who receive an Emotional Support Dog, Veterans who receive a Service Dog will have decreased thoughts of suicide, as assessed by the Columbia-Suicide Severity Rating Scale (C-SSRS).

Hypothesis 2c: Compared to Veterans who receive an Emotional Support Dog, Veterans who receive a Service Dog will have a decrease in depression as assessed by the PHQ-9.

Hypothesis 2d: Compared to Veterans who receive an Emotional Support Dog, Veterans who receive a Service Dog will have improved sleep outcomes as measured by the Pittsburgh Sleep Quality Index (PSQI)

Objective 3: To characterize and compare how health care utilization and costs are affected by the provision of a Service Dog or Emotional Support Dog.

Hypothesis 3a: Compared to Veterans who receive an Emotional Support Dog, Veterans who receive a Service Dog will have lower health care utilization as quantified by inpatient and outpatient visits to healthcare providers and to mental health providers.

Hypothesis 3b: Compared to Veterans who receive an Emotional Support Dog, Veterans who receive a Service Dog will have decreased medication usage as assessed by the medical record.

Hypothesis 3c: Compared to Veterans who receive an Emotional Support Dog, Veterans who receive a Service Dog will have decreased use of sleep medications as assessed by the medical record.

Objective 4: To characterize and compare how employment and productivity are affected by the provision of a Service Dog or Emotional Support Dog.

Hypothesis 4a: Compared to Veterans who receive an Emotional Support Dog, Veterans who receive a Service Dog will be more likely to be employed at follow-up.

Hypothesis 4b: Compared to Veterans who receive an Emotional Support Dog, Veterans who receive a Service Dog will have greater work productivity as quantified by the Work Productivity and Activity Impairment Questionnaire: General Health Problem V2.0.

2.2 Endpoints

2.2.1 Primary Outcome Measures

The primary outcome measures include limitations of activity and quality of life.

For hypothesis 1a: The outcome measure will define activity limitations as measured by the World Health Organization Disability Assessment Scale II (WHODAS 2.0). The WHODAS 2.0 is a structured 36 item instrument which assesses difficulties in six domains of life during the last 30 days. The six domain scores and the total WHO-DAS 2.0 score will be utilized for this hypothesis.

For hypothesis 1b: The outcome measure will be the summary measures from the VR-12 instrument of health-related quality of life as measured by the Physical Component Summary (PCS) and Mental Component Summary (MCS).

2.2.2 Secondary and Tertiary Outcome Measures

Secondary outcomes include PTSD severity and symptoms, depression, sleep, suicide intent and healthcare utilization and cost, and employment.

Hypothesis 2a**:** The outcome measure for hypothesis 2a will be the Posttraumatic Stress Disorder Checklist (PCL-5). The PCL-5 is a 20-item self-report measure of PTSD symptoms (in the past month) that assesses the 20 DSM-5 symptoms of PTSD. The 20 questions are responded to with a 5-point Likert scale response format.

Hypothesis 2b will examine suicide ideation, which will be assessed by the Columbia-Suicide Severity Rating Scale (C-SSRS).

Hypothesis 2c examines depression, which will be assessed by the Patient Health Questionnaire (PHQ-9).

Hypothesis 2d measures sleep as assessed by the Pittsburgh Sleep Quality Index (PSQI).

Hypotheses 3a, 3b and 3c examine healthcare utilization and costs. Information for these hypotheses will be collected from VA administrative data sets and using the Health Economics Resource Center (HERC)-developed standard questions regarding outpatient and inpatient utilization.

Hypotheses 4a and 4b examine employment outcomes. The Work Productivity and Activity Impairment Questionnaire: General Health Problem V2.0 will be used for this outcome.

2.3 Derived variables

2.3.1 WHODAS 2.0 Summary and Domain Scores

The World Health Organization Disability Assessment Scale II (WHO-DAS 2.0) is a structured 36-item instrument, which assesses difficulties in six domains of life during the last 30 days. A total disability score is produced as well as domain scores. The domains include:

1. Cognition- understanding and communicating with the world
2. Mobility- moving and getting around
3. Self-care- attending to one's hygiene, dressing, eating and staying alone
4. Interpersonal interactions- getting along with people
5. Life activities- domestic responsibilities, leisure, and work
6. Participation in society-joining in community activities

The summary score is calculated using the 36 questions on the instrument. Summary scores can range from 0 to 100. Initially if any of the domain scores are missing, then the WHODAS summary score will be considered missing. However, if significant missing data is found, a sensitivity analyses will be computed, and the executive committee will determine if data should be imputed using standard multiple imputation methods.

2.3.2 VR-12

The scoring of the PCS and MCS for the VR-12 is based on weights derived from the VR-36 instrument administered to 1.4 million Veteran enrollees with 877,775 respondents in the 1999 Large Health Survey of Veteran Enrollees (Veterans Health Study) (Iqbal, 2009). Higher PCS and MCS scores reflect greater quality of life. Imputation methods may be used to calculate scores for Veterans who do not complete all 12 questions included in this measure.

2.3.3 PTSD Checklist – PCL-5

The PCL-5 is a 20-item self-report measure that assesses the 20 DSM-5 symptoms of PTSD. The PCL-5 is similar to the PCL-S (specific) version. The wording of PCL-5 items reflects both changes to existing symptoms and the addition of new symptoms in DSM-5. The PCL-5 is scored by giving a **score of 0 through 4 for each symptom**, reflecting a change from 1-5 in the DSM-IV version. Rating scale descriptors are the same: "Not at all," "A little bit," Moderately," "Quite a bit," and "Extremely." There are several different ways to score the PCL-5: 1) Calculate a total symptom severity score (ranging between 0 and 80) by summing the scores for each of the 20 items; 2) Calculate DSM-5 symptom cluster severity scores by summing the scores for the items within a given cluster [i.e., cluster B (items 1 – 5), cluster C (items 6-7), cluster D (items 8 – 14) and cluster E (items 15-20)]; or 3) calculate a provisional PTSD diagnosis by treating each item rated as a 2 ("Moderately") or higher as a symptom endorsed, then follow the DSM-5 diagnostic rule which requires at least: 1 B item (question 1 – 5), 1 C item (question 6-7), 2 D items (question 8-14), 2 E items (questions 15-20).

Overall, optimal PCL-5 cut-points appear to be 11 – 14 points lower than PCL for DSM-IV cup-points. A PCL-5 cut-point of 38 appears to be a reasonable value to propose until further psychometric work is available. The recommended minimum threshold for determining whether an individual has responded to treatment is 5 points (Weathers et al., 2013).

2.3.4 Columbia Suicide Severity Rating Scale (C-SSRS)

The C-SSRS is a 4-page form asking questions about suicidal ideation, intensity of ideation, and suicidal behavior. Developed by Posner and collaborators at the New York State Psychiatric Institute (Oquendo et al 2003), the scale is intended for use by trained administrators. The questions contained in the C-SSRS are suggested probes. Ultimately, the determination of the presence of suicidality depends on clinical judgment.

2.3.5 Patient Health Questionnaire (PHQ-9)

The PHQ-9 consists of 9 questions to measure depression. The questions are answered on a scale from 0 (not at all) to 3 (nearly every day). The total score is calculated by adding together the selected rating for each of the 9 questions. Summary PHQ-9 scores can range from 0 to 27. Initially, if any of the nine questions on the PHQ-9 are missing, then the total PHQ-9 score will be missing. However, if significant missing data is found, a sensitivity analyses will be computed, and the executive committee will determine if data should be imputed using standard multiple imputation methods.

2.3.6 Pittsburgh Sleep Quality Index (PSQI)

The PSQI is a 24-item self-administered survey used to assess sleep-related problems during the past month. The first 19 items are completed by the subject and there are five items completed by a bed partner or roommate. The five items answered by a bed partner or roommate are used as clinical information and are not included in scoring. The first 19 items are grouped into seven components:

1. sleep quality
2. sleep latency

3. sleep duration
4. habitual sleep efficiency
5. sleep disturbances
6. use of sleep medication
7. daytime dysfunction

Each of the seven components are weighted equally on a 0-3 scale with 0 (better) to 3 (worse). The seven component scores are then summed to yield a global score, which has a range of 0-21; higher scores indicate worse sleep quality.

