



Published in final edited form as:

J Consult Clin Psychol. 2023 November ; 91(11): 665–679. doi:10.1037/ccp0000832.

The Role of Therapy Delivery and Clinic Organizational Factors in Explaining Therapist Effects for Trauma-Focused Psychotherapies in the Veterans Health Administration

Nina A. Sayer^{1,2}, Shannon Wiltsey Stirman^{3,4}, Craig S. Rosen^{3,4}, Shannon Kehle-Forbes^{1,5}, Michele R. Spont^{1,2}, Afsoon Eftekhari³, Kathleen M. Chard^{6,7}, Adam Kaplan¹, David B. Nelson^{1,5}

¹Center for Care Delivery and Outcomes Research, Minneapolis VA Health Care System

²Departments of Medicine and Psychiatry, University of Minnesota

³Dissemination & Training Division, National Center for PTSD, VA Palo Alto Health Care System

⁴Department of Psychiatry & Behavioral Sciences, Stanford University School of Medicine

⁵Department of Medicine, University of Minnesota

⁶PTSD and Anxiety Disorder Division, Cincinnati Veterans Affairs (VA) Medical Center

⁷University of Cincinnati College of Medicine

Abstract

Objective: This study estimated the size of therapist effects (TEs) for dropout and clinical effectiveness of two trauma-focused psychotherapies (TFPs) and evaluated whether therapy delivery and clinic organizational factors explained observed TEs.

Method: Participants were 180 therapists (54.4% psychologists, 42.2% social workers) from 137 Veterans Health Administration facilities and 1,735 patients (24.7% women; 27.2% people of color) who completed at least two TFP sessions. Outcomes were dropout (< 8 TFP sessions) and, for a subsample ($n = 1,273$) clinically meaningful improvement and recovery based on PTSD Checklist-5 scores. Therapist-level predictors were ascertained through survey, manual chart review and administrative data. Multilevel models estimated TEs.

Results: Over half (51.2%) of patients dropped out and those who dropped out were less likely to meet criteria for clinically meaningful improvement or recovery ($ps < 0.001$). Adjusting for case-mix and TFP type, therapists accounted for 5.812% ($p < 0.001$) of the unexplained variance in dropout. The average dropout rate for the 45 therapists in the top performing quartile was 27.0% while the average dropout rate for the 45 therapists in the bottom performing quartile was 78.8%. Variation between therapists was reduced to 2.031% ($p = .140$) when therapists' mean of days between sessions, adherence, implementation climate and caseload were added to multilevel models. TEs were non-significant for clinically meaningful improvement and recovery.

Correspondence concerning this article should be addressed to Nina A. Sayer, Ph.D., Center for Care Delivery and Outcomes Research, Minneapolis VA Health Care System, One Veterans Drive, Minneapolis, MN 55417, United States. nina.sayer@va.gov

The authors report no conflict of interest.

Conclusions: Interventions targeting therapy delivery and clinic organization have the potential to reduce variation between therapists in TFP dropout so that more patients stay engaged long enough to experience clinical benefit.

Keywords

therapist effects; trauma-focused psychotherapies; dropout; treatment outcomes; implementation

The inclusion of Posttraumatic Stress Disorder (PTSD) as a distinct mental health diagnosis in the Diagnostic and Statistical Manual for Mental Disorders in 1980 catalyzed efforts to identify PTSD-specific therapies. Since that time, research has shown that trauma-focused psychotherapies (TFPs) confer the greatest benefit compared with other psychotherapeutic modalities. In 2005, the US Veterans Health Administration (VHA) began rollout of two TFPs – Cognitive Processing Therapy (CPT) and Prolonged Exposure (PE), both of which are recommended as first-line treatments in all Clinical Practice Guidelines for PTSD (Hamblen et al., 2019). This rollout included policy requiring the availability of CPT and PE at every medical center, competency-based training of its mental health workforce, the development of templates for documenting CPT and PE delivery in the electronic medical record, and ongoing support for therapists through mentoring and consultation (Karlin et al., 2010).

