A randomized controlled trial comparing a manual and computer version of CALM in VA community-based outpatient clinics

Michael A. Cucciare[a,b,c,*], **Kathy Marchant**[a], **Traci Abraham**[a,b,c], **Anthony Ecker**[d,e,f], **Xiaotong Han**[a,b,c], **Penny White**[c], **Michelle G. Craske**[g,h,i], **Jan Lindsay**[d,e,f]

[a]Center for Mental Healthcare and Outcomes Research, Central Arkansas Veterans Affairs Healthcare System, North Little Rock, AR, United States

[b]Department of Psychiatry, University of Arkansas for Medical Sciences, Little Rock, AR, United States

[c]VA South Central Mental Illness Research, Education and Clinical Center, North Little Rock, AR, United States

[d]Houston Veterans Affairs Health Services Research and Development Service Center for Innovations in Quality, Effectiveness, and Safety, Michael E DeBakey Veterans Affairs Medical Center, Houston, TX, United States

[e]VA South Central Mental Illness Research, Education and Clinical Center, Houston, TX, United States

[f]Menninger Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine, Houston, TX, United States

[g]Department of Psychology, University of California-Los Angeles, Los Angeles, CA, United States

[h]Department of Psychiatry and Biobehavioral Sciences, University of California-Los Angeles, Los Angeles, CA, United States

[i]Department of Psychology, Anxiety and Depression Research Center, University of California-Los Angeles, Los Angeles, CA, United States

Abstract

*Corresponding author: Department of Psychiatry, University of Arkansas for Medical Sciences, 4301 West Markham Street #755, Little Rock, AR 72205, United States. macucciare@uams.edu (M.A. Cucciare).

Background: This study compared a computer and manual version of a tailored Coordinated Anxiety Learning and Management (VA CALM) protocol on provider fidelity to CBT and patient outcomes.

Methods: This study was a cluster randomized controlled trial. Providers ($\underline{N}$ = 32) were randomized to deliver VA CALM by computer or manual. Veteran patients ($\underline{N}$ = 135), treated by study providers, were recruited. The primary outcome was CBT fidelity, measured by rating audiotaped sessions. Secondary outcomes were Veterans' general (BSI-18 GSI, SF-12) and disorder-specific (GAD-7, PCL-5, PHQ-9) outcomes assessed at baseline, three and six month follow-up.

Results: We found a large ($\underline{d}$ = 0.88) but not statistically significant difference in mean fidelity rating scores between conditions. Compared with the manual, participants with generalized anxiety disorder receiving VA CALM by computer reported lower GAD-7 scores at three (−5.88; 95% $\underline{CI}$=−11.37, −0.39) and six month (−5.25; 95% $\underline{CI}$=−10.29, −0.22) follow-ups ($\underline{d}$ = 0.37 to 0.55). Participants in the computer and manual conditions reported lower PHQ-9 (−3.11; 95% $\underline{CI}$=−5.51, −0.71; −4.06; 95% $\underline{CI}$=−7.22, −0.90, respectively) and BSI-18 GSI (0.78; 95% $\underline{CI}$=0.68,0.90; 0.71; 95% $\underline{CI}$=0.58, 0.87, respectively) scores from baseline to six month follow-up. We did not find statistically significant differences over time or between conditions on SF-12 or PCL-5 scores.

Limitations: This study was underpowered to test the primary outcome. Small samples sizes in the disorder-specific subgroup analysis may limit the generalizability of findings.

Conclusions: Neither modality proved to be superior on VA CALM fidelity. The computer version of VA CALM, compared to the manual, may provide modest benefit to Veterans with GAD.

1. INTRODUCTION

Evidence-based psychotherapies (EBPs), particularly Cognitive Behavioral Therapy (CBT) (Cape et al., 2010; Hoifodt et al., 2011), are highly effective for treating anxiety and depression, the most common mental health disorders in the primary care setting (Leon et al., 1995; Ansseau et al., 2004; Kroenke et al., 2007). As a result, the Department of Veterans Affairs (VA) has made it a national health care priority to increase the availability and accessibility of EBPs across VA medical facilities and clinics, to Veterans in need of mental health care (DVA, 2008).

Broadly disseminating and implementing (i.e., "scaling up") EBPs is a challenge in many large health care systems (Hepner et al., 2010), including the VA (Bidassie et al., 2015; Waxmonsky et al., 2014). Although training providers is necessary for scaling up, training alone is insufficient to ensure EBPs are implemented with enough fidelity to optimize their clinical effectiveness (Stetler et al., 2006). Without adequate implementation and fidelity,

patients may realize few benefits and the resources devoted to training providers may fail to improve outcomes.

Implementing EBPs in small and/or remote primary care clinics, such as VA Community-based Outpatient Clinics (CBOCs), is especially challenging. In a study of VA CBOCs, only 22% of Veterans with depression, anxiety, or posttraumatic stress disorder (PTSD) received at least one session of psychotherapy and rural, compared to urban, Veterans were less likely to receive any psychotherapy (Cully et al., 2010). Further, the quality of psychotherapy received in CBOCs remain unknown. Many mental health providers working in CBOCs, especially those located in rural geographical areas, are often the only mental health provider in the clinic and can be isolated from peers and resources needed to sustain their use of an EBP. Further, some CBOC providers have been trained in EBPs (Karlin et al., 2012; Drapkin et al., 2016, 2013) but may not have the resources or time needed to become proficient in these skills. Thus, efforts are needed to implement tools that support and assist CBOC mental health providers, especially those in working rural communities, in delivering EBPs, such as CBT, with high fidelity.

The National Institute of Mental Health-funded Coordinated Anxiety Learning and Management (CALM; Roy-Byrne et al., 2010; Craske et al., 2011; Sullivan et al., 2007) study faced a similar challenge in implementing CBT in non-VA primary care clinics across the United States. Most providers enrolled in the study, usually nurses and social workers, had no prior training in CBT. To meet this challenge, researchers developed CALM - a computer-delivered CBT program - to support primary care mental health providers in delivering CBT to patients with a range of disorders, including panic disorder (PD), generalized anxiety disorder (GAD), social anxiety disorder (SAD), posttraumatic stress disorder (PTSD), and depression (Craske et al., 2009).

The CALM program was created to guide and train mental health providers in using CBT. It involves the patient and provider viewing the computer screen together and proceeding through treatment modules at their own pace (Craske et al., 2009). Continuous use of the CALM program to guide delivery of treatment has the benefit of helping non-expert providers maintain fidelity to the CBT treatment (Craske et al., 2011).

The CALM program is unique in that it can be used to treat a range of common mental health disorders. This is accomplished through use of basic CBT content targeting symptoms that are shared across disorders plus additional branching content that is disorder-specific (Craske et al., 2011). CALM is acceptable to providers and primary care patients, and improves treatment engagement, homework compliance (Craske et al., 2009), and clinical outcomes, relative to standard care (medication, counseling in primary care, and/or referral to specialty mental health care), for persons with the target disorders (i.e., PD, GAD, SAD, PTSD, and depression) (Roy-Byrne et al., 2010; Craske et al., 2011). Moreover, clinical outcomes achieved appear to result from receipt of CBT delivered via the CALM computer program (Roy-Byrne et al., 2010). Thus, implementation of a program that guides CBT treatment delivery could assist VA CBOC providers in delivering high quality treatment leading to improved outcomes.