2.3.7 Dimensions of Anger Reactions (DAR):

This seven item scale measures anger disposition that is directed to other individuals (Forbes et al, 2004). It has been shown to be a reliable and sensitive measure in populations that have PTSD.

2.3.8 Non-VA Healthcare Utilization

Non-VA Healthcare Utilization will be assessed using the "HERC non-VA utilization survey." This survey was created in 2011 by VA Health Economics Resource Center (HERC) investigators. The HERC non-VA utilization survey is self-administered and asks about outpatient and inpatient (including Emergency Department) visits to non-VA providers. Inpatient visits are characterized by hospital name, location, length of stay, and type of hospital (e.g., general medical, nursing home, psychiatric or substance abuse facility, etc).

2.3.9 Employment/Productivity

Veteran's employment status and work productivity will be assessed through the Work Productivity and Activity Impairment Questionnaire: General Health Problem V2.0 (WPAI:SHP). The WPAI is the most frequently-used instrument to assess work productivity (Prasad et al., 2004). The WPAI: GHP is a 6-item questionnaire that can be interviewer-administered or completed by a respondent (Reilly et al., 1993). The questionnaire asks about the number of hours of work missed due to health problems as well as the effect of the health problems on productivity while at work.

2.3.10 Treatment Groups

Veterans will be randomized and paired to:

- A Service Dog which has been trained for specific tasks to assist with the Veteran's PTSD disability.

- An Emotional Support Dog (dogs with suitable behavior characteristics, Canine Good Citizen trained dogs to provide emotional comfort).

3 STUDY METHODS

3.1 General Study Design and Plan

This is a longitudinal randomized trial, with two randomized arms which will follow Veterans with PTSD for 18 months post-pairing to examine how the provision of a Service Dog or Emotional Support Dog impacts their function by assessing longitudinal change in functioning over time. The treatment arms consist of Veterans with PTSD partnered with Service Dogs (SERV) and Veterans with PTSD partnered with Emotional Support Dogs (EMOT). This study aims to enter approximately 220 participants in three participating VA medical centers over approximately a 36-month (3 years) recruitment period. For challenges in the recruitment of participants or in the pairing (or availability) of dogs, sample size will be revisited. All participants will be followed for a total of 18 months post-pairing with a Service Dog or Emotional Support Dog for the determination of primary and secondary outcome measures.

3.2 Screening and Home/Baseline Visits

Following initial screening for inclusion criteria, participants will sign the consent form and will begin the screening visits of the study. During this visit, inclusion criteria will be verified by performing the Clinician Administered PTSD Scale for DSM 5 (CAPS-5) and the M.I.N.I. International Neuropsychiatric Interview. Within approximately one week of the screening visit, a home visit will occur to assess the home environment and to verify it is safe for a dog. During the home visit, or within approximately one week, the baseline assessment on functioning, mental health, psychosocial well-being and socioeconomic and health care utilization will be given.

3.3 Inclusion-Exclusion Criteria

3.3.1 Inclusion Criteria

1) Males and Females greater than 18 years of age
2) Referral from Mental Health provider that documents PTSD.
3) PTSD as a result of any trauma as determined by meeting DSM 5 diagnostic criteria.
4) Enrolled in mental health services at VA and has attended at least one visit in the 90 days prior to consent.
   - If individual not currently enrolled in mental health treatment decides to enroll in such then he/she may become eligible to participate in the study.
   - If individual enrolled in mental health treatment schedules and attends a mental health visit then he/she may become eligible to participate in the study.
5) Agrees to remain in mental health treatment throughout the duration of the study
6) Can adequately care for and handle the dog
   - Adequately caring for a dog requires that participants will be responsible for and able to provide food, water, protection, shelter, exercise, transportation, and treatment related to their assigned dog.
   - Adequately handling the dog means having the ability to give and reinforce obedience commands and control the dog using a leash.

7) Home environment is suitable for a dog.
   - If the home environment can be remedied the potential participant may become eligible to participate in the study.
   - If a participant moves home while enrolled in the study the new home must be suitable for a dog
8) Home environment is structurally and geographically accessible for study staff
   - If the home is geographically inaccessible to study staff and, the individual cannot remedy the situation unless he/she moves home. The study team will not encourage this. If a move takes place, it will be the individual's responsibility to re-contact the study team.
   - If the individual changes home residence while enrolled in the study, the new home must be geographically accessible to study staff. If it is inaccessible, the dog will be removed, and the individual will be withdrawn from the study.
9) Is willing to accept randomization outcome.
10) Has someone to care for the dog during extended absence of the participant.
   - If no one is available to care for the dog but the situation changes then the participant may become eligible to participate.
11) Others in home are agreeable to having dog
   - If others in the home are not agreeable but at a later date the situation changes, then the potential participant may become eligible to participate
12) Is willing and able to travel (by air or car) to the dog vendor training site for pairing if assigned to receive a service dog.
   - If individual's unwillingness to travel to a training site changes, he/she may become eligible to participate. In this instance, it will be the individual's responsibility to re-contact the study team.
13) Individual has no pet in the home to threaten the bonding and obedience training of an assigned study dog.
   - If a household dog lives inside the home and the home is partitioned such that there are two or more separate living spaces served by independent entrance/exits, and the individual does not live in a partition with a dog, then the individual can be eligible. If a household dog lives primarily outside the home in a rural area and the individual is not primarily responsible for feeding the dog on a daily basis, then the individual can be eligible.
   - If an individual has pets other than dogs that could interfere with bonding, the individual will be scheduled for the screening visits and the relationship will be assessed by the dog trainer.
   - If an individual has a household dog or other pet that prevents participation in the study but the situation changes, the individual may become eligible to participate. In this instance, it will be the individual's responsibility to re-contact the study team.
14) Individuals can verbalize understanding of consent form, is willing to provide written informed consent and to follow study procedures.

3.3.2 Exclusion Criteria

1) Hospitalization for mental health reasons in the past 6 months

- Once six months since hospitalization have passed, the individual may become eligible to participate in the study

2) Aggressive behavior that would make it unsafe for dog
3) Diagnosis of psychosis, delusions, dementia, moderate or severe alcohol/substance disorder (SUD), or moderate to severe traumatic brain injury as determined by the presence or absence of a condition following scoring of MINI responses or as documented in chart notes.

   SUD assessment (alcohol/non-alcohol):

   - Ineligibility is based on the presence of a Moderate (4-5 symptoms) to Severe (6+ symptoms) SUD as identified by the MINI within the previous 12-month period starting from date of the study MINI screening.
   - If a Moderate to Severe SUD has been documented or communicated by the referring clinician or potential participant, or is noted in the EMR prior to the initial MINI screening visit, individuals should be scheduled for their initial screening visit on a timeline commensurate with meeting the 12-month SUDs eligibility window.
   - If an individual is identified as ineligible during the initial screening visit (i.e. MINI SUDs score ≥4) he/she may be re-evaluated later at the discretion of the study team. Re-evaluations should be scheduled based on a timeline commensurate with meeting the 12-month SUDs eligibility window (absence of a Moderate to Severe SUD for the previous 12 months). If at re-evaluation the individual has <4 symptoms, he/she may become eligible to participate in the study.