Although CPT and PE are among the TFPs for PTSD with the strongest evidence base, not all patients experience a therapeutic benefit. This may be particularly true among military populations (Kitchiner et al., 2019; Steenkamp et al., 2015). A recent comparative effectiveness study in a large veteran sample found that 27% – 40% failed to respond to these TFPs (Schnurr et al., 2022). Furthermore, therapy non-completion, a problem across therapies (Cooper et al., 2018; Fernandez et al., 2015), is common for CPT and PE (Steenkamp et al., 2020). An understanding of the reasons that patients dropout or fail to benefit even when they receive an adequate dose can inform treatment decisions and guide the design of strategies to reduce dropout and improve clinical outcomes from TFPs.

Most of the research examining the variability in dropout and effectiveness from CPT and PE focused on patient characteristics and yielded inconsistent results (Eftekhari et al., 2020; Hale et al., 2019; Kehle-Forbes et al., 2016; Maguen et al., 2019, 2020; Roberge et al., 2022). The most notable exception is that younger age or having served in the more recent wars consistently predicted increased dropout from CPT and PE (Goetter et al., 2015). Much less attention has been paid to the therapists providing TFPs, variation in how they deliver these treatments, and whether this variation affects retention and treatment effectiveness. As the largest integrated healthcare system in the US, variability between VHA therapists and the clinics in which they work is inevitable.

Systematic variation between therapists in patient outcomes including retention and clinical improvement is referred to as *therapist effects* (TEs). When TEs are present, certain therapists consistently achieve better results than others (Baldwin & Imel, 2013). Two practice-based studies have examined TEs in the context of TFPs. One study in a single VHA PTSD clinic involved 25 therapists and 192 patients who completed CPT and found that therapists accounted for approximately 12% of the variability in post-treatment PTSD

symptom levels (Laska et al., 2013). The second study used VHA administrative data to examine 2,709 therapists who provided CPT or PE to 18,461 patients and reported that therapists accounted for almost 9% of the unexplained variance in therapy completion (Sayer et al., 2022). The first study did not look at dropout and the second study did not examine clinical outcomes. Neither study controlled for potentially confounding case-mix variables (e.g., patient characteristics associated with outcomes and that may be unevenly distributed across therapists).

While quantifying the extent of TEs is an essential first step, the more actionable research questions center on the reasons some therapists are more effective than others. One might expect therapists who provide CPT or PE as tested in RCTs and prescribed in treatment manuals to have better patient outcomes than those who do not. In general, the literature on the relationship between fidelity and therapy outcomes has yielded mixed results (Keefe et al., 2022; Webb et al., 2010). The small number of studies that have examined this issue in the context of CPT reported that certain components of treatment fidelity (e.g., adherence to the protocol or competence in providing key components of the protocol) are associated with clinical improvement (Farmer et al., 2017; Holder et al., 2018; Keefe et al., 2022; Marques et al., 2019). The CPT and PE treatment manuals used in VHA also specify that CPT is to be delivered in 60-minute and PE in 90-minute weekly sessions and efficacy trials for these TFPs often used a twice per week session structure. However, in routine care, therapists may have difficulty seeing patients this frequently or consistently. Consistent with the possibility that session spacing matters, a study of women who participated in a RCT comparing CPT to PE found that higher average days between sessions and, to a lesser extent, inconstancy in days between sessions were associated with less symptom reduction over treatment (Gutner et al., 2016). Additionally, converging evidence indicates that TFP dropout is lower when sessions are prescribed at least twice weekly compared with less frequently (Galovski et al., 2022; Levinson et al., 2022). We are not aware of studies that have examined the association between session spacing or fidelity and TFP dropout in the context of routine outpatient settings where there is less treatment monitoring than in RCTs.