We conducted a randomized controlled trial comparing two methods (computer and manual) of delivering the VA CALM program on pre-registered outcomes (i.e., outcomes were registered with clinicaltrials.gov prior to completing the study) of CBT fidelity, and general and disorder-specific mental health symptoms in Veterans with GAD, SAD, PD, PTSD, and/or depression (Cucciare et al., 2016). The version of CALM evaluated in the present study was adapted specifically for Veterans (Abraham et al., 2018). The adapted version of CALM was developed into two modalities presenting the same content – one version delivered content by computer and one version delivered the same content in portable document format (PDF) that could be printed out (henceforth referred to as the manual). Our primary hypothesis was that, relative to the manual, providers using the computer-version of VA CALM would show increased fidelity to the CBT treatment model. This hypothesis was based on prior findings showing that the computer version of CALM can help non-expert providers adhere to and competently apply CBT (Craske et al., 2011). In VA, manuals have been the primary method for disseminating EBPs (Stewart et al., 2014; Drapkin et al., 2016) including CBT (Karlin et al., 2010). Therefore, it is important to determine whether a computer version of VA CALM confers any additional benefit to providers and patients over a manual version. If the computer version were to improve fidelity and/or outcomes, it could be implemented in VA to support CBOC mental health providers in delivering CBT.

Secondary hypotheses were that Veterans receiving VA CALM by computer, relative to the manual, would report fewer general mental health and disorder-specific symptoms at follow-ups. The two secondary outcomes were included given findings showing that patients receiving a computer version of CALM, relative to standard care (e.g., counseling in primary, specialty mental health) showed improvements in general (Roy-Byrne et al., 2010) and disorder-specific mental health symptoms (Craske et al., 2011).

2. Methods

2.1. Study design

This study was a randomized controlled trial conducted in 20 VA CBOCs in the South and Mid-South regions of the United States. The VA is a nationalized system of public health care available to individuals who have served in the military. CBOCs are smaller (compared to VHA medical centers) VA clinics designed to provide health care to the more than 4.7 million Veterans living in urban and rural locations (VHA ORH, 2018).

In this study, we compared two methods of delivering VA CALM. Thirty-two CBOC mental health providers were recruited and randomized to deliver VA CALM by computer or manual. A total of 135 Veteran participants, referred directly by enrolled providers, were also recruited. All provider and Veteran participants were informed that they were participating in a research study to evaluate a process for implementing VA CALM in VA CBOCs.

We obtained human participants approval from the Research and Development Committees at the Central Arkansas Veterans Healthcare System and VA Central Institutional Review Board and registered the study at clinicaltrials.gov.

2.2. Participants and recruitment

2.2.1. Provider recruitment—Between October 2016 and January 2017, mental health providers at CBOCs within Veterans Integrated Service Network (VISN) 16 (Arkansas, Louisiana, Mississippi) and 17 (Texas) were recruited to participate. VISNs are VA's geographically divided administrative areas within the United States. Providers were eligible if they provided behavioral health care in a VISN 16 or 17 VA CBOC. Recruitment was initiated by first obtaining permission from mental health leadership within each study VISN. VISN leadership forwarded an email containing a brief description of the study, its purpose, provider eligibility, and a description of provider involvement to all CBOC mental health providers. Interested mental health providers were asked to contact study staff at the Central Arkansas Veterans Healthcare System, the coordinating site for the study. Providers were asked to obtain approval from their immediate supervisor and clinic director before enrolling in the study. Study staff obtained providers' informed consent over the telephone. Thirty-four VA CBOC mental health providers from various disciplines gave their consent to participate in the study. Two providers dropped out of the study (due to changing jobs) prior to initiating the two-day VA CALM training workshop, resulting in 32 provider participants - social workers ($\underline{n} = 17$), psychologists ($\underline{n} = 11$), a psychiatrist ($\underline{n} = 1$), and nurses ($\underline{n} = 3$)

2.2.2. Veteran recruitment—Between March 2017 and December 2018, 135 Veterans were enrolled in the study (Fig. 1). Veterans were eligible if they had been clinically diagnosed with one of the target mental health disorders (GAD, SAD, PD, PTSD, or depression), were a patient of a CBOC mental health provider enrolled in the study, and agreed to receive up to eight weeks of CBT delivered by manual or computer. Veterans receiving on going CBT were excluded (none were deemed ineligible for this reason). Providers obtained Veterans' permission for the study team to contact them. Once consent to contact was obtained, a study team member, blinded to the referring provider's condition assignment, described the study and obtained consent by telephone.

2.3. Power analysis for primary outcome

We aimed to recruit a total sample of 34 (17 per condition) providers to achieve a statistical power of 0.94 for analysis comparing our primary outcome of CBT fidelity between the two conditions. We used a general linear mixed model to account for patients clustered within the same providers and a type I error rate of 0.05. This initial sample size was also determined by assuming an effect size of 1.0 (1.3 point difference on a scale of CBT fidelity; 5.3 of 6 for the computer condition versus 4 of 6 for manual condition; Craske et al., 2011), a medium intraclass correlation of 0.5, and four patients per provider.

A total of 16 of 32 clinicians self-selected to provide audiotaped sessions that were assessed for fidelity. Assuming an effect size of 1.0 for condition on the fidelity outcome, an intraclass correlation of 0.5, four participants per provider, and a type I error rate of 0.05, our statistical power for comparing our primary outcome of fidelity between conditions using a general linear mixed model is 0.65.

2.4. Procedure, randomization, and data collection

Upon receiving informed consent, providers were asked to participate in three phases of CALM training: (phase one) a two-day, in-person workshop to learn the CALM protocol, (phase two) three months of CALM supervision-consultation, and (phase three) six months of external facilitation (with technical support for those assigned to the computer condition). A detailed description of the training model used in this study is described in detail elsewhere (Cucciare et al., 2019).

In phase one, providers participated in a two-day workshop led by two experts in CALM. The workshop occurred at the Central Arkansas Veterans Healthcare System in North Little Rock, Arkansas. The first day and a half of the workshop consisted of an overview of the CALM treatment model such as basic CBT skills, an overview of disorders targeted in the CALM protocol, role-plays to practice skills, and question and answer sessions. Providers were also provided relevant articles, a CALM training workbook, and a description of the CALM treatment program with a summary of each CALM treatment session. During day two, providers were randomized to the computer ($n = 17$) or manual ($n = 15$) version of VA CALM and provided instruction on how to deliver the protocol using either modality. Given that prior experience with CBT may impact providers' ability to deliver CBT for some disorders (Craske et al., 2011), randomization was done within each of the two conditions (i.e., self-reported formal training or no formal training on prior CBT experience; referred to as strata). Randomization was conducted by the study statistician using Randomizer.org. Training in either modality included didactics, role-plays of "mock" therapy sessions, small group discussions, and problem-solving anticipated challenges to implementing the specific modality for each disorder in their practice.