4) Active suicidal intent as determined by a CPRS flag for suicidal intent or an endorsement of yes to question 5 (active suicidal ideation with specific plan and intent) on the C-SSRS completed at the Clinic Qualifying Visit.
   - An endorsement of yes to question 4 (Active Suicidal Ideation with Some Intent to Act, without Specific Plan) without endorsement of question 5 indicates that the individual needs additional assessment to determine eligibility.
5) Homicidal intent or cognitive disabilities that would preclude safety of dog and/or ability to participate in the study.
6) Social, mental or physical condition that prevents the individual from either giving informed consent or participating in the study.
7) Participation in another unapproved research trial.
   - If the individual is in another unrelated study and both the study Chair/PI of this and the other study consider participation in both studies to be acceptable then the individual may become eligible to participate in this study.
   - If the study Chair/PI of this and/or the other study consider participation in both studies to be unacceptable then, once participation in the other study is complete, the participant may become eligible to participate in this study. At that time, it will be the individual's responsibility to re-contact the study team.
8) Has CPRS flag for violent/disruptive behavior.
9) Potential participants who are pregnant/who have a partner who is pregnant, or who currently have one or more children younger than age 5 in the household for more than 8 hours per day, one day a week will be excluded from the study.
   - If a participant or anyone else in the household becomes pregnant during the observation period, the participant will be excluded from the study.

- Participants who have children in their home/become pregnant after being paired with a dog will be evaluated on a case-by-case basis (see Safety Monitoring of Children in the Home below)
- After a total of 10 dogs have been placed with participants who have children between the ages of 5 and 10 years, and after each pairing has successfully reached and passed the 2-month home visit, this exclusion criterion will be revisited for potential inclusion of participants with children younger than 5 years.

3.4 Randomization and Blinding

Once a Veteran has been approved to participate in the study, the participant will be randomized using the Perry Point Interactive Touch Tone Randomization System (ITTRS). During this call the Veteran will be randomized to a dog type (SERV or EMOT). Randomization will utilize a random block scheme to assign the Veteran to the dog type. At this point the vendor, Veteran, and site, will not be notified which group (SERV or EMOT) the Veteran was randomized to. The CSPCC will inform the contracting officer's representatives (COR) of the dog type randomization assignment, who will in turn relay the information on the vendor assigned to the site and the CSPCC.

Once randomized, the site will provide the Veteran contact information for the vendor. The Veteran will be expected to contact the vendor within two weeks of receipt of information. Once the Veteran contacts the vendor, an interview will take place between the Veteran and vendor so that the vendor may gain information regarding the personality of the Veteran in order to find the correct match of dog. This interview should take no more than three hours, and the vendor will have up to two weeks to complete.

After completing the interview there will be no further interaction between the Veteran and the vendor without the specific approval of the study team and COR. During any contact between vendor and Veteran, members of the ORD study team may be present. If the vendor needs additional information from the Veteran, they will contact the CORs for appropriate protocol. After the interview has been completed, the vendor will be unblinded to the dog type the Veteran was randomized to (SERV or EMOT). The site and the Veteran will not be told of the randomized dog type until the clearing visit.

Sealed randomization envelopes containing the dog type (SERV or EMOT) will be provided to the site after each randomization. If the participant remains eligible for the study after the final home clearing visit, then the study team member(s) present will open the envelope with the participant to find out if the participant was randomized to the Service Dog or Emotional Support Dog group. At the point, neither the LSI, study team, Chair's Office nor the Veteran will be blind to the treatment group.

3.5 Observational Phase

Following randomization, participants will be followed for three months or more for the observation phase of the study. After three months and once the assigned dog from the Vendor has become available and been approved by the VA, the participant is eligible for the clearing visits. During the clearing visits, eligibility criteria will be verified, the assessments on functioning, mental health, psychosocial well-being and socioeconomic and health care utilization will be administered and a home visit will verify the home remains suitable for a dog.

3.6 Forms for Screening, Baseline, Observational and Clearing Visits

The following forms will be used during the screening visit, home visit, observational phase and clearing visit of the study.

Table 1. Screening and Clearing Data Forms

Form #	Form Name	Screening	Home/Baseline	Clearing Visit 1	Clearing Visit 2
	PHASE	0	0	1	1
	VISIT #	01	02	01	02
	Location	Clinic	Home	Clinic	Home
86	Informed Consent Confirmation	R			
1	Demographics	R			
2	Veteran Characteristics	R			
3	CAPS	R			
4	MINI	R			
5	Suitability for a Dog		R		R
7	WHO-DAS 2.0		R	R	
8	PCL		R	R	
9	PSQI		R	R	
10	VR-12		R	R	
11	C-SSRS	R		R	
12	PHQ-9		R	R	
13	DAR		R	R	
14	HERC Non-VA Utilization		R	R	
15	WPAI: GHP v2.0		R	R	
16	Inclusion/Exclusion & Randomization				R
17	Medication Log	A	A	A	A
18	Payment Log	A	A	A	A
25	Protocol Deviation	A	A	A	A
26	Adverse Events	A	A	A	A

26a	Adverse Events for Dogs	A	A	A	A
27	Serious Adverse Events	A	A	A	A
27a	Serious Adverse Events for Dogs	A	A	A	A
28	AE/SAE Follow-up	A	A	A	A
28a	SAE Follow-up for Dogs	A	A	A	A

R= REQUIRED

A = AS NEEDED

3.7 Follow-up

The primary objective of this study is to determine how the provision of a Service Dog or an Emotional Support Dog impacts function and quality of life for Veterans with PTSD. Primarily, the objective is to compare functioning and quality of life change, relative to baseline, over the 18-month intervention period between the two groups of Veterans with PTSD, i.e., those who receive Service Dogs and those who receive Emotional Support Dogs. The secondary outcomes in this study are: 1) a reduction in PTSD symptoms as measured by the PTSD Civilian Checklist (PCL); 2) improvement in depression as measured by the PHQ-9; 3) improvement in sleep as measured by the Pittsburg Sleep Quality Index (PSQI); 4) decreased thoughts of suicide as measured by the Columbia-Suicide Severity Rating Scale (C-SSRS); and 5) to characterize and compare how healthcare costs and utilization are impacted. All primary and secondary outcome measures will be collected at baseline, clearing, and at 3, 6, 9, 12, 15 and 18 month post-pairing. The following forms will be used during the follow-up phase of the study.

Table 2. Follow-up Phase Data Forms

Form #	FORM Name	1 week	2 week	1 month	2 month	3 month	6 month	9 month	12 month	15 month	18 month
	PHASE	2	2	3	3	3	3	3	3	3	3
	VISIT #	01	02	01	02	03	06	09	12	15	18
	LOCATION	Home	H/P	Home	H/P	Clinic	Home	Clinic	Home	Clinic	Home
3	CAPS										R
7	WHO-DAS 2.0					R	R	R	R	R	R
8	PCL-5					R	R	R	R	R	R
9	PSQI					R	R	R	R	R	R
10	VR-12					R	R	R	R	R	R
11	C-SSRS					R	R	R	R	R	R
12	PHQ-9					R	R	R	R	R	R
13	DAR					R	R	R	R	R	R
14	HERC Non-VA Utilization					R	R	R	R	R	R
15	WPAI: GHP v2.0					R	R	R	R	R	R
17	Medication Log	A	A	A	A	A	A	A	A	A	R
18	Payment Log	A	A	A	A	A	A	A	A	A	R
19	Post Pairing Evaluation	R	A	A	A	A	A	A	A	A	A
20	Veteran and Service/ES Dog Visit Report	R	A	R	A	A	R	A	R	A	R
21	Dog Related Questions	R	R	R	R	R	R	R	R	R	R
22	Dog Return	A	A	A	A	A	A	A	A	A	A
23	Veterinarian Checklist	R					R		R		R
24	Study Completion/Termination	A	A	A	A	A	A	A	A	A	R
24a /b	Exit Interview – SERV or Exit Interview – EMOT										R
24c	Dog Trainer Evaluation										R
25	Protocol Deviation	A	A	A	A	A	A	A	A	A	A
26	Adverse Events	A	A	A	A	A	A	A	A	A	A
26a	Adverse Events for Dogs	A	A	A	A	A	A	A	A	A	A
27	Serious Adverse Events	A	A	A	A	A	A	A	A	A	A
27a	Serious Adverse Events for Dogs	A	A	A	A	A	A	A	A	A	A
28	AE/SAE Follow-up	A	A	A	A	A	A	A	A	A	A
28a	SAE Follow-up for Dogs	A	A	A	A	A	A	A	A	A	A