The settings in which therapists work vary in terms of climate, culture, policies, resources and procedures. Despite this variability, the clinic and organizational factors that affect therapists' performance are understudied. For example, policy regarding therapist caseload may be driven more by the need to ensure patient access than from evidence on the association between caseload and therapists' effectiveness. TFPs are time intensive (eight to 15 weekly sessions) and therapists report lack of time as a barrier to using TFPs (Finley et al., 2015). It would be useful to determine whether caseload also affects therapists' ability to effectively deliver TFPs. The effect of clinic type on patient outcomes is also unknown. VHA has a system of PTSD specialty care clinics in addition to general mental health programs, and the majority of research on CPT and PE in VHA has focused on patients in PTSD specialty care. Information on whether clinic type accounts for differences between therapists in patient outcomes has implications for both research and clinic design.

Leadership plays an important role in ensuring that time and resources are available for evidence-based practices and shapes the extent to which use of an intervention is expected, supported, and rewarded by colleagues and supervisors, which collectively is referred to

as the *implementation climate* (Moullin et al., 2018). First-level leaders and a supportive implementation climate can increase adoption of evidence-based practices, including TFPs, in mental health treatment settings (Cook et al., 2015; Williams et al., 2020). It is unknown, however, whether these factors also augment intervention effectiveness. In the context of TFPs, this might occur if therapists who work in clinics that encourage TFP delivery convey positive treatment expectations to their patients, use measurement-based care to inform shared decision making, and seek to improve skills through consultation and additional training. If this were the case, then perceived support for TFP delivery may be associated with lower dropout and better clinical outcomes rates and variation in implementation support might at least partially explain differences between therapists.

The present study is based on a national sample of VHA therapists providing CPT and PE to patients with PTSD as part of routine care. The objectives of this research were to use multilevel modeling to estimate the size of TEs for TFP dropout and clinical effectiveness and to evaluate whether therapy delivery, namely therapist adherence and session spacing, and clinic organizational factors explained part of any observed TEs. We hypothesized there would be systematic differences between therapists in patient dropout, clinically meaningful improvement and recovery from PTSD, with some therapists having better results after controlling for patient case-mix. Our second hypothesis was that session spacing intensity and therapist adherence would at least partially account for differences between therapists found in the examination of the first hypothesis. Exploratory analyses examined the association between clinic organizational factors (clinic type, implementation leadership, implementation climate, caseload) and patient outcomes and whether these associations explained TEs.

Method

The Minneapolis VA Health Care System Institutional Review Board approved this research and determined that the criteria for waivers of HIPAA authorization and informed consent for patients as well as a waiver of documentation of informed consent for therapists were met. This article followed Journal Article Reporting Standards (Appelbaum et al., 2018).

Participants and Recruitment

To obtain a representative sample of therapists, we stratified the population of 2,962 licensed mental health professionals who provided individual CPT or PE in 2018 into 12 strata based on type of TFP they provided (CPT, PE, both) and US geographic region (West, South, Midwest, Northeast). We used the proportions of the 2,962 in each stratum to identify the sample stratum target proportions for study recruitment. We randomly ordered the 1,934 therapists providing TFPs to at least three patients in 2018 for recruitment into the study. In each of the two years prior to recruitment, on average, those therapists provided TFPs to more than 11 patients. To meet enrollment goals for the smaller strata (e.g., PE therapists in the Northeast), we added 743 therapists providing TFPs to at least three patients in 2019 to the sampling frame. Power analysis showed that for TEs from 5% to 10%, the range observed in most prior studies (Baldwin & Imel, 2013; Johns et al., 2019; Schiefele et al., 2017), a sample of 200 therapists, with an average of 10 patients, would provide power of at

least .85 to detect TEs for outcome rates between 30% and 70% in the planned analyses. Our target therapist sample size for recruitment was 350 to allow for exclusion of therapists who substantively reduced (< 3 patients) or stopped providing TFPs over the study period.