In phase two, providers in both conditions were asked to participate in three months of weekly, group telephone delivered supervision-consultation meetings led by supervisors with expertise in CALM. Providers were grouped by condition. All providers attended at least two weeks of the supervision-consultation calls, with most attending at least six of the 12 calls. Weekly attendance on the supervision-consultation calls ranged from 58% to 78%.

In phase three, providers in both conditions received six months of external facilitation and technical support (for those in the computer condition) to promote the implementation of each modality of CALM at their CBOC. External facilitation began simultaneously with the supervision-consultation phase and occurred through telephone calls, emails, and face-to-face meetings. External facilitation was used to encourage and mentor providers, provide technical and logistical help, and problem-solve barriers to implementing CALM in clinical practice. To facilitate analysis of the primary outcome (CALM fidelity), all providers were asked to provide audio recordings of all CALM treatment sessions.

Veteran participants were referred directly from an enrolled provider and were not randomized into study conditions. A study staff member, blinded to the referring provider's condition, conducted the initial eligibility screening. The referring provider was also contacted to obtain the primary diagnosis for which the person was being treated. Once consented, participants completed a 45 to 60 min baseline assessment over the telephone. The same assessment, conducted by study staff blinded to participant condition, was

conducted over the telephone at three and six month follow-up. Follow-up rates were 85% at three month follow-up and 90% at six month follow-up. Participants were compensated US$25 for the baseline assessment, US$25 for the three month assessment, and US$30 for the six month assessment.

2.5. VA CALM intervention

VA CALM is based on the original CALM program which is a computer program that involves the patient and provider viewing the computer screen together and proceeding through CBT content at a personalized pace (Craske et al., 2011). CALM includes up to eight sessions of CBT content with modules that cover: symptom education; self-monitoring; hierarchy development; breathing retraining; relapse prevention; cognitive restructuring; and exposure to feared internal and external stimuli (Roy-Byrne et al., 2010). CBT content is tailored to each target disorder. When comorbid disorders were present, providers asked Veteran participants to select their most distressing and disabling disorder to be the primary target of treatment. Modifications to the CALM program for the current study were mostly related to its "look and feel", e.g., adding a VA web template, videos of Veterans describing their experiences with and in treatment for the target disorders, images that are representative of Veterans, and, based on Veteran preferences, allowing some CBT homework assignments to be audio recorded instead of written (Abraham et al., 2018).

In the present study, two modalities were used to deliver the same CALM content-the computer and manual. Providers in the manual condition received a printed pdf version of the same content included in the computer version. In both modalities, the same printable patient handouts (e.g., symptom education and tracking) were available to providers to give to Veteran participants to use in-session and/or take home.

2.6. Outcomes and measures

2.6.1. CALM fidelity—The primary outcome was fidelity to the CALM CBT treatment model (Cucciare et al., 2016). Fidelity was measured using the CALM Mental Health Provider Proficiency Evaluation Manual used to evaluate CBT fidelity in a prior study of CALM with civilians (Craske et al., 2011). Fidelity was defined as providers' overall competence in delivering the VA CALM protocol, for each rated session (i.e., How skilled was the mental health provider in delivering the information in this section?), rated on a Likert scale of 0 to 6 (higher scores indicate greater fidelity).

Two experts in CBT served as fidelity raters. Both raters received training in measuring CALM fidelity by two expert raters. Training included an overview of the CALM protocol and fidelity manual, practice rating VA CALM treatment sessions, and problem-solving any rating issues (e.g., questions about how to apply the fidelity rating criteria). To establish initial inter-rater reliability, both raters, blinded to study condition, independently rated an audiotaped treatment session and then shared their overall competence ratings to determine a consensus rating. Raters were asked to first determine if providers delivered the relevant core components of each VA CALM treatment session (e.g., the provider covered how to make an Avoidance List for the appropriate disorder; yes or no; and indicate overall competence on a scale of 0 [not at all] to 6 [fully]). Raters were then asked to evaluate the overall competence

of the provider in delivering VA CALM content for each treatment session. Practice ratings were discontinued once both raters produced two initial consecutive sets of ratings with a percentage agreement of 80%.

To measure fidelity to VA CALM, we obtained audiotaped treatment sessions from study providers. A total of 379 audiotaped (77.5% of 489 total) treatment sessions were obtained. Eleven audiotaped treatment sessions were deemed to be of too poor quality to rate leaving a total of 368 audiotaped sessions to be considered for rating. First and last sessions were not collected for rating due to those sessions primarily providing orientation to VA, the treatment process, and treatment discontinuation. Following approaches used in prior studies (Diebold et al., 2020; Stirman et al., 2018; Craske et al., 2011), we used stratified sampling to randomly select 19% (68/368; 32 and 36 audiotaped treatment sessions in the manual and computer conditions, respectively) of the audiotaped VA CALM sessions across sessions, and randomly assigned audiotapes to two coders to conduct the fidelity rating.

2.6.2.   General mental health symptoms—Two secondary outcomes were the Brief Symptom Inventory (BSI-18) Global Severity Index (GSI) (Recklitis et al., 2017) and the Short-Form (SF-12) mental health composite (Ware et al., 1996). The BSI-18 is designed to measure general psychological distress and consists of three six-item subscales (somatization, anxiety, and depression). Respondents are asked to indicate, using a 5-point Likert scale, how much they have been bothered by each symptom over the past week. All 18-items can be summed to derive the GSI score which is an indicator of overall level of psychological distress.

The SF-12 was designed to measure physical and mental health of persons in the United States (Ware et al., 1996), with all items weighted and summed to construct summary scores representing both aspects of overall health. Only the mental health composite was used in the present study.

2.6.3.   Disorder-specific symptoms—Three secondary, disorder-specific outcomes were also specified (Cucciare et al., 2016). The Generalized Anxiety Disorder-7 (GAD-7) is a reliable and valid assessment of GAD symptoms (Spitzer et al., 2006). Participants indicate how bothered (0=not at all, 3=nearly every day) they have been by anxiety symptoms (e.g., trouble relaxing) over the past two weeks. The Posttraumatic Stress Disorder Checklist for the Diagnostic and Statistical Manual for Mental Disorders (Fifth Edition) (PCL-5; Weathers et al., 2013) was used to assess symptoms of PTSD. Participants are asked to indicate how much they have been bothered (0 = not at all, 4 = extremely) by each item (e.g., feeling distant or cutoff from other people). The Patient Health Questionnaire (PHQ-9) was used to measure the severity of depression symptoms. The PHQ-9 is a valid measure for assessing how often participants have been bothered (0 = not at all, 3 = nearly every day), in the past two weeks, by depression symptoms (e.g., feeling tired or having little energy) (Kroenke et al., 2001).

There were too few participants with SAD and PD to conduct a sub analysis of those disorders (Table 1). Thus, disorder-specific outcomes were examined for participants with GAD, PTSD, and depression.

2.7. Independent variable and potential covariates

The independent variable was a binary variable indicating the study conditions (1 = computer and 0 = manual). Potential covariates included demographics for patient participants consisting of age, gender, race/ethnicity (Hispanic, Black, White and other), marital status (married or living as married, separated/divorced/widowed, and single), education (below college, some college, and college graduate or above), employment status (employed full/part time, unemployed, disabled, and other), individual income (under $10,000, $10,000–25,000, $25,001–50,000, and $50,001 or above) and number/type of health insurance. Patient clinical characteristics (diagnosis, number of treatment sessions attended) were also obtained.

2.8. Statistical analysis

For the primary outcome of providers' CBT fidelity, descriptive statistics were calculated for the entire sample and by session. The association between condition and the outcome of fidelity was examined using a general linear mixed model to account for Veteran participants clustered within providers. The treatment session and strata variables were also included in the model.

For the patient-level secondary outcomes, bivariate analysis was performed using generalized estimating equations due to the clustered structure of Veterans within the same providers. Associations between condition and potential covariates and between outcomes and potential covariates were examined. Covariates with p-values less than 0.10 were included in the multivariate models for associations between condition and outcomes over time. Generalized linear mixed models were used to account for the correlations for patients within providers as well as the correlations of multiple assessments within patients. Gamma distribution was specified for BSI-18 GSI scores after a small rescale for zero value due to its violation of normality and normal distribution was specified for the remaining outcomes as they were approximately normally distributed. All the models included the condition indicator variable, time (for the three interviews), strata, and covariates identified in the bivariate analysis. The covariates associated with condition (gender and primary diagnosis) were included in all the models with the exception of primary diagnosis not being included in the subgroup specific diagnosis group analysis. The covariates associated with the outcomes were also included in the corresponding outcome models. The interaction between condition and time was included in all of the models as hypothesized. General linear mixed models were also fit for disorder specific outcomes for subgroups of Veterans with the corresponding specific disorders as they were approximately normally distributed.

The LS mean differences (or ratios depending on the outcomes) between the two conditions and their corresponding 95% confidence intervals were calculated for evaluating the effect of condition. Similar differences (or ratios) between each follow-up and baseline by each condition and their corresponding 97.5% confidence intervals were also calculated for evaluating the effect of time. A narrower confidence interval (equivalent to a p-value of 0.025) was used to adjust for multiple comparisons. All the analyses were performed using SAS 9.4.

3. Results

3.1. Participant characteristics

Table 1 shows that most demographics characteristics were similar between patient participants in the computer and manual condition. Gender and principal diagnoses showed large differences between the two conditions ($p < .10$) and were included in all models as covariates except that principal diagnosis was not included in the subgroup analysis. The sample was mostly male (73%), with lower rates of males in the computer condition (70%) when compared to the manual condition (80%). Participants were primarily treated for depression (34%), PTSD (33%), or GAD (26%). The sample was mostly White (50%) and Hispanic (27%). The mean age of participants in the study was 44 ($\underline{SD}$=13).

3.2. Treatment participation

Prior to initiating treatment, in both conditions providers educated patients on the components of VA CALM and encouraged them to complete the initial eight sessions over eight to twelve weeks, but flexibility was permitted. The average number of VA CALM treatment sessions was 3.3 ($\underline{SD}$ = 2.8) for the computer condition and 4.1 ($\underline{SD}$ = 3.2) for manual condition which was not statistically different based on a generalized linear mixed model adjusting for strata ($\underline{p}$= .57; data not shown in tables).

3.3. Primary outcome

Table 2 presents descriptive statistics for fidelity by sessions for the two conditions. The general linear mixed models did not show a statistically significant effect of condition on overall fidelity ($p = .1701$).

3.4. Secondary outcomes

3.4.1. General mental health symptoms—We did not observe statistically significant effects of condition on either the BSI-18 GSI or SF-12 mental health composite scores. However, BSI-18 GSI scores reduced significantly over time for participants in both the computer and manual conditions at six month follow-up compared to baseline (Table 3). For BSI-18 GSI scores, effect sizes for time between baseline and six month follow-up were 0.39 and 0.56 for the computer and manual conditions, respectively.

3.4.2. Disorder-specific symptoms—For participants with a principal diagnosis of PTSD, we did not observe statistically significant effects of condition or time on PCL-5 scores (Table 4). For participants with a principal diagnosis of GAD, GAD-7 scores differed significantly between the two conditions at both follow-ups, with participants in the computer condition reporting significantly lower GAD-7 scores than participants in the manual condition at three and six month follow-up. The effect sizes for condition on GAD-7 scores at three month and six month follow-up were 0.37 and 0.55, respectively. GAD-7 scores also decreased significantly over time for participants in the computer condition, but not the manual condition, at both three and six month follow-up. The effects sizes for time between baseline and three month and baseline and six month follow-up were 0.55 and 0.45 for the computer condition, respectively. For participants with a principal diagnosis of depression, we did not observe statistically significant effects of condition on PHQ-9

scores. We did find that PHQ-9 scores decreased significantly over time for participants in the manual condition at six month follow-up and in the computer condition at three and six month follow-up compared to baseline. Effect sizes for time between baseline and three month and between baseline and six month follow-up for the computer condition were 0.51 and 0.49, respectively. The effect size for time between baseline and six month follow-up for the manual condition was 0.71.

4. Discussion

This study examined the effect of an adapted version of CALM (VA CALM), developed specifically for US military Veterans, on VA provider fidelity to CBT and patient outcomes. Our findings suggest that, contrary to our hypothesis, the computer version of VA CALM did not show a statistically significant improvement in our primary outcome (CALM fidelity) relative to the manual. However, we observed a large effect size (0.88) for the comparison between the computer and manual conditions on CBT fidelity, indicating that using the computer version of VA CALM may have provided some benefit relative to the manual in helping providers adhere to the CBT model. This finding is consistent with those of a prior study showing that the computer version of CALM helped primary care providers, with little to no prior experience with CBT, adhere to and competently deliver CBT to persons with anxiety disorders (Craske et al., 2011).

Average CBT fidelity ratings of providers in the computer condition (mean = 3.99) were similar to those observed in a prior study of CBT naïve providers learning and delivering CALM in a primary care setting (Craske et al., 2011; Roy-Byrne et al., 2010). This finding suggests that providers in the computer condition may have been more successful in delivering VA CALM than providers in the manual condition (average score = 3.22), although the differences between conditions on this outcome was not statistically significant. It is important to note that our qualitative findings show that providers were comfortable using both the manual and computer versions of VA CALM (Ecker et al., 2020). And, providers reported that both modalities helped them treat mental health conditions that they were not prepared to treat (Ecker et al., 2020).

In terms of secondary outcomes, we found no statistically significant differences between the two groups on general mental health symptoms (SF-12 MCS), and symptoms of depression or PTSD. Having an active usual care condition (i.e., manualized VA CALM) likely contributed to findings showing no differences between conditions on these outcomes, and for BSI-18 GSI scores and depression symptoms, improvements in both conditions at six month follow-up. For participants with PTSD, no improvements in symptoms were observed over time or between groups. It is generally assumed that PTSD is among the more difficult diagnoses to treat for inexperienced therapists. About half of our providers had little to no training in CBT. To the extent this variable serves as a proxy for "experience", it is possible that our training procedures were not sufficient to support providers in treating patients with this diagnosis.

In contrast, the computer condition was statistically superior to the manual condition at three and six month follow-ups for GAD. The effect sizes for the differences between conditions

on GAD symptoms were small to medium. A prior study comparing CALM to standard care for treating civilians in primary care showed similar effect sizes for a computer version of CALM on GAD symptoms, relative to standard care, at six month follow-up (Craske et al., 2011). Our findings suggest that the computer version of CALM, compared to the manual, may improve providers' ability to treat GAD.

Reasons for improved GAD outcomes in the VA CALM computer, compared to manual, condition remain unclear. It is unlikely that this finding is due to the small but not statistically significant difference in the fidelity outcome observed between conditions. Qualitative interviews with providers who participated in the VA CALM study may help shed some light on this finding. Providers reported that perceived benefits of the computer version of VA CALM were that it: (a) helped providers deliver CBT content in a flexible and individualized manner; (b) had a structure that was particularly useful for providers with little to no prior training or experience delivering CBT; (c) was helpful in "walking patients through the [CBT] treatment process"; (d) had a "novel" aspect to it, contributing to some patients feeling "excited" about using the computer in their treatment; and (e) seemed to be particularly well received among patients with anxiety disorders. It is possible that the structure of the computer modality offered important benefits (e.g., flexibility to the target disorder, structured help in supporting the provider and patient in delivering and receiving CBT content), not as easily accessed in the manual modality, that contributed to better outcomes for Veterans with GAD.

This study has several strengths and limitations. As a multisite randomized controlled trial, this study has strong internal validity and findings should generalize to other VA CBOCs. Eligibility criteria of any Veteran with a clinical diagnosis of one of the target disorders, as opposed to meeting strict screening or interview criteria, may also enhance the generalizability of findings to Veterans with the targeted diagnoses receiving mental health care at other VA CBOCs. The present study also tested pre-registered outcomes which enhances study transparency and methodological rigor (Lassman et al., 2017). Also, follow-up rates were high at three and six month follow-up which can reduce bias and enhance the validity of the study findings (Norvell et al., 2016).

A limitation of this study is that Veteran patients were not randomized so there remains the possibility that the study conditions were not balanced on important unmeasured factors that may have contributed to the results (Suresh, 2011). In addition, we were unable to evaluate VA CALM's effectiveness for Veterans with other anxiety disorders (e.g., SAD, PD) due to small samples sizes. Also, providers in both conditions received an intensive training program in VA CALM that may have led to an underestimate of the effect of condition (or modality) on the fidelity outcome: that is, the training components (e.g., supervision-consultation, facilitation) may have provided substantial help in delivering CBT thereby reducing the effect of modality (computer or manual) on CBT fidelity. This study was underpowered to test our primary hypothesis. Is it possible that had this study been fully powered, we would have observed statistically significant differences in CBT fidelity between the two conditions. Also, our disorder-specific subgroups contained small sample sizes which may limit the generalizability of our findings. Limitations also include providers being asked to identify participants which may have introduced bias in the sample, providers

being aware that their audiotaped sessions would be evaluated which may have impacted their in-session behavior, and the possibility of fidelity raters, originally blinded to study condition, becoming aware of study condition when listening to audiotaped therapy sessions.

Although we observed differences between the two conditions (favoring the computer condition) on VA CALM fidelity these differences were not statistically significant. Larger studies are needed to determine whether the differences observed on fidelity outcomes replicate and are statistically significant. Further, both conditions were associated with better patient outcomes and the computer version of VA CALM provided modest benefit to Veterans with GAD compared to the manual. Larger studies are needed to fully determine whether the modality of delivering VA CALM impacts provider fidelity to CBT and/or outcomes for patients with a wide variety of mental health diagnoses. Should future studies uphold these findings, providers might be encouraged to choose the modality for delivering VA CALM based on what is most feasible for them to use in their clinical practice and/or is preferable for their patient population as both versions are associated with positive patient outcomes. Further, as the use of telehealth within VA continues to grow, CALM could be adapted for use with video to home technologies which allow providers to deliver mental health treatment content by sharing their computer screen with Veterans (Boykin et al., 2019).

Acknowledgements

References

Abraham TH, Marchant-Miros K, McCarther MB, Craske MG, Curran GM, Kearney LK, Greene C, Lindsay JA, Cucciare MA, 2018. Adapting coordinated anxiety learning and management for veterans affairs community-based outpatient clinics: iterative approach. JMIR Ment. Health5 (3), e10277 [PubMed: 30135051]

Ansseau M, Dierick M, Buntinkx F, Cnockaert P, De Smedt J, Van Den Haute M, Vander Mijnsbrugge D., 2004. High prevalence of mental disorders in primary care. J. Affect. Disorder78 (1), 49–55

Bidassie B, Williams LS, Woodward-Hagg H, Matthias MS, Damush TM, 2015. Key components of external facilitation in an acute stroke quality improvement collaborative in the Veterans Health Administration. Implement. Sci10, 69 [PubMed: 25971405]

Boykin DM, Keegan F, Thompson KE, Voelkel E, Lindsay JA, Fletcher TL, 2019. Video to home delivery of evidence-based psychotherapy to Veterans with posttraumatic stress disorder. Front. Psychiatry6 (10), 1–8

Cape J, Whittington C, Buszewicz M, Wallace P, Underwood L., 2010. Brief psychological therapies for anxiety and depression in primary care: meta-analysis and meta-regression. BMC Med. 8, 38 [PubMed: 20579335]

Craske MG, Rose RD, Lang A, Welch SS, Campbell-Sills L, Sullivan G, Sherbourne C, Bystritsky A, Stein MB, Roy-Byrne PP, 2009. Computer-assisted delivery of cognitive behavioral therapy for anxiety disorders in primary-care settings. Depress. Anxiety26 (3), 235–242 [PubMed: 19212970]

Craske MG, Stein MB, Sullivan G, Sherbourne C, Bystritsky A, Rose RD, Lang AJ, Welch S, Campbell-Sills L, Golinelli D, Roy-Byrne P., 2011. Disorder-specific impact of coordinated anxiety leaning and management treatment for anxiety disorders in primary care. Arch. Gen. Psychiarty68 (4), 378–388

Cucciare MA, Curran GM, Craske MG, Abraham T, McCarthur MB, Marchant-Miros K, Lindsey JA, Kauth MR, Landes SJ, Sullivan G., 2016. Assessing fidelity of cognitive behavioral therapy in rural VA clinics: design of a randomized implementation effectiveness (hybrid type III) trial. Implement. Sci11, 65 [PubMed: 27164866]

Cucciare MA, Marchant K, Lindsay J, Craske MG, Ecker A, Day S, Hogan J, Henn J, LeBeau RT, Rabalais A, Rose RD, Qualls M, Treanor M, Abraham TH, 2019. An evidenced-based model for disseminating-implementing coordinated anxiety learning and management in department of veterans affairs' community-based outpatient clinics. J. Rural Health36 (3), 371–380 [PubMed: 31508861]

Cully JA, Jameson JP, Phillips LL, Kunik ME, Fortney JC, 2010. Use of psychotherapy by rural and urban Veterans. J. Rural Health26 (3), 225–233 [PubMed: 20633090]

Diebold A, Ciolino JD, Johnson JK, Yeh C, Gollan JK, Tandon SD, 2020. Comparing fidelity outcomes of paraprofessional and professional delivery of a perinatal depression preventive intervention. Adm. Policy. Ment. Health47 (4), 597–605 [PubMed: 32086657]

Drapkin ML, Wilbourne P, Manuel JK, Baer J, Karlin B, Raffa S., 2016. National dissemination of motivation enhancement therapy in the Veterans Health Administration: training program design and initial outcomes. J. Subst. Abuse Treat. 65, 83–87 [PubMed: 26951921]

Ecker AH, Abraham TH, Martin LA, Marchant-Miros K, Cucciare MA, 2020. Factors affecting adoption of Coordinated anxiety learning and management (CALM) in veterans' affairs community-based outpatient clinics. J. Rural Health10.1111/jrh.12528.

Hepner KA, Azocar F, Greenwood GL, Miranda J, Burnam MA, 2010. Development of a clinician report measure to assess psychotherapy for depression in usual care settings. Adm. Policy. Ment. Health37, 221–229 [PubMed: 20186569]

Hoifodt RS, Strom C, Kolstrup N, Eisemann M, Waterloo K., 2011. Effectiveness of cognitive behavioural therapy in primary health care: a review. Fam. Pract28 (5), 489–504 [PubMed: 21555339]

Karlin BE, Brown GK, Trockel M, Cunning D, Zeiss AM, Taylor CB, 2012. National dissemination of cognitive behavioral therapy for depression in the Department of Veterans Affairs Health Care System: therapist and patient-level outcomes. J. Consult. Clin. Psychol80 (5), 707–718 [PubMed: 22823859]

Karlin BE, Ruzek JI, Chard KM, Eftekhari A, Monson CM, Hembree EA, Resick PA, Foa EB, 2010. Dissemination of evidence-based psychological treatments for posttraumatic stress disorder in the Veterans Health Administration. J. Trauma Stress23 (6), 663–673 [PubMed: 21171126]

Karlin BE, Trockel M, Taylor CB, Gimeno J, Manber R., 2013. National dissemination of cognitive behavioral therapy for insomnia in veterans: therapist- and patient-level outcomes. J. Consult. Clin. Psychol81 (5), 912–917 [PubMed: 23586730]

Kroenke K, Spitzer RL, Williams JB, 2001. The PHQ-9: validity of a brief depression severity measure. J. Gen. Intern. Med16 (9), 606–613 [PubMed: 11556941]

Kroenke K, Spitzer RL, Williams JB, Monahan PO, Lowe B., 2007. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann. Intern. Med146 (5), 317–325 [PubMed: 17339617]

Lassman SM, Shopshear OM, Jazic I, Ulrich J, Francer J., 2017. Clinical trial transparency: a reassessment of industry compliance with clinical trial registration and reporting requirements in the United States. BMJ Open7 (9), e015110. https://doi.org/10.1136/bmjopen-2016-015110https://doi.org/https://doi.org/.

Leon AC, Olfson M, Broadhead WE, Barrett JE, Blacklow RS, Keller MB, Higgins ES, Weissman MM, 1995. Prevalence of mental disorders in primary care. Implications for screening. Arch. Fam. Med4 (10), 857–861 [PubMed: 7551133]

Norvell DC, Dettori JR, Chapman JR, 2016. Enhancing clinical study retention rates to avoid follow-up bias: how do we keep our study participants from the "The Land of the Lost?Glob. Spine J6, 519–521

Roy-Byrne P, Craske MG, Sullivan G, Rose RD, Edlund MJ, Lang AJ, Bystritsky A, Welch SS, Chavira DA, Golinelli D, Campbell-Sills L, Sherbourne CD, Stein MB, 2010. Delivery of evidence-based treatment for multiple anxiety disorders in primary care: a randomized controlled trial. JAMA303 (19), 1921–1928 [PubMed: 20483968]

Recklitis CJ, Blackmon JE, Chang G., 2017. Validity of the Brief Symptom Inventory-18 (BSI-18) for identifying depression and anxiety in young adult cancer survivors: comparison with a structured clinical diagnostic interview. Psychol. Assess29 (10), 1189–1200 [PubMed: 28080106]

Spitzer RL, Kroenke K, Williams JB, Lowe B., 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch. Intern. Med166 (10), 1092–1097 [PubMed: 16717171]

Stetler CB, Legro MW, Wallace CM, Bowman C, Guihan M, Hagedorn H, Kimmel B, Sharp ND, Smith JL, 2006. The role of formative evaluation in implementation research and the QUERI experience. J. Gen. Intern. Med21 (Suppl 2), S1–S8

Stewart MO, Raffa SD, Steele JL, Miller SA, Clougherty KF, Hinrichsen GA, Karlin BE, 2014. National dissemination of interpersonal psychotherapy for depression in veterans: therapist and patient-level outcomes. J. Consult. Clin. Psychol82 (6), 1201–1206 [PubMed: 25045906]

Sullivan G, Craske MG, Sherbourne C, Edlund MJ, Rose RD, Golinelli D, Chavira DA, Bystritsky A, Stein MB, Roy-Byrne PP, 2007. Design of the Coordinated Anxiety Learning and Management (CALM) study: innovations in collaborative care for anxiety disorders. Gen. Hosp. Psychiatry29 (5), 379–387 [PubMed: 17888803]

Suresh K., 2011. An overview of randomization techniques: an unbiased assessment of outcome in clinical research. J. Hum. Reprod. Sci4 (1), 8–11 [PubMed: 21772732]

U.S. Department of Veterans Affairs, 2008. Veterans Health Administration: VHA Handbook 1160.01: Uniform Mental Health Services in VA Medical Centers and Clinics. Washington, DC.

Veterans Health Administration Office of Rural Health (ORF). Caring for rural veterans [internet]. ORF; 2018. https://www.ruralhealth.va.gov/docs/ORH_Brochure_2018_v3_1_print_508.pdf. AccessedSeptember 6, 2019.

Ware J Jr., Kosinski M, Keller SD, 1996. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med. Care34 (3), 220–233 [PubMed: 8628042]

Waxmonsky J, Kilbourne AM, Goodrich DE, Nord KM, Lai Z, Laird C, Clogston J, Kim HM, Miller C, Bauer MS, 2014. Enhanced fidelity to treatment for bipolar disorder: results from a randomized controlled implementation trial. Psychiatr Serv. (Washington, DC)65, 81–90

Weathers FW, Litz BT, Keane TM, Palmieri PA, Marx BP, Schnurr PP (2013). The PTSD Checklist for DSM-5 (PCL-5). Scale available from the National Center for PTSD at www.ptsd.va.gov.

Stirman SW, Marques L, Creed TA, Gutner CA, DeRubeis R, Barnett PG, Kuhn E, Suvak M, Owen J, Vogt D, Jo B, Schoenwald S, Johnson C, Mallard K, Beristianos M, La Bash H., 2018. Leveraging routine clinical materials and mobile technology to assess CBT fidelity: the Innovative Methods to Assess Psychotherapy Practices (imAPP) study. Implement. Sci13 (1)
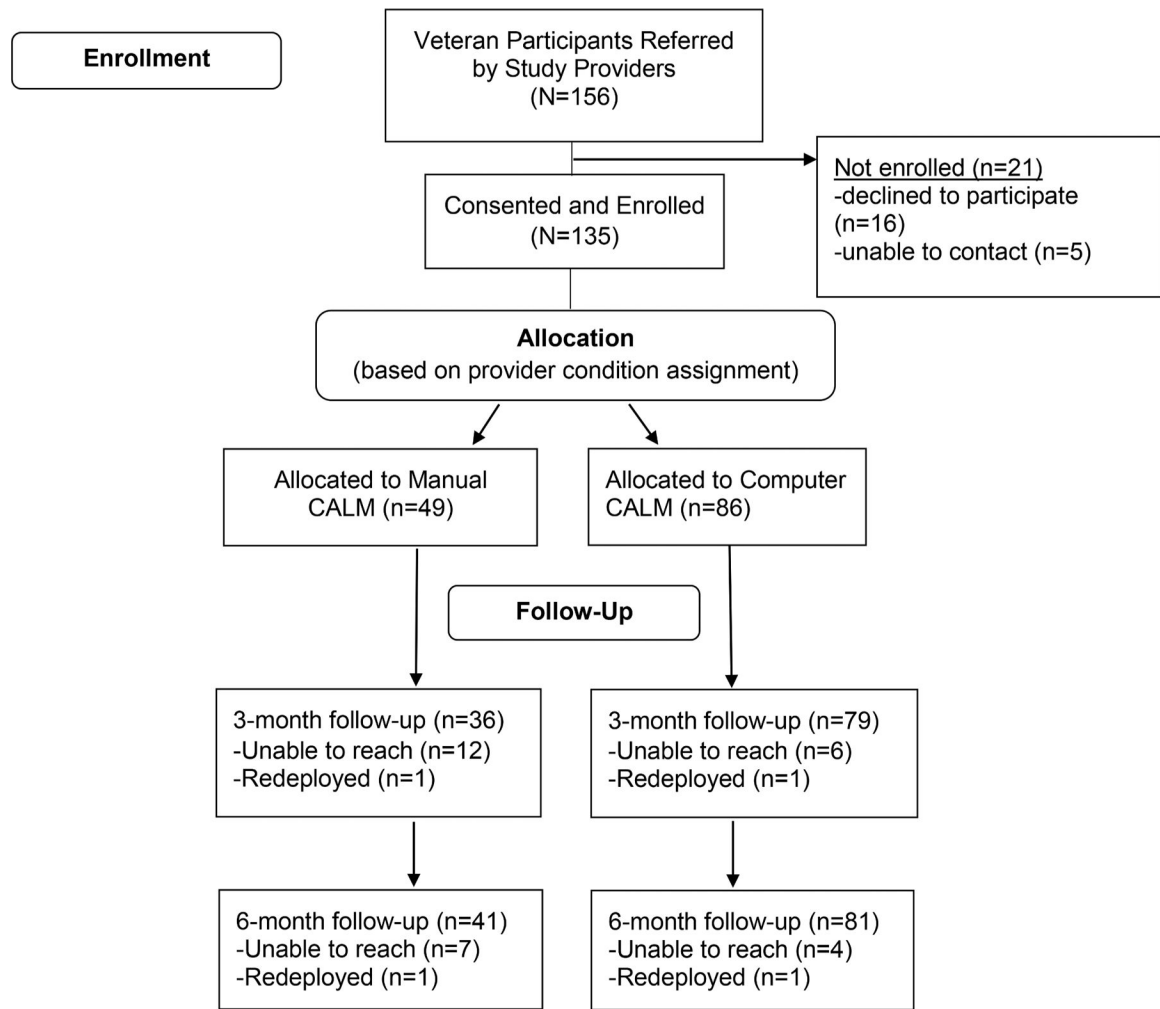
Fig. 1.
CONSORT flow diagram of Veteran participants.

Table 1

Descriptive statistics by condition – CALM-delivered via computer or manual N (%).

Variable		Manual *n* = 49 (36.3%)	Computer *n* = 86 (63.7%)	Total *N* = 135	p value
Strata	None or informal training	30 (61.22)	36 (41.86)	66 (48.89)	0.5635
	Formal training	19 (38.78)	50 (58.14)	69 (51.11)	
Gender	Male	39 (79.59)	60 (69.77)	99 (73.33)	0.0122
	Female	10 (20.41)	26 (30.23)	36 (26.67)	
Race	Hispanic	21 (42.86)	15 (17.44)	36 (26.67)	0.1862
	Other	6 (12.24)	6 (6.98)	12 (8.89)	
	Black	5 (10.20)	14 (16.28)	19 (14.07)	
	White	17 (34.69)	51 (59.30)	68 (50.37)	
Marital status	Married or living as married	22 (44.90)	40 (46.51)	62 (45.93)	0.9136
	Separated/divorced/Widowed	17 (34.69)	27 (31.40)	44 (32.59)	
	Single	10 (20.41)	19 (22.09)	29 (21.48)	
Education	Below college	11 (22.45)	14 (16.28)	25 (18.52)	0.4018
	Some college	21 (42.86)	44 (51.16)	65 (48.15)	
	College graduate or above	17 (34.69)	28 (32.56)	45 (33.33)	
Employment status	Employed full/part time	16 (32.65)	34 (39.53)	50 (37.04)	0.7913
	Unemployed	8 (16.33)	17 (19.77)	25 (18.52)	
	Disabled	13 (26.53)	20 (23.26)	33 (24.44)	
	Other	12 (24.49)	15 (17.44)	27 (20.00)	
Individual income	Under $10,000	9 (18.75)	11 (13.10)	20 (15.15)	0.6662
	$10,000 – 25,000	10 (20.83)	24 (28.57)	34 (25.76)	
	$25,001 – 50,000	16 (33.33)	29 (34.52)	45 (34.09)	
	$50,001 or above	13 (27.08)	20 (23.81)	33 (25.00)	
Principal diagnosis	Generalized Anxiety Disorder (GAD)	7 (14.29)	28 (32.56)	35 (25.93)	0.0586
	Social Anxiety Disorder (SAD)	1 (2.04)	2 (2.33)	3 (2.22)	
	Depression	18 (36.73)	28 (32.56)	46 (34.07)	
	Panic Disorder (PD)	4 (8.16)	3 (3.49)	7 (5.19)	
	Posttraumatic Stress Disorder (PTSD)	19 (38.78)	25 (29.07)	44 (32.59)	
Age (mean (SD))		44.63 (14.56)	43.20 (12.35)	43.72 (13.16)	0.5347
Insurance sum		1.41 (0.79)	1.30 (0.67)	1.34 (0.71)	0.5418
Outcomes					
BSI-18 GSI (mean, SD)	Baseline	26.65 (14.09)	23.79 (14.42)	24.83 (14.31)	0.3524
	3 month (*n* = 114)	27.78 (13.83)	22.70 (14.71)	24.32 (14.57)	0.0834
	6 month (*n* = 120)	22.10 (15.53)	19.30 (12.62)	20.23 (13.66)	0.5784
SF-12 MCS (mean, SD)	Baseline	34.20 (6.84)	33.80 (6.33)	33.95 (6.50)	0.6959
	3 month (*n* = 114)	35.25 (5.37)	35.14 (6.08)	35.17 (5.84)	0.9397
	6 month (*n* = 120)	35.73 (6.94)	35.53 (6.28)	35.60 (6.48)	0.9369
PCL-5 (mean, SD) (PTSD only)	Baseline	51.63 (9.59)	46.28 (17.37)	48.59 (14.63)	0.1519
	3 month (*n* = 114)	47.53 (10.45)	44.68 (17.35)	45.84 (14.84)	0.7306

Variable		Manual *n* = 49 (36.3%)	Computer *n* = 86 (63.7%)	Total *N* = 135	p value
	6 month (*n* = 120)	46.00 (17.34)	42.57 (17.09)	44.05 (17.04)	0.6552
GAD-7 (mean, SD) (GAD only)	Baseline	13.43 (3.99)	12.75 (5.57)	12.89 (5.25)	0.7022
	3 month (*n* = 114)	13.00 (6.82)	9.83 (5.55)	10.38 (5.78)	0.2047
	6 month (*n* = 120)	12.00 (5.51)	9.88 (5.81)	10.33 (5.73)	0.4045
PHQ-9 (mean, SD) (Depression only)	Baseline	16.67 (5.76)	16.39 (6.76)	16.50 (6.32)	0.7652
	3 month (*n* = 114)	15.87 (5.33)	13.62 (6.43)	14.44 (6.08)	0.2714
	6 month (*n* = 120)	12.33 (5.85)	13.32 (6.99)	12.98 (6.56)	0.7167

Table 2

Descriptive statistics for CALM fidelity competence ratings by each treatment session and overall (scores can range from 0 to 6; higher scores = more competence) Means (SD).

Session #	Manual	Computer
2	3.06 (1.04)	4.04 (0.71)
3	3.27 (0.58)	4.22 (0.69)
4	2.89 (1.19)	3.87 (1.21)
5	3.81 (0.55)	3.96 (0.98)
6	3.38 (0.18)	3.92 (0.17)
7	3.50 (0.71)	4.00 (0.82)
Overall	3.23 (0.88)	3.99 (0.86)

Table 3

LS-mean differences (and 95% confidence intervals) from generalized linear mixed models for associations between outcomes and study condition (computer or manual) with interactions between study condition and time included in the models. Mean difference (95% CI).

Outcomes	Computer-Manual baseline	three month	six month	Follow-ups-Baseline[b] three month-baseline computer	manual	six month-baseline computer	manual
BSI-18 GSI[a]	0.90 (0.66,1.23)	0.84 (0.60,1.16)	1.00 (0.73,1.38)	0.91 (0.78,1.05)	0.97 (0.79,1.20)	0.78 * (0.68,0.90)	0.71 * (0.58,0.87)
SF-12 Mental health composite	0.14 (−2.97,2.52)	0.15 (−2.47, 2.77)	−0.28 (−2.86, 2.29)	1.30 (−0.45, 3.04)	1.29 (−1.21,3.78)	1.70 (−0.06,3.46)	2.12 (−0.33,4.57)

Covariates included variables associated with the study condition or outcomes at $p < .10$. The variables of strata (by design), gender and primary diagnosis were included in all models based on the associations between study condition and covariates. The interaction between study condition and time was also included as hypothesized. Covariates were added based on their associations with each outcomes. For the model for GSI, employment was also included. For MCS, race and marital status were also included

[a]For GSI, the values are ratios of LS-means for the computer condition divided by the manual condition. This outcome is distributed with gamma distribution with log link

[b]For the time effect, the 97.5% CI was presented to control for multiple comparisons between three month and baseline and between six month and baseline.

*
$p < .05$ for comparison between computer and manual conditions and $p < .025$ for comparisons between each follow-up and the baseline.

Table 4

LS-mean differences (and 95% confidence intervals) from generalized linear mixed models for associations between study condition and outcomes (for specific principal diagnoses) with the interaction between condition and time included. Mean difference (95% CI).

Outcomes	Computer-Manual baseline	three month	six month	Follow-ups – Baseline[a] three month-baseline computer	manual	six month-baseline computer	manual
PCL-5 (PTSD condition)	-6.34 (-16.86, 4.19)	-1.44 (-12.58, 9.70)	-3.68 (-14.78, 7.42)	-1.03 (-7.92, 5.86)	-5.93 (-14.18, 2.33)	-3.84 (-10.85, 3.17)	-6.50 (-14.56, 1.56)
GAD-7 (GADcondition)	-3.87 (-8.89, 1.15)	-5.88*(-11.37, -0.39)	-5.25*(-10.29, -0.22)	-2.66*(-5.09,-0.24)	-0.65 (-5.82, 4.52)	-2.82*(-5.18, -0.45)	-1.43 (-6.03, 3.17)
PHQ-9 (Depression condition)	0.37 (-3.41,4.14)	-2.96 (-7.40, 1.48)	-0.038 (-4.77, 4.02)	-2.50*(-4.97, -0.03)	-0.87 (-4.02, 2.29)	-3.11*(-5.51, -0.71)	-4.06*(-7.22, -0.90)

Covariates included variables associated with the study condition or outcomes at $p < .1$. The variables of strata (by design) and gender were included in all models based on the association between the study condition and covariates. The interaction between study condition and time was also included as hypothesized. Covariates were added based on their associations with each outcomes; For the model for PTSD and GAD-7, education and employment status were also included; For the model for PHQ-9, income and employment status were also included.

[a] For the time effect, the 97.5% CI was presented to control for multiple comparisons between three month and baseline and between six month and baseline.

*
$p < .05$ for comparison between computer and manual conditions and $p < .025$ for comparisons between each follow-up and the baseline.