H/P = HOME/PHONE
R = REQUIRED
A = AS NEEDED

3.8 Sample Size

The goal for this study is a sample size of 220 including adjustment for subjects who withdraw or are lost-to-follow-up and interim analyses. In order to detect a 15% difference in mean scores for MCS (outcome requiring largest sample) over the 18-months of follow-up, at a statistical significance level of 0.05 (two-tailed test) and a power of 0.85, 82 participants per group will be needed, and 110 participants per treatment (220 total) will be required to account for a maximum of 25% post-pairing participant loss or dropout rate (see protocol). This sample size results in an expectation of 110 subjects in each of the treatment groups. For challenges in the recruitment of participants or in the pairing (or availability) of dogs, sample size will be revisited and adjusted accordingly taking into account the study's post-pairing termination rate. For 80%, a sample size of 72 per group is needed, and 90 participants per treatment (180 total) to account for a 20% post-pairing termination rate or 85 (170) to account for a 15% termination rate. When such challenges occur, a sample size of 180 will be the aim.

4 STUDY AND DATA MANAGEMENT

4.1 Study Management at the CSPCC

A Cooperative Studies Program Coordinating Center (CSPCC) study team has been assigned to, this study for providing data management, statistical, and administrative supports to the study executive committee for a smooth conduct and timely completion of the study. The study team is comprised of:

Biostatistician and Team Lead	Eileen Stock, Ph.D.
Project Manager	Leslie Norman, LGSW
Statistical Programmer	Frances McSherry
Database Programmers	Joseph Tadalan
Computer Assistants	Ellen Sterrett

Other core CSPCC staff, for example, Quality Assurance, Travel Clerk, Printer, Secretary, etc., will provide help based on the need of the study.

The Biostatistician is the study team leader and has the overall responsibility for the conduct of the study at the CSPCC. S/he is the CSPCC's spokesperson to the Study Group; s/he represents the CSPCC on the study's Executive Committee and along with the Study Chairpersons, s/he is responsible for representing the study at the Data Monitoring Committee meetings. The Biostatistician is also responsible for providing the Study Group with statistical and clinical trial advice, for working with other CSPCC team members in the preparation of routine interim reports, and for conducting the final analyses at the end of the study.

The Project Manager is responsible for the administrative coordination of the study by the CSPCC. S/he serves as the Biostatistician's Administrative Assistant and works with the CSPCC study team to ensure that all reports, study materials, and meeting arrangement notices are sent to the proper

individuals in a timely fashion. S/he will work closely with the National Study Coordinator in the Chairman's office to ensure that the study runs smoothly and will be in contact with both the National Study Coordinator and the Local Research Coordinators at the participating centers at least monthly to discuss any problems that they may be having, including those with the CSPCC. S/he will also work with the local VA R&D Offices at the participating centers to obtain R&D and IRB approvals at the beginning of the study and annually as well as the preparation of study budgets yearly during the ongoing phases of the study.

The Statistical Programmer is responsible for the preparation of the tables and analyses for all of the routine study reports. These include Study Group, Executive Committee, Data Monitoring Committee, and the mid-study report to CSSEC. S/he also prepares the tables and reports for the final analyses. S/he works closely with the Biostatistician on these analyses.

The Database Management System (DBMS) Programmer is the lead of the data management support group and works closely with the assigned computer assistant(s) to address the data management need for the assigned study. S/he is responsible for establishing, updating and maintaining the study's database. In addition, s/he will write edit program based on an agreed upon edit plan that will thoroughly check the data for errors and missing information. S/he is also responsible for programming and maintaining the randomization system for the study.

The *Computer Assistant(s)* are responsible for setting up the data definition table for the study, for building the Study Definition Editor and also for laying out the electronic case report forms in the form design software. They are also responsible for training the study staff at each site on how to properly manage the data collection process and how to appropriately respond to data edits. The computer assistant(s) are also responsible for working with the sites to resolve the data queries generated based on the incomplete and/or inaccurate data submitted to the study database.

4.2 Data Management

The data management system for this study will be built using DataFax. Each site will fax completed data collection forms to the CSPCC where they will be reviewed and validated. When a fax arrives at the CSPCC, DataFax breaks it into pages and proceeds to process each page. It identifies which study each page belongs to, reads the data from each page, enters the data into the study database, and stores all pages. CSPCC Computer Assistants use split-screen validation to review all pages received, complete data entry, and flag any problems they identify (e.g., missing data) with Quality Control (QC) notes. Randomization will be completed at the Perry Point CSPCC using ENVOX, an automated interactive telephone system.

Tables 1 and 2, above, list the case report forms and the assessments that will be used in this study and indicates when each will be administered. The Local Site Investigator (LSI), Study Coordinator (SC) or other local study team members at each participating site will record participant data on the study

forms. The final responsibility for the completeness and accuracy of all study data collected at a participating site resides with the LSI. After checking the forms for completeness and accuracy, the LSI or other designee will be required to sign and date each form. Data forms are faxed to the Perry Point CSPCC where data are processed and data checks are implemented to prevent errors and to prompt the sites for resolution. Data will be transmitted directly to the CSPCC improving the ability to monitor site progress and the status of the trial and participants.

When a participating site has a potential study participant that meets all of the eligibility requirements, the LSI or SC (or other local study team member designee) will assign a unique participant ID and Alpha code to the participant. This unique Participant ID number and alpha code will be entered on all study related forms for the duration of the study. Examples of study-related forms include the participant's data collection forms, data quality reports, etc. for the remainder of the study.

The DataFax system will have built-in edit checks on data fields to minimize data errors, such as missing, inconsistent, or extremely unusual data. The data management staff at the coordinating center will generate data query forms which will be sent to the sites for clarifications. The study database will be continuously updated with new data and changes to previously submitted data. To notify the participating sites about missing or late forms, reports with pertinent information will be generated at a regular interval and will be posted on the study SharePoint site.

In addition, a summary report of all data submitted and problems identified will be generated for each participating site. This report will provide each site with a summary of their progress. The National Study Coordinator in the Chairman's Office will also be reviewing each site's progress to ensure that there are no unforeseen problems with the forms or with a particular participant.

5 GENERAL CONSIDERATIONS

5.1 Timing of Analyses

The Study Group (all of the LSIs) and Executive Committee will meet six to nine months after participant recruitment begins and at annual intervals thereafter until the end of the study. Three weeks prior to these meetings and at six-month intervals between the meetings, these groups will be provided a report that will allow them to assess study progress. These reports will contain information on:

- Screening, Enrollment and Retention
- Participant background characteristics at entry
- Data quality and protocol adherence
- Participant characteristics during observation phase

Reports will be created for the Data Monitoring Committee starting approximately six months after the official start of the study. Data monitoring committee reports will be created every six months thereafter.

The final analysis will be performed on data which will have been documented as meeting the cleaning and approval requirements of CSP SOPs and after the finalization and approval of this SAP document.

5.2 Analysis Populations

The exact process for assigning the statuses will be defined and documented prior to final analyses along with any predefined reasons for eliminating a subject from a particular population.

All Study Participants
This population includes all participants who signed the informed consent form to participate in the study.

Intent-to-Treat (ITT)
This population includes all subjects who are randomized to the study.

Per-Protocol Population (PP)
This population includes all participants who are paired with a dog.

Safety
This population includes all subjects who have signed informed consent. Adverse events are recorded for the duration of the participant's study participation and may include serious adverse events for up to one month after participation ends.

5.3 Missing Data and Imputations

The primary analyses will analyze the data as observed.. In the hopefully unlikely event that a dog needs to be replaced after pairing (e.g., dog dies, gets hurt, becomes aggressive), then the participant's data will be included in the analysis up to this point. The participant will still be offered another dog and will remain enrolled in the study for the remainder of the 18 months. Data will be collected after the dog has been replaced, however it will not be used in the primary analysis.

