



**Office of Research and Development  
Veterans Health Administration  
Department of Veterans Affairs**

January 5, 2021

David Allison, PhD  
Chair, Review Committee, VA PTSD/Service Dog Study  
Institute for Laboratory Animal Research  
National Academies of Science, Engineering, and Medicine

Re: NASEM Review of “A randomized trial of differential effectiveness of service dog pairing versus emotional support dog pairing to improve quality of life for Veterans with PTSD”

Dr. Allison:

On behalf of the VA study team, I would like to express my appreciation to the committee for its timely review of the 2<sup>nd</sup> draft of Monograph 1, revised in response to the committee’s review of the original draft. We have carefully reviewed the committee’s letter dated 12/22/20 and made further revisions to Monograph 1 to resolve the outstanding concerns and suggestions for improvement, as much as reasonably possible. This 3<sup>rd</sup> and final draft is attached for committee’s reference, as well as an appendix detailing VA response actions to all suggestions and concerns. One change not requested by the committee is the addition of language describing the C-SSRS data, which needed to be addressed directly in the Abstract due to its importance to Veterans. The added language is shown in Row 8 of the appendix.

VA thanks the committee for their acknowledgement that the study was well executed and that great care was taken to address and safeguard ethical and animal welfare concerns. The NASEM committee’s comments prompted substantial improvements to the manuscript, and the study team took all suggestions and concerns very seriously. We appreciate the committee’s acknowledgement that the manuscript was substantially improved. Underlying all our responses is an appreciation for the committee’s diligent review and constructive criticism throughout this process. The study findings are expected to impact eligibility criteria for VA’s service dog insurance benefit program.

In closing, we sincerely appreciate the committee’s efforts and the time commitment that made the review possible, particularly in light of pandemic-related disruptions.

Sincerely,

A handwritten signature in black ink that reads "Joan T. Richerson". The signature is written in a cursive, flowing style.

Joan T Richerson  
Study Chair

Appendix: VA Responses to NASEM Committee Review of the 2nd Draft of Monograph 1  
(Letter dated 12/22/2020)

## Appendix: VA Responses to NASEM Committee Review of the 2<sup>nd</sup> Draft of Monograph 1 (Letter dated 12/22/2020)

<b>High-level Concerns</b>	
	<b>Specificity of Language and Interpretation of Results</b>
1	<p><u>Committee concern:</u> More effort should be made to reduce the use of causal language when describing the results of the study; for example, replacing instances where the word “benefit” is used to describe “improvements” that occurred in both intervention groups and which might not be attributable to the interventions in this study.</p> <p><u>VA Response:</u> To prevent confusion, we have now defined the word benefit in the <i>Glossary of Terms</i> as follows:</p> <p><b>“Benefit – as defined for this study, an advantage or an improvement gained, which for study design reasons, may or may not be directly attributable to the intervention of a service dog as compared to an emotional support dog.”</b></p>
2	<p><u>Committee concern:</u> The addition of one sentence stating that the PCL-5 finding is not demonstrable of a clinical advantage does not sufficiently offset this shortcoming. As PCL-5 is a secondary outcome, following at least three primary tests of association for the primary outcomes (World Health Organization Disability Assessment Schedule 2.0; Veterans RAND 12 Item Health Survey [VR-12] Physical Component Score, VR-12 Mental Component Score), the importance of the intervention effect on PCL-5 must be interpreted in the context of the magnitude needed for clinical importance. In short, a statistically significant difference is not evidence of a clinically significant difference. The VA author team should consider further revisions to clarify this for readers.</p> <p><u>VA Response:</u> We have added an additional sentence in the Discussion section (second paragraph, page 103) to further clarify that the change in baseline, although statistically significant, is not necessarily considered clinically significant. The bolded red text has been added:</p> <p>“Examining the within-group findings reveals that Veterans receiving either intervention experienced improvements in disability and quality of life measures; the main group difference was reflected in the SERV intervention favoring a reduction in PTSD symptoms using the PCL-5. Specifically, participants paired with a SERV showed a 34.5% decrease in PCL-5 from baseline to study completion compared to a 25.0% PCL decrease over the same time period for participants paired with an EMOT. At month 18, the final PCL-5 score for those</p>