Therapist enrollment took place between May 2, 2019 and October 9, 2019. We randomly ordered therapists within strata and recruited within a stratum until we reached the target sample size. We emailed recruitment material to 1,139 therapists and excluded those who communicated that they were no longer providing CPT or PE ($n = 18$). Of the remaining 1,121 therapists recruited, 358 (31.9%) completed online consent. Among those who consented, we excluded therapists who did not complete the therapist survey ($n = 12$), did not have any TFP patients within the following year according to electronic medical record data ($n = 69$) or did not have at least three qualifying patients ($n = 97$).

At the time of this study, the vast majority of individual CPT and PE sessions were documented using templates that generate data elements stored in administrative data (Shiner et al., 2022). We used these templates and current procedure technology codes for individual psychotherapy to prospectively identify the 2,678 patients who began individual CPT or PE with the consented therapists within the year of therapists' consent. We used manual chart note review to exclude patients who: (a) were seen by unlicensed mental health professionals (e.g., psychology interns) working under the consented therapist ($n = 527$), (b) received psychotherapies other than individual CPT or PE even though the CPT or PE template was used ($n = 14$), (c) switched therapists ($n = 3$), (d) had only one TFP session ($n = 261$), or were seen by therapists who did not meet the above inclusion criteria ($n = 138$).

Procedures

Immediately after consenting, enrolled therapists were prompted to complete a 15-minute online survey to assess demographics and characteristics of their work environment. They were also prompted to watch a five-minute online tutorial on CPT and PE documentation using CPT and PE templates and emailed a one-page summary on "essential elements for documenting".

CPT and PE chart note templates include checklists for documenting the session-specific essential elements of each TFP. We manually rated therapy adherence based on documentation generated by the checklists in the templated notes (Wiltsey Stirman et al., 2022). Because it was infeasible to rate all sessions, we implemented a sampling strategy intended to provide stable estimates of therapists' adherence which, for cognitive therapy, can be achieved with five sessions per patient and four to five patients per therapist (Dennhag et al., 2012). We further considered the need to evaluate adherence across sessions that covered the different prescribed elements of each TFP (Barber et al., 2007). Therapists introduce new material through session 4 for PE and session 7 for CPT. Thus, to ensure an adequate number of sessions per patient and coverage of the different prescribed intervention elements, we rated the first seven TFP sessions per patient. To provide stable therapist-level estimates, we planned to rate the first 10 patients per therapist or all patients for therapist who provided TFPs to fewer than 10 patients. However, we rated more than 10 patients for some therapists in our training set. Because we did not rate all patients for therapists who had more than 10 TFP patients, some patients did not contribute to the adherence measure.

Trained raters were randomly assigned patients and rated all required sessions for those patients so they could get a full picture of the progression of each patient's therapy course. For training, we double coded all elements of all sessions for 106 (56 PE and 50 CPT) patients. Because agreement on individual items for each session exceeded 95% when raters agreed on the protocol session number (e.g., content from CPT protocol session 3 was covered), we had two raters code the session number (e.g., CPT session 3) but only one of them also coded the individual items within a session. We continued to double code a subset of all records to ensure ongoing calibration. We checked agreement in 11 batches. We used consensus coding when the agreement check found that the raters did not agree on the protocol session number (e.g., did not agree that content from CPT protocol session 3 was covered) or when it was requested by one rater because the documentation was confusing. The raters met weekly to review cases, resolve discrepancies and document coding decisions. Through this process, we rated 7,273 sessions for 1,253 (72.2%) of the 1,735 patients seen by the therapists in our sample. The 482 unrated patients were seen by therapists who had 10 patients whose sessions were rated for adherence.

Measures

Outcomes

Dropout: Prior research has classified dropout as completing fewer than 8 CPT or PE sessions (Kehle-Forbes et al., 2016; Maguen et al., 2019). However, those patients who have fewer than 8 sessions because they no longer need treatment should not be classified as having dropped out. Thus, we classified as *early completers* those patients who had fewer than 8 sessions but final PCL-5 scores of ≤ 18 (Schnurr et al., 2022) or a templated chart note documenting "early completer/symptoms remitted...additional sessions are not needed." Early completers were grouped with completers for analysis.

Clinically Meaningful Improvement and Recovery: The CPT and PE protocols used within VHA include assessment of PTSD symptoms using the PTSD Checklist for DSM-5 (PCL-5), a validated 20-item self-report measure to assess the DSM-5 symptoms of PTSD (Blevins et al., 2015; Bovin et al., 2016). Higher scores indicate more severe PTSD symptoms. Therapists enter item scores into a mental health database linked to the template when writing the session therapy note. We used data extraction from the electronic health record supplemented with manual chart review to identify total PCL-5 scores associated with each CPT or PE session. When a PCL-5 score was unavailable for an initial CPT or PE session, we extracted PCL-5 scores from the preceding two weeks to establish the baseline PCL-5. To calculate PTSD symptom change, we used the PCL-5 score closest to the last CPT or PE session. Consistent with recommended interpretation of PCL-5 scores at the time of study development (International Society for Traumatic Stress Studies, 2022; National Center for PTSD, 2022), we used a score of 32 or higher for PTSD diagnosis, a change of at least five points for reliable change (i.e., not due to chance) and a reduction of 10 or more points for clinically meaningful improvement. Patients meeting criteria for reliable change and loss of diagnosis were classified as having recovered (Jacobson & Truax, 1991; Wise, 2004). Because we extracted total scores from electronic medical records, we were not able to calculate coefficient alpha for this sample. Coefficient alpha was .96 in a psychometric evaluation of the PCL-5 in US veterans (Bovin et al., 2016)

Therapy Delivery Factors

Session Spacing.: We calculated the mean and the standard deviation of days between TFP sessions per patient. We created summary measures for each therapist to operationalize session spacing intensity and consistency, respectively, comprising the average of their patients' mean days between sessions and the average of the standard deviation of their patients' days between sessions. The means for these therapists' spacing measures were 11.49 days ($SD = 3.03$) for the intensity and 7.45 days ($SD = 2.53$) for the consistency measure. The median for session spacing intensity and consistency were 11.18, ($IQR = 2.66$) and 7.22 ($IQR = 2.87$) days, respectively.

Adherence.: Whether therapists administer the active elements of psychotherapy prescribed in the manual is a critical component of therapist adherence. Because TFPs are highly specified, these "unique and essential" elements are very well defined (Barber et al., 2007). Raters evaluated the categorical presence or absence of each session-specific unique and essential element in the templated notes using modified versions of adherence forms from a CPT and PE comparative effectiveness study (Schnurr et al., 2022). For example, for PE session 3, unique and essential elements include explaining, carrying out and processing imaginal exposure. We had separate adherence forms for CPT with and without written trauma accounts. The adherence rating forms are provided in Supplemental Material 1. When the templated notes revealed a break in the sequencing of sessions (e.g., we found templates for CPT sessions 3 and 5 but not session 4), we reviewed chart notes that were not templated and rated adherence in untemplated CPT and PE notes for that patient. The vast majority (95.3%) of the CPT and PE sessions in this sample were templated.

We calculated adherence scores for each session as the number of the unique and essential items present for a session out of the total number of unique and essential items included in the template for that session. If a therapist skipped a protocol session (e.g., documented content from sessions 1, 2, 4, but skipped 3), the removed session was scored 0% adherence. When a therapist repeated a session (e.g., provided session 2 content in two separate sessions), we combined the unique and essential elements documented across sequentially repeated sessions. For each therapist, we calculated the mean of the adherence scores for their TFP patients. Therapists' adherence scores ranged from 46.5% to 100% and were negatively skewed ($M(SD) = 86.92\% (9.36\%)$; $mdn (IQR) = 88.99\% (10.85\%)$).

Clinic Organizational Factors

Implementation Leadership.: We used the Implementation Leadership Scale (ILS) to assess the degree to which the leader of the clinic in which the therapist worked was proactive, knowledgeable, supportive and perseverant in regard to evidence-based practice implementation. The ILS has demonstrated excellent internal consistency as well as convergent and discriminant validity (Aarons et al., 2014). Participants rated their agreement with 12 statements about the clinic's leadership on a 5-point Likert scale ranging from "not at all" (0) to "very great extent" (4). We used two versions of the ILS, one for staff to report about their clinic leader, and another for clinic leaders to report about themselves. The term "evidence-based practice" was replaced with "CPT and PE" for this study. Total scores,

ranging from 0 to 4, were computed as the mean across all 12 items. Internal consistency was .96 in this sample.

Implementation Climate.: We used the Implementation Climate Scale (ICS) to assess staff perceptions of clinic policies, practices, procedures, and behaviors that are rewarded, supported, and expected to facilitate effective TFP implementation. The ICS is a psychometrically validated and reliable measure (Ehrhart et al., 2014). Participants rated agreement with 18 statements describing the implementation climate for the clinic in which they provided CPT or PE on a 5-point Likert scale ranging from “not at all” (0) to “very great extent” (4). Because of our interest in the overall implementation climate rather than sub-dimensions, we calculated total scores, ranging from 0 to 4, by computing a mean across all 18 items, as done in prior work (Williams et al., 2018, 2020). Internal consistency was 0.92 in this sample.

Caseload.: Caseload was computed as the average number of unique patients per month per therapist during the year following consent, regardless of the patients’ diagnoses or the clinical interventions received. The $M(SD)$ of caseload per month was 16.08 (7.99) patients.

PTSD Specialty Care Involvement.: We used VHA clinic stop codes to determine the clinic setting (PTSD specialty care vs. other) for each TFP patient. Therapist’s level of involvement in PTSD specialty care was calculated as the percent of their TFP patients seen in a PTSD specialty care clinic with scores ranging from 0 to 100%. Fifty six (31.1%) therapists saw patients in PTSD specialty care clinics only while 77 (42.8%) saw patients in other mental health clinics only. The remaining 47 (26.1%) therapists saw patients in both clinic types.

Participant Characteristics

Therapists.: We extracted the following variables on therapists from administrative data: VA facility of employment, census region of therapists’ facility, discipline, number of TFP patients in year prior to recruitment, and sex (gender is not available for staff). Therapist characteristics ascertained from the survey included supervisory responsibility (clinic director, staff member, other), number of years treating Veterans with PTSD, number of years since finishing professional degree, preferred theoretical orientation, race and ethnicity.

Patients.: We extracted from administrative databases the following variables anchored to TFP session 1: age, military service era, sex (gender is inconsistently available in administrative records), race, ethnicity, disability status for military-related PTSD (referred to as PTSD service connection), past year psychiatric comorbidities, past year medical comorbidities, past year psychiatric hospitalization, suicide risk flag, baseline and final session PCL-5 total scores and baseline depression measured with the Patient Health Questionnaire 9 (PHQ-9) (Kroenke et al., 2001). To avoid redundancy among the clinical variables used for case-mix adjustment, we created a composite baseline psychiatric severity indicator with high severity defined as baseline PCL ≥ 50 , or baseline PHQ-9 ≥ 15 , or prior year psychiatric hospitalization, or suicide risk flag. We used past year medical

comorbidities to compute Charlson Comorbidity Index scores (Charlson et al., 1987; Quan et al., 2005). We also extracted diagnoses associated with each TFP session. We manually extracted the following from clinical notes: education, employment, marital status, housing stability, index trauma type and history of childhood trauma and multiple trauma. Last, because the COVID-19 pandemic began during data collection for patients, we created a variable to classify each patient into one of three pandemic periods depending on the date of TFP initiation. The *pre-pandemic period* included those patients who began CPT or PE in 2019 ($n = 871$); the *early pandemic period* included those patients who began CPT or PE in January or February, 2020 ($n = 344$); the *pandemic period* included those patients who began CPT or PE after March, 2020 ($n = 520$).

Analysis

Initial analyses summarized data at the patient level to assess the distribution of patient dropout, clinically meaningful improvement and recovery. We used univariate logistic regression to examine the association between patient characteristics, TFP type (CPT vs. PE) and our outcomes and retained those characteristics significant at $p < .20$ for case-mix adjustment (Mickey & Greenland, 1989). For categorical variables (e.g., index trauma), a reference category was identified (e.g., military sexual trauma), and when a non-reference category (e.g., other sexual trauma) demonstrated no statistical difference from the reference category, that category was merged with the reference category (e.g., military and other sexual trauma versus combat trauma).

With the retained case-mix variables, we constructed multi-variable multilevel logistic regression models. This allowed us to model the structure of clinical care with patients nested within therapists, and to partition the total variance in outcomes between patient and therapist levels. We refer to the case-mix adjusted model as the *base model*. The multilevel model for an outcome was specified by $\text{logit } p_{ij} = \beta_0 + \tau_i + \beta_j'x_j$, $\tau_i \sim N(0, \sigma_\tau^2)$ where p_{ij} is the probability of the outcome for participant j treated by therapist i , β_0 is the model intercept, β_j the vector of coefficients for the case mix predictors x_j and τ_i the random effects, or residuals, for therapists.

These models can be viewed in a latent variable framework where for a given binary outcome Y , we assume there is a latent variable Y^* such that Y takes value 1 when Y^* is greater than 0. This framework seems applicable for the dropout and PCL-5 change outcomes considered here. Assume $Y^* = a_0 + a_j'x_j + e_i + e_{ij}$ with the $e_i \sim N(0, \sigma_\tau^2)$ as above and the e_{ij} following a logistic distribution with variance $\pi^2/3$. The TE variance estimated in fitting the multilevel logistic regression model, $\hat{\sigma}_\tau^2$, is an estimate of the variance of the therapist effects for the latent outcome. A common approach to gauge the amount of variance explained by the therapist effects is to estimate the proportion of remaining variance in the latent outcome stemming from the therapist effects, estimated by $\hat{\sigma}_\tau^2 / \left(\frac{\pi^2}{3} + \hat{\sigma}_\tau^2 \right)$

where $\hat{\sigma}_\tau^2 + \frac{\pi^2}{3}$ is the remaining variance in the latent outcome conditional on the included fixed effects.

The estimated TE variance $\hat{\sigma}_T^2$ and this estimated proportion of outcome variance at the therapist level are both summary measures of the TE for an outcome. For each fitted model, we used likelihood ratio tests to test for the presence of TEs. The therapist residuals represent the degree to which each therapist varies in their impact on outcomes relative to the average therapist, controlling for the included case-mix and fixed effect variables. The size of the estimated residuals, $\hat{\tau}_i$, can be used to make comparisons between therapists. For these models, the exponentiated therapists' residuals represent the odds ratio for the given outcomes associated with a given therapist relative to the average therapist.

A further aspect of these models is that with fixed effects in the models, these proportions of remaining variance are not directly analogous to R^2 measures and that proportions of variance explained in these models that appear small compared to R^2 measures can be associated with large differences between therapists. For example, a variance of $\sigma_T^2 = 1/2$ for the TE would yield an estimate of 13.2% of the remaining variance in latent outcomes stemming from TEs but the odds ratio comparing the rates of successful outcomes between therapists at the 75th and 25th percentile values for the τ_i would be approximately $\exp(1.35\sqrt{.5}) \approx 2.6$.

In subsequent analyses, we used logistic regression to examine the association between therapy delivery and clinic organizational factors and our three outcomes. These were therapist-level variables, meaning that the measure was computed specific to a therapist or had the same value for all the therapist's patients. To evaluate whether the therapy delivery and clinic organizational factors at least partially accounted for TE for a given outcome, we added each variable individually as a predictor to the base model and compared TEs with and without these predictors. Additionally, we examined the conjoint effect of predictors that explained some of the variation between therapists. The multilevel model for the outcomes here are altered from those above to $\text{logit } p_{ij} = \beta_0 + \tau_i + \beta_j'x_j + \beta_i'x_i$, where β_j is the vector of coefficients for the therapist-level predictors x_j . The Akaike information criterion (AIC) measures and likelihood ratio tests indicated that inclusion of the fixed effects for patient characteristics, therapy delivery and clinic organizational variables improved model fit compared with model that only included random effects.

The measure of therapy dropout was not missing for any patient. However, we did not have both baseline and final PCL-5 scores for 462 (26.6%) patients. Rather than impute missing PCL-5 data for one quarter of the sample, we based analyses of clinically meaningful improvement and recovery on this reduced, observed-data sample. Addressing these missing data could be done by imputing the outcome data using patient characteristics but this assumes that there is no therapist effect. Using methods to include TEs in the imputation process or a missing not at random estimation process seemed circular or problematic for an analysis intended to estimate the TEs. Analyses were implemented in SAS 9.4 and R 4.02.

Results

Participant Characteristics

One hundred eighty of the 358 therapist who consented met inclusion criteria. These therapists were working in 155 clinics across 137 VHA facilities including 17 in the Northeast, 37 in the Midwest, 39 in the West, and 54 in the South. The therapist sample largely comprised psychologists ($n = 98$, 54.4%) and social workers ($n = 76$, 42.2%) and most ($n = 156$, 86.7%) were clinical staff without leadership or administrative responsibilities. Most were White ($n = 157$, 87.2%) and non-Hispanic ($n = 171$, 95.0%). Two thirds ($n = 116$, 64.4%) of the sample had been treating veterans with PTSD for at least six years and the remaining one third had been treating veterans with PTSD for at least one year. About half ($n = 85$, 47.2%) had more than 11 years of clinical experience since finishing their professional degree. The preferred theoretical orientation was behavioral or cognitive behavioral for three quarters of the sample ($n = 136$, 75.6%). Fifty-eight therapists provided CPT only, 12 provided PE only and 110 provided both CPT and PE during the year following consent. The mean number of TFP patients per therapist was 9.64 ($SD = 7.63$). Thirteen therapists (7 CPT therapists, 1 PE therapist, and 5 who had provided both CPT and PE) were dropped from the clinically meaningful improvement and recovery analyses because they did not document final PCLs for any TFP patients.

The 180 therapist participants did not differ in terms of census region where they worked or discipline from the 941 therapists who were eligible for recruitment but not included in this study ($ps > .36$). They did differ ($ps < .05$) in terms of the number of TFP patients and type of TFP they provided in 2018, the year prior to recruitment. Specifically, compared with eligible therapists who did not enroll, therapists included in the sample provided TFPs to more patients ($M(SD) = 11.55 (10.16)$ vs. $8.19 (6.81)$), were more likely to provide both CPT and PE (52.8% vs. 27.1%) and less likely to provide CPT only (40.0% vs. 63.6%) in 2018.

The patient sample (Table 1) for dropout included 1,735 patients who received CPT ($n = 1,298$) or PE ($n = 437$) from the enrolled therapists. The patient sample for clinically meaningful improvement and recovery included the 1,273 of these patients who had an initial and final session PCL-5 scores. These 1,273 differed from the 462 without PCL-5 change scores in terms of Hispanic ethnicity (7.7% vs. 14.50% Hispanic), housing instability (7.4% vs. 3.7% with stable housing), unemployment (32.9% vs. 26.0%), childhood trauma history (37.2% vs. 31.2%), prior year rates of depressive disorders (70.9% vs. 65.4%) and substance-related and addictive disorders other than alcohol (15.2% vs. 8.7%), Charlson Comorbidity Index