However, some missing data will arise. When missing data are encountered in the analyses and considered substantial, a detailed sensitivity analysis will be conducted on the effects of various assumptions about the missing data. Within the taxonomy of Little and Rubin, mechanisms will be a mixture of missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR). MAR assumes that missing status is due purely to observed variables while MNAR allows missing status to also depend on the value of the missing datum. An example of MCAR is when a participant moves away. The sensitivity of results to assumptions about the missing-value mechanism

(MAR and MNAR) will be examined by analyzing the data with methods appropriate to each. Multiple imputation will be performed under the assumption that missing data are MAR (Little and Rubin, 2002). Assuming MNAR, we will fit the data to pattern-mixture models, assessing sensitivity to various identifying restrictions, as needed (Thijs et al., 2002).

5.4 General Considerations

This section details general policies to be used for the statistical analyses. Departures from these general policies may be given in the specific detailed sections of this statistical analysis plan. When this situation occurs, the rules set forth in the specific section take precedence over the general policies. The following policies will be applied to all data presentations and analyses.

- All p-values will be rounded to 3 decimal places. All p-values that round to 0.000 will be presented as '<0.001' and p-values that round to 1.000 will be presented as '>0.999'. Any p-value $\leq \alpha$ will be considered statistically significant and will be marked with one asterisk (e.g., 0.026*).
- Summary statistics will consist of the number and percentage of responses in each category for discrete variables, and the mean, median, standard deviation (SD), minimum, and maximum for continuous variables.
- All mean and median values will be formatted to one more decimal place than the measured value. Standard deviation values will be formatted to two more decimal places than the measured value.
- All percentages will be rounded to one decimal place. The number and percentage of responses will be presented in the form XX (XX.X), where the percentage is in the parentheses. The decimal of the percentage may be dropped due to space constraints when creating a table.
- All listings will be sorted for presentation in order of treatment group, site number, subject number, and date of procedure or event.
- When necessary for analysis purposes, partial dates will be completed (i.e., turned into complete dates) using the most conservative approach.
- All analysis and summary tables will have the population sample size for each group or treatment group in the column heading.
- The baseline for the WHO-DAS 2.0, PCL, PSQI, VR-12, C-SSRS and PHQ-9 is the Home Visit (or Baseline Visit).
- Version 9.3 of SAS® or higher will be the statistical software package used to produce all summaries, listings, statistical analyses, and graphs.
- Updated version of MedDRA will be used for adverse event and pre-treatment coding.
- The current version of the World Health Organization (WHO) drug dictionary will be used for the coding of medications.

5.5 Interim Monitoring

The analysis of study data to monitor for adverse or beneficial treatment effects will be the responsibility of the CSPCC. During the course of this study, reports will be prepared and distributed to the various groups that are responsible for the conduct of the study. These groups are the Executive Committee, the Data Monitoring Committee, the VA's Central IRB, and the Cooperative Studies Scientific Evaluation Committee (CSSEC). The Study Chairmen, the Executive Committee and the Cooperative Studies Scientific Evaluation Committee will not be privy to any of the results that compare the outcomes of the two treatment groups.

5.5.1 Data Monitoring Committee

An independent oversight committee called the Data Monitoring Committee (DMC) will monitor study progress. This committee meets on the same basic schedule as the Study Group and Executive Committee, i.e., at 6 to 9 months after the start of participant recruitment and yearly thereafter. Initially, the DMC will meet to become acquainted with the study and to establish monitoring guidelines.

The main responsibility of the DMC members is to make a recommendation to the Director of the Cooperative Studies Program on whether the study should continue or not based on the reviews of the progress reports submitted to them. The study could be recommended for termination due to poor recruitment, treatment differences so large that it would be possible to reach a final decision about the main question of the study, treatment differences so small that continuation would be irresponsible, or due to safety concerns. The DMC also reviews the participating sites' performance in terms of recruitment, adherence to the protocol etc., and makes recommendations on them. Their final responsibility is to review all proposed protocol changes and suggested sub-protocols and to make recommendations in regards to their acceptability.

In order for the DMC to carry out its responsibilities, the CSPCC Study Team will provide the committee with a report approximately three weeks prior to their meetings. The report will consist of the tables describing number of participants enrolled, expected enrollment, study progress, baseline demographics, and observational phase data as well as tables presenting outcomes. It is the responsibility of the CSPCC Study Team to provide the DMC with whatever information the committee feels that it needs to successfully monitor the study. Thus, additional tables will be added as required by the DMC. In addition to the reports for the yearly meetings, the DMC will also be provided with reports between meetings at 6-month intervals.

The Data Monitoring Committee will receive reports from the analyses about three weeks prior to their annual meetings and at six-month intervals between meetings. These reports will include data on the participants enrolled, with an emphasis on a description of the participants at entry into the study, description of reasons for exclusion, dog behavior, and study and individual center performance with regards to participant intake and follow-up. Adverse event (AE) and serious adverse event (SAE) data will also be presented.

5.5.1.1 Stopping Rules

The DMC determined at their initial meeting that they would not utilize an interim analysis of the primary outcome to determine stopping rules. Instead they will analyze the safety data from the study, including AEs and SAEs, to determine if the study should terminate. This will continue throughout the study.

5.5.2 Executive Committee

The Executive Committee is the management and decision-making body for the operational aspects of the study and will monitor the performance of participating medical centers and the quality of data collected. The Executive Committee will formulate publication plans and will oversee the publication and presentation of all data from the study. The Executive Committee is comprised of the Chairmen, PPCSP Biostatistician and Project Manager, LSIs and Subject Matter Experts.

The Executive Committee will be presented with data that will allow the group to assess the quality of the data being submitted and how well the sites, in general, are adhering to the protocol. These data will also be presented by site, so sites performing substantially below average can be identified and remedial action can be taken to improve performance. One piece of information that will be routinely provided is the number of forms that are missing according to the participant's assessment schedule.

6 SUMMARY OF STUDY DATA

All continuous variables will be summarised using the following descriptive statistics: n (non-missing sample size), mean, standard deviation, median, maximum and minimum. The frequency and percentages (based on the non-missing sample size) of observed levels will be reported for all categorical measures. In general, all data will be listed, sorted by site, treatment (or screening) and subject. Summary tables with treatment should be structured with a column for each treatment group. For reporting data at study phases and visits, the order In Screening, Observation Phase, Clearing (waiting for pairing), and pairing visits should be used. Summary tables will be annotated with the total population size relevant to that table/treatment group.

6.1 Subject Disposition

Subject disposition will be summarized for all participants that signed the consent form, were randomized, and were paired. The following data will be presented:

- The number of participants enrolled and expected to be enrolled by site (TABLE 3).
- The number of participants screened and randomized by month for each site (TABLE 4).
- A cumulative summary of the participants in screening by site (TABLE 5).

- A listing of participants that discontinued from the study during screening. The listing will include the reasons for the exclusion of the participant from the study (TABLE 6).
- A figure of participants enrolled and expected by site (FIGURE 1).
- Subject disposition by site for all active participants (TABLE 7). The number of subjects that were randomized at each study center and the number and percentage of subjects that completed or discontinued after randomization at each study center will be summarized for each treatment group and for all participants.
    - The termination CRF will be used to determine randomized subjects who discontinued prematurely from the study.

6.2 Demographics

Subject demographics will be summarized for all study participants. Demographic characteristics of all study participants will include age, gender, ethnicity, income, and marital status. The summary will include:

- The number and percentage of participants with each category of gender, ethnicity, marital status, and income (TABLE 8).
- The sample size, mean, median, SD, minimum and Maximum values for the following:
    - Age
        - Age will be calculated using the date the participant was consented (year on Form 01) and the participant birth date (birth date on Form 01).