	<p>participants paired with a SERV was 31.66 versus 35.25 for those paired with an EMOT. <b>The data from this study did not necessarily demonstrate a clinically significant change from baseline. Nonetheless,</b> examinations of the PCL-5 have indicated that a score of 31 is a clinically relevant threshold associated with probable diagnosis of PTSD, and scores below 31 represent a symptom burden that may not require clinical intervention. Because the mean PCL-5 score for SERV participants is still slightly above the cut point of 31 (31.66) we could not substantiate that the SERV group has a demonstrable clinical advantage over the EMOT group (35.25) in PCL-5 scores (Blevins, et al., 2015).”</p>
	<p><b>Reiterating the Importance of the Intent to Treat Analysis</b></p>
<p>3</p>	<p><u>Committee concern:</u> The committee strongly recommends that the VA focus on the intent to treat (ITT) population in the analysis of all of the outcome measures, as the committee requested in its first report.</p> <p><u>VA Response:</u> Although we appreciate the committee’s concern, the ITT analysis did not result in any substantial differences for the primary outcomes versus the modified ITT analysis. We do not see any advantage to focusing on the ITT population, which requires a large amount of data imputation, which has the potential to introduce a substantial amount of error. The modified ITT analysis the team performed is an accepted analysis in randomized clinical trials involving more complex study designs. In section 4.A (Results: Enrollment, Randomization, and Pairing; page 59), additional text has been added to explain our approach:</p> <p><b>“In this study, participants randomized and not paired with a dog were not followed, and therefore, no primary or secondary outcome information was collected on these participants. Imputing 18 months of follow-up information solely based on baseline data was not reasonable as it potentially introduces a large amount of error, limits interpretation and creates a challenge from which to draw conclusions. If these participants who did not actually receive any treatment were included with those receiving treatment, it may result in a more conservative estimate of treatment effect, introduce heterogeneity (e.g., noncompliant and dropout participants mixed with compliant participants in the final analysis), and thus indicate very little about the efficacy of the treatment and make interpretation more challenging. Further, mITT is an acceptable analysis in RCTs involving more complex study designs such as this one. However, ITT analyses were performed, and found no substantial differences from that reported in this monograph observed for the study’s primary outcomes; therefore, the ITT population with all missing data imputed was not included.”</b></p>
	<p><b>Transparency</b></p>
<p>4</p>	<p><u>Committee concern:</u> The committee recognizes and respects the need for study participant privacy, yet the points regarding data sharing raised in the committee’s first report remain. Properly de-identified or perturbed data could be posted</p>

publicly. Additionally, sharing the code used for statistical analysis would raise no ethical concerns and would enable others to vet and understand the models used. Lastly, publication of the full statistical analysis plan (including dated amendments) would better enable evaluation of the monograph within the scientific community. These points are raised in the spirit of considering the monograph as a stand-alone document that contains all of the needed information in one place. For example, there are details about the construction of the imputation model included in the VA response document that should be incorporated into the monograph, including a more complete description of the statistical model, the covariates included in the model, the number of iterations generated, and the software and version utilized.

VA Response: We understand the committee’s reservations regarding data sharing, but we do have a process for sharing data with appropriate VA and non-VA groups, and our concerns with Veteran privacy take precedent over other considerations.

The details about the construction of the imputation model included in the 11/10/20 VA response document were previously incorporated in the monograph and are shown below. We believe that the detail provided is adequate, but of course we would always be willing to provide further information to interested parties upon request. The existing text is found on page 96 in the Results section, G. Attrition and Missing Data:

“Reasons for dropout in the two groups suggested no evidence of missingness not occurring at random, thus multiple imputation assuming missingness at random was employed. Multiple imputation (MI) using fully conditional specification (FCS) was performed for incomplete outcome data using the MI procedure available in SAS version 9.4 to assess the impact of missing data from differential dropout on outcome results. MI by FCS was performed using available outcome data, including complete and incomplete outcome variables, baseline assessment data, gender, and center. Imputation models included regression for continuous variables and discriminant function for binary/categorical variables, using 10 imputations. Ten replicant datasets were created where observed data were constant across the 10 datasets and missing data were imputed for each dataset. Data were pooled using the MIANALYZE procedure and analyzed using a generalized linear mixed model assuming a binomial distribution and logit link for the C-SSRS outcome and linear mixed repeated measures and random intercept models for all other outcomes. This creates unbiased estimates, assuming the differential dropout was missing at random. The results for the PP population are shown in Table XX. The results from multiple imputation using FCS for PP and PPDR (Appendix B – Part B: Table OOOO) revealed similar treatment effects to those of the linear mixed repeated measures models for all primary and secondary outcomes with the exception of the PCL-5, which was non-significant in imputed models.”

**Recognition of Ongoing Treatment**

5	<p><u>Committee concern:</u> The revised monograph, the cover letter, and the VA response document are all very clear that it would be “unethical to ask participants to stop utilizing their existing PTSD therapies to allow a comparison of a dog only group (utilizing an unproven potential mitigation for PTSD)” (revised monograph, p. 28). However, the impact of ongoing treatment on the interpretation of the study should be more explicitly stated and accounted for within the text.</p> <p><u>VA Response:</u> We have added additional language in the Discussion to address this concern (page 111). The bolded red text has been added:</p> <p>“Finally, due to the nature of the study design, all participants were required to stay in active treatment engagement with their assigned MH team (requirements were the same between the two interventions). <b>Active treatment engagement defined as their mental health care treatment as usual, and this concurrent PTSD treatment, including medications and/or therapy, was not controlled for over the course of the study.</b> This usual care study design requirement may have also contributed to improvements in PTSD functioning and disability secondary to adherence with MH care. <b>As noted previously, it would be unethical to ask participants to stop utilizing their existing PTSD therapies to allow a comparison of an unproven potential mitigation for PTSD; therefore, the impact of ongoing usual care treatment on the interpretation of the study cannot be readily quantified.</b>”</p>
---	--