Estimates of the baseline demographic characteristics will be presented for all study participants. Tables summarizing the important background characteristics by site will be prepared and submitted to the Study Group to provide an idea of the population being studied and based on this information, comparisons of the participant characteristics among the sites will be possible.

6.3 Baseline Variables

6.3.1 WHO-DAS 2.0

The sample size, mean, median, SD, minimum, and maximum values for the WHO-DAS summary scores will be presented for the baseline visit.

6.3.2 Depression

The sample size, mean, median, SD, minimum, and maximum values for the PHQ-9 will be presented for the baseline visit.

6.3.3 PTSD

The sample size, mean, median, SD, minimum, and maximum values for the PCL-5 will be presented for the baseline visit using the total summary score.

6.3.4 VR-12

The sample size, mean, median, SD, minimum, and maximum values for the PCS and MCS scores from the VR-12 will be presented for the baseline visit.

6.3.5 PSQI

The sample size, mean, median, SD and minimum and maximum values for the 7 component scores including sleep duration, sleep disturbance, sleep latency, day dysfunction due to sleepiness, sleep efficiency, overall sleep quality and needs medications to sleep will reported for the baseline visit. In addition, the sample size, mean, median, SD and minimum and maximum values for the overall PSQI total score will be presented for the baseline visit. See Appendix 2 for PSQI scoring algorithms.

6.3.6 C-SSRS

The number and percentages of participants with suicidality will be reported for the baseline visit.

6.4 Treatment Exposure

The number and percentages of randomized subjects in each treatment group by Month.

- Months of treatment will be calculated using date of pairing (day the training begins) for the Service Dog group.
- Months of treatment will be calculated using date dog was received for Emotional Support Dog group.

7 STATISTICAL ANALYSES

7.1 Primary Outcomes

The primary objective of this study is to determine how the provision of a Service Dog or an Emotional Support Dog impacts activity and quality of life for Veterans with PTSD. Primarily, the objective is to compare activity limitations and quality of life change, relative to baseline, over the 18-month intervention period between the two groups of Veterans with PTSD, i.e., those who receive Service Dogs and those who receive Emotional Support Dogs. There are two primary outcomes in this study: 1) improvement in activity as assessed by the total WHO-DAS 2.0 score and 2) improvement in quality of life as assessed by the Physical Component Scale (PCS) and Mental Component Scale (MCS) of the VR-12. The outcome measure for the first primary objective will be the change in the WHO-DAS 2.0 score over the 18 months of the intervention phase adjusted by baseline scores. The outcome measures for the other primary objectives will be the relative change in the PCS and MCS scores over the 18 months of the intervention phase adjusted by the baseline scores. The WHO-DAS 2.0 and VR-12 will be assessed at baseline, clearing and at 3, 6, 9, 12, 15 and 18-month post-pairing visits.

To begin the analyses, the distribution of each variable will be examined for outliers and to determine if transformations (to normalize data) are necessary. The mean and standard deviation of the activity level (WHO-DAS 2.0 summary score) and quality of life (PCS and MCS score) will be calculated at each time point during the screening/clearing and then during the follow-up phase (TABLES 9 & 10). During the follow-up phase, the activity level will be calculated separately for those veterans paired who were assigned to the Service Dog group as well as the Emotional Support Dog group.

The following are hypothesized for this study:

- For the WHO-DAS 2.0 outcome: Veterans who receive Service Dogs will have decreased activity limitations over time, as compared to Veterans who receive Emotional Support Dogs.

- For the VR-12 outcomes, Veterans who receive Service Dogs will have improved quality of life over time, as compared to Veterans who receive Emotional Support Dogs.

For the first hypothesis, a generalized linear mixed model (GLMM) will be used to determine changes over time between groups. Thus the level-1 units consist of the repeated measures, WHODAS 2.0 score, for each subject, and the level-2 unit is the individual or subject. In addition to estimating overall parameter estimates, multilevel modeling for repeated measures allows regression equations at the level of the individual (Curran, 2010). One of the major advantages of multivariate multiple regression is that outcomes with more missing data can "borrow information" from outcomes with less missing data (McCulloch, 2008). The use of mixed models allows for control of covariance data expected in clustered and repeatedly sampled data, and missing data. The statistical tests for the primary outcome measure will be two-sided at a 5% Type I error rate. Confidence intervals will be two-sided with a 95% confidence level.

The primary analysis will include gender and site in the model. Other variables will be examined, including other demographic factors, to determine if any significant differences exist. If a variable is considered to be a potential confounder, it may be included in the model as a covariate. See TABLE 13 for example reporting tables.

For the second hypotheses, a generalized linear mixed model (GLMM) will be used to determine changes over time on PCS and MCS scores between groups. The level-1 unit will be the repeated measure. In the first model, PCS will be the level-1 unit; in the second model, MCS will be the level-1 unit. The level-2 unit in both models will be the individual or subject. The statistical tests for the primary outcome measure will be two-sided at a 5% Type I error rate. Confidence intervals will be two-sided with a 95% confidence level. As described above, any variable considered to be a potential confounder may be added to the models.

An additional analysis for all hypotheses will use the linear repeated mixed models, with random intercepts. When random coefficients are specified, each subject has its own regression equation, making it possible to evaluate whether subjects differ in their means and/or response patterns over time. The WHODAS 2.0 summary scores will be regressed on time and the group x time interaction with random intercepts added for participants that will account for the correlation among repeated measures.

7.2 Secondary and Tertiary Outcomes

7.2.1 To examine how mental health is impacted by the provision of a Service Dog versus an Emotional Support Dog.

Mental health will be examined using the PCL-5 total score, C-SSRS based suicidality, the PHQ-9 total score, and the seven components and the total score of the PSQI. To begin the analyses, the distribution of each variable will be examined for outliers and to determine if transformations (to normalize data) are necessary. The mean and standard deviation of PCL-5 total scores, PHQ-9 total scores, PSQI seven component scores and the PSQI total score will be calculated for each time point. The number and percentage of participants reporting suicidality will also be reported for each time point. Data will be presented in two ways: 1) for all participants paired and 2) by treatment group.

Secondary continuous variable analyses will include linear repeated mixed model analysis on PTSD Symptom Severity using PCL-5, depression, sleep and thoughts of suicide. Any variable considered to be a potential confounder will be included in the model. Suicidality will be examined using logistic regression methods.

Open-ended questions are asked in the Dog Related Questionnaire. The purpose of these questions is to gain further information about the human-animal bond that occurs for Veterans who have PTSD and are paired with a Service Dog or Emotional Support Dog. The questions will be entered as text-based fields into the database. The categorized responses to these questions will help guide future planning needs, if a program is implemented.

7.2.2 To characterize and compare how health care utilization and costs are affected by the provision of a Service Dog or Emotional Support Dog.

The health care utilization and cost analyses will utilize VA databases, participant records and the Health Economics Research Center (HERC) developed non-VA utilization questionnaire. These analyses will be conducted by HERC. For a detailed description of the healthcare utilization and cost analyses, see the protocol.

7.2.3 To characterize and compare how employment and productivity are affected by the provision of a Service Dog or Emotional Support Dog

The employment and productivity analyses will utilize the WPAI: SHP V2.0 questionnaire and will be conducted by the Health Economics Research Center (HERC). For a detailed description of the healthcare utilization and cost analyses, see the protocol.

8 ADVERSE EVENTS

An Adverse Event (AE) can be any unfavorable or unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the study interventions. For the purposes of this study, the study interventions are 1) being paired with a Service Dog and 2) being paired with an Emotional Support Dog. In this study, information on AEs related to or possibly related to study intervention, for the participant, participant family members or the study dog, and on all serious adverse events (SAEs) will be collected and recorded.

Adverse events will be recorded during screening, observation phase and during the 18 months of follow-up. The reporting period for AEs begins when the participant signs the informed consent form and continues until the earlier of the 30 days after the participant's completion or early termination of study participation or the end of the study. All adverse events with a reasonable causal relationship to the study intervention should be considered "related". A definite relationship does not need to be established. The following levels of relatedness will be used in this trial:

- Not attributed to a study intervention (study dog)
- Possibly attributed to a study intervention (study dog)
- Attributed to a study intervention (study dog)

Only possibly related or related events must be reported on a study form.

Incidence of adverse events will be summarized for each treatment group by body system and MedDRA term. The number and percentage of participants with each body system and MedDRA term will be presented for each treatment group. Tables to summarize the incidence rates will be created for each of the following groups:

- Adverse events
- Adverse events by relationship to treatment
- Adverse events leading to premature discontinuation
- Adverse events presented in descending order of frequency by MedDRA term (no System Organ Class shown)
- Serious adverse events
- Serious adverse events by type
- Serious adverse events by relationship to treatment

Adverse events that led to premature discontinuation from the study will be listed. Serious adverse events will also be listed. These listings will contain details about the adverse event such as intensity and relationship to study treatment. All adverse events will be coded with MedDRA (updated version).

9 STATISTICAL PROGRAM VALIDATION PLAN

Perry Point Work Instruction (WI) 202 – Validation Plan for SAS Statistical Programs will serve as the validation plan for validation of the statistical programs created for the analyses of data collected during this study.

REFERENCES

Curran PaJ, Obeidat, K, Losardo, D. (2010). "Twelve Frequently Asked Questions About Growth Curve Modeling". *Journal of Cognition and Development* **11** (2): 121–136.

Forbes D. Hawthorne G. Elliott P McHugh T. Biddle D. Creamer M. Novaco RW. A concise measure of anger in posttraumatic stress disorder. J O Traumatic Stress, 2004, 17: 249-256.

Gellis LA, Mavandadi S, Oslin DS. Functional Quality of life in full versus partial post traumatic stress disorder among veterans returning from Iraq and Afghanastan. Prim Care Companion, J Clin Psychiatry, 2010, 12(3).

Iqbal SU, Rogers W, Selim A, Qian S, Lee A, Ren XS, Rothendler J, Miller D, Kazis LE
The Veterans Rand 12 item health survey (VR-12): What it is and how it is used.
Center for Health Quality, Outcomes, and Economic Research, A Health Services Research and Development Center of Excellence, VA Medical Center, Bedford, MA, USA, 2009.

Little, R.J.A., Rubin, D.B. (2002). Statistical analysis with missing data, second edition. Hoboken: John Wiley & Sons, Inc.

McCulloch, C. (2008). Joint modeling of mixed outcome types

Oquendo MA, Halberstam B, Mann JJ. Risk Factors for Suicidal Behavior: The Utility and Limitations of Research Instruments. In M.B. First (Ed.). Standardized Evaluation in Clinical Practice, Chapter 4, pp. 103-130, 2003.

Reilly MC, Zbrozek AS, Dukes E: The validity and reproducibility of a work productivity and activity impairment measure. *Pharmacoeconomics.* 1993;4(5):353-365.

Thijs, H., Molenberghs, G., Michiels, B., Verbeke, G., & Curran, D. (2002). Strategies to fit pattern-mixture models. Biostatistics, 3, 245-265.

Weathers, F.W., Litz, B.T., Keane, T.M., Palmieri, P.A., Marx, B.P., & Schnurr, P.P. (2013). *The PTSD Checklist for DSM-5 (PCL-5)*. Scale available from the National Center for PTSD at www.ptsd.va.gov.

APPENDIX 1: TABLES

TABLE 3: Number of Participants Entered and Number Expected

Site	Number Consented	Number Randomized	Number Expected	Percent of Expected
Atlanta				
Iowa City				
Portland				
Total				

TABLE 4: Participants Screened/Randomized by Month

Site	Month 1		Month 2		Month 3		Total	
	Screened	Randomized	Screened	Randomized	Screened	Randomized	Screened	Randomized
Atlanta								
Iowa City								
Portland								
TOTAL								

FIGURE 1: Observed Versus Expected Participant Recruitment

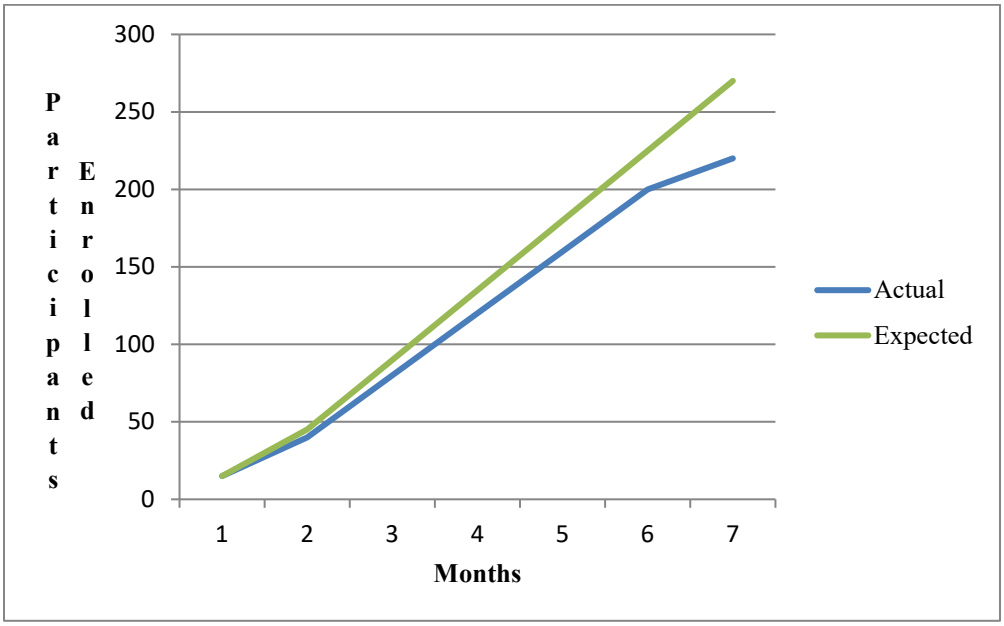

TABLE 5: Cumulative Screening Summary: All Participants by Site

		Status of Consented Participants			Status of Randomized Participants	
Site	Consented	Screen Failure	In Screening	Randomized	In obs. Phase	Paired
Atlanta						
Iowa City						
Portland						
Total						

TABLE 6: Summary of Ineligibility: Reasons for Exclusion, Total and Site

TOTAL NUMBER SCREENED = _____

Reason	No. Excluded	% of Screened
Less than 18 years of age		
No Referral from Mental Health provider, which documents PTSD diagnosis.		
Does not have PTSD as a result of a trauma by meeting DSM 5 criteria		
Not Enrolled in mental health at VA or has not attended at least one visit in the past 90 days prior to consent.		
Does not agree to remain in mental health therapy throughout the duration of the study.		
Cannot adequately care for a dog.		
Home environment is not suitable for a dog		
Home environment is not accessible for study staff.		
Does not have anyone to care for dog in the long-term absence of Veteran		
Others in home do not agree to have a dog		
Not willing and/or able to travel (by air or car) to training site for pairing		
Subject has cats, dogs or other household pets in home		
Not able to verbalize understanding of consent form or provide written informed consent		
Was hospitalized in the past 6 months for mental health reasons		
Participant has aggressive behavior what would make it unsafe for dog		
Diagnosis of psychoses, delusions, dementia, active alcohol/substance dependence or moderate to server TBI.		

Active suicidal intent as determined by a CPRS flag for suicidal intent or a score ≥ 5 on the C-SSRS completed at the baseline visit.		
Homicidal intent or cognitive disabilities that would preclude safety of dog and/or ability to participate in study.		
Social, mental or physical condition that prevents participant from giving informed consent		
Participation in another research trial		
Has chart note flag for violent/disruptive behavior		
Has children younger than age 10 in the household for more than 8 hours per day, on day a week or more.		

TABLE 7. Participant Disposition

Site	Consented	In Screening	Randomized	In Observation Phase	Waiting for Clearing Visit	Waiting for Pairing	Emotional Support Dog	Service Dog	Completed Study
Atlanta									
Iowa City									
Portland									
TOTAL									

TABLE 8: Demographic Variables by Site

Variables	Participating Site				
	1	**2**	**3**	**TOTAL**	**p-value**
Age (Mean(SD))					
Gender (N (%))					
Male					
Female					
Ethnicity (N (%))					
Asian-American					
African-American					
Caucasian					
Hispanic					
Native-American					
Other					
Marital Status (N (%))					
Married					
Divorced					
Civil Committment					
Never Married					
Co-habitating					
Separated					
Widowed					
Income (N (%))					
< $10,000					
$10,001 - $20,000					
$20,001 - $30,000					
$30,001 - $40,000					
$40,001 - $50,000					
$50,001 - $60,000					
$60,001 - $70,000					
> $70,001					

TABLE 9. Baseline, Observational Phase and Clearing Visit WHODAS 2.0 Scores

	Baseline Mean ± SD	Clearing Visit Mean ± SD
Understanding & Communicating		
Mobility		
Self-Care		
Getting along with Others		
Life Activities		
Participation in Society		
Summary Score		

TABLE 10. WHODAS 2.0 Scores by Treatment Group

	3 month Mean ± SD		6 month Mean ± SD		9 month Mean ± SD		12 month Mean ± SD		15 month Mean ± SD		18 month Mean ± SD	
	SERV	EMOT	SERV	EMOT	SERV	EMOT	SERV	EMOT	SERV	EMOT	SERV	EMOT
Understanding & Communicating												
Mobility												
Self-Care												
Getting along with Others												
Life Activities												
Participation in Society												
Summary Score												

TABLE 11. Baseline, Observational Phase and Clearing Visit VR-12 Scores

	Baseline Mean ± SD	Clearing Visit Mean ± SD
MCS		
PCS		

TABLE 12. VR-12 Scores by Treatment Group

	3 month Mean ± SD		6 month Mean ± SD		9 month Mean ± SD		12 month Mean ± SD		15 month Mean ± SD		18 month Mean ± SD	
	SERV	EMOT	SERV	EMOT	SERV	EMOT	SERV	EMOT	SERV	EMOT	SERV	EMOT
MCS												
PCS												

TABLE 13. GLMM Results for daily activities and quality of life

		Treatment Study Month					
	Baseline	3	6	9	12	15	18
EMOT %							
N							
SERV %							
N							

Fit Statistics	
-2 Res Log Pseudo-Likelihood	
Generalized Chi-Square	
Gener. Chi-Square/DR	

Covariance Parameter Estimates			
Cov Parm	**Subject**	**Estimate**	**Standard Error**
SP(EXP)	Intercept		
Residual			

	Type III Tests of Fixed Effects			
MODEL	Num DR	Den DF	F	PR > F

APPENDIX 2:

Pittsburgh Sleep Quality Index (PSQI)

Form Administration Instructions, References, and Scoring

Form Administration Instructions

The range of values for questions 5 through 10 are all 0 to 3.

Questions 1 through 9 are not allowed to be missing except as noted below. If these questions are missing then any scores calculated using missing questions are also missing. Thus it is important to make sure that all questions 1 through 9 have been answered.

In the event that a range is given for an answer (for example, '30 to 60' is written as the answer to Q2, minutes to fall asleep), split the difference and enter 45.

Reference

Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ: The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Research* 28:193-213, 1989.

Scores – reportable in publications

On May 20, 2005, on the instruction of Dr. Daniel J. Buysse, the scoring of the PSQI was changed to set the score for Q5J to 0 if either the comment or the value was missing. This may reduce the DISTB score by 1 point and the PSQI Total Score by 1 point.

PSQIDURAT **DURATION OF SLEEP**

IF Q4 $\geq$ 7, THEN set value to 0
IF Q4 < 7 and $\geq$ 6, THEN set value to 1
IF Q4 < 6 and $\geq$ 5, THEN set value to 2
IF Q4 < 5, THEN set value to 3
Minimum Score = 0 (better); Maximum Score = 3 (worse)

PSQIDISTB **SLEEP DISTURBANCE**

IF Q5b + Q5c + Q5d + Q5e + Q5f + Q5g + Q5h + Q5i + Q5j (IF Q5JCOM is null or Q5j is null, set the value of Q5j to 0) = 0, THEN set value to 0

IF Q5b + Q5c + Q5d + Q5e + Q5f + Q5g + Q5h + Q5i + Q5j (IF Q5JCOM is null or Q5j is null, set the value of Q5j to 0) $\geq$ 1 and $\leq$ 9, THEN set value to 1

IF Q5b + Q5c + Q5d + Q5e + Q5f + Q5g + Q5h + Q5i + Q5j (IF Q5JCOM is null or Q5j is null, set the value of Q5j to 0) > 9 and $\leq$ 18, THEN set value to 2

IF Q5b + Q5c + Q5d + Q5e + Q5f + Q5g + Q5h + Q5i + Q5j (IF Q5JCOM is null or Q5j is null, set the value of Q5j to 0) > 18, THEN set value to 3

Minimum Score = 0 (better); Maximum Score = 3 (worse)

PSQILATEN **SLEEP LATENCY**

First, recode Q2 into Q2new thusly:

IF Q2 ≥ 0 and ≤ 15, THEN set value of Q2new to 0
IF Q2 > 15 and ≤ 30, THEN set value of Q2new to 1
IF Q2 > 30 and ≤ 60, THEN set value of Q2new to 2
IF Q2 > 60, THEN set value of Q2new to 3
Next
IF Q5a + Q2new = 0, THEN set value to 0
IF Q5a + Q2new ≥ 1 and ≤ 2, THEN set value to 1
IF Q5a + Q2new ≥ 3 and ≤ 4, THEN set value to 2
IF Q5a + Q2new ≥ 5 and ≤ 6, THEN set value to 3

Minimum Score = 0 (better); Maximum Score = 3 (worse)

PSQIDAYDYS **DAY DYSFUNCTION DUE TO SLEEPINESS**

IF Q8 + Q9 = 0, THEN set value to 0
IF Q8 + Q9 ≥ 1 and ≤ 2, THEN set value to 1
IF Q8 + Q9 ≥ 3 and ≤ 4, THEN set value to 2
IF Q8 + Q9 ≥ 5 and ≤ 6, THEN set value to 3
Minimum Score = 0 (better); Maximum Score = 3 (worse)

PSQIHSE **SLEEP EFFICIENCY**
Diffsec = Difference in seconds between day and time of day Q1 and day Q3
Diffhour = Absolute value of diffsec / 3600
newtib =IF diffhour > 24, then newtib = diffhour – 24
IF diffhour ≤ 24, THEN newtib = diffhour
(NOTE, THE ABOVE JUST CALCULATES THE HOURS BETWEEN GNT (Q1) AND GMT (Q3))
tmphse = (Q4 / newtib) * 100

IF tmphse ≥ 85, THEN set value to 0
IF tmphse < 85 and ≥ 75, THEN set value to 1
IF tmphse < 75 and ≥ 65, THEN set value to 2
IF tmphse < 65, THEN set value to 3
Minimum Score = 0 (better); Maximum Score = 3 (worse)

PSQISLPQUAL **OVERALL SLEEP QUALITY**

Q6
Minimum Score = 0 (better); Maximum Score = 3 (worse)

PSQIMEDS **NEED MEDS TO SLEEP**
Q7
Minimum Score = 0 (better); Maximum Score = 3 (worse)

PSQI **TOTAL**
DURAT + DISTB + LATEN + DAYDYS + HSE + SLPQUAL + MEDS
Minimum Score = 0 (better); Maximum Score = 21 (worse)
Interpretation: TOTAL $\leq$ 5 associated with good sleep quality

TOTAL > 5 associated with poor sleep quality