**Minor Points**

6	<p><u>Committee Concern:</u> Page 69 (lines 1822-1831)</p> <ol style="list-style-type: none"> <li>1. It would be helpful to clearly define linear repeated measures mixed model and contrast it with the linear repeated measures mixed model with random intercepts. The reader may presume that the first, without random intercept, may be a standard linear repeated measures model with some covariance structure. The covariance structure assumptions should also be specified.</li> <li>2. It would be helpful to state here what type of contrast coding is used for inference (dummy, sum, effect coding, etc.).</li> </ol> <p><u>VA Response:</u> Language clarifying the points listed in items 1 and 2 have now been added on page 59 in the Methods section, ii, Mental Health Outcomes. Red bolded text has been added:</p> <p>“For all the mental health outcomes except suicidality, a <b>traditional</b> linear repeated measures mixed model was then used to determine changes over time between the SERV and EMOT groups with gender, center and the baseline score of the outcome measure included as covariates, as well as a time by treatment group interaction (with time as a categorical variable) <b>and using an unstructured covariance structure to model within-subject variation.</b> Adjustment using center, gender, and baseline scores were prespecified. Center was a stratification factor, and gender was included to account for any potential confounding gender effects. Using the same set of covariates, an additional analysis using a linear</p>
---	--

repeated measures mixed model with random intercepts was also employed **modeling between-subject variation**. The random intercept was based on the participant and corrects the standard errors given the repeated observations for each participant. Suicidality (per C-SSRS) was examined using a generalized linear mixed model, assuming a binomial distribution and logit link. **Linear contrasts testing for a difference between groups across time and at 18 months employed effect coding.**"

7 Committee Concern: Page 150 (line 3256) Reference listed as M.B. McConnell. This should be P.B. McConnell.

VA Response: Thanks for catching this error. The reference citation has been corrected on page 121:

- *McConnell, P.B. (2003). The Other End of the Leash: Why We Do What We Do Around Dogs. New York, NY: Ballantine Books.*

### **Additional Text Added by VA to Abstract**

8. Additional Text added by VA: Due to its importance and relevance to Veterans, the team added text to the Abstract as follows:

1. *Results* section, page 6. Red bolded text was added.

"227 participants were randomized to either the service dog intervention (n=114) or emotional support dog (n=113) intervention. 46 participants terminated prior to pairing; (n=17) participants assigned to the service dog intervention versus (n=29) participants assigned to the emotional support dog intervention. 97 participants were paired with a service dog; 84 participants were paired with an emotional support dog. 9 participants paired with a service dog terminated after pairing; 19 participants paired with an emotional support dog terminated after pairing. Participants paired with a dog were on average 50.6 years old (SD=13.6; range 22-79), mostly male (80.1%), white (66.3%), and non-Hispanic (91.2%). After adjusting for baseline score, center, and gender, the linear mixed repeated measures (LMRM) model for WHO-DAS 2.0 (disability) showed no statistical difference between the two intervention groups nor did the mixed models for quality of life (VR-12) show statistical differences between the two groups for either PCS (physical health) or MCS (mental health). Of the secondary outcome measures, only PCL-5 (PTSD symptoms) using the adjusted LRMR model showed a statistically significant difference between intervention groups. Participants receiving the service dog intervention had a 3.7-point improvement (lower score=less symptoms of PTSD) in the PCL-5 total score over time as compared to the emotional dog intervention. **Contrasts testing for a difference in the service dog group versus the emotional support dog group for suicidal ideation and behavior (per C-SSRS) did not show a significant difference between groups across time, however, it did show a difference between groups at 18 months with the service dog group having fewer suicidal behaviors and ideations.** In both groups, WHO-DAS 2.0 scores at 18 months decreased (less disability) from scores at 3 months post pairing; improvement in VR-12 MCS also showed some improvement over time in both groups."

2. Conclusions section, page 7. Red bolded text has been added:

“While both groups appeared to have experienced some benefit, an improvement in overall disability and quality of life among Veteran participants with PTSD was not observed with the provision of a service dog relative to provision of an emotional support dog. Among secondary outcome measures, participants paired with a service dog experienced a reduction in the severity of PTSD symptoms (PCL-5) compared to participants paired with an emotional support dog, **and had fewer suicidal behaviors and ideations, particularly at 18 months post-pairing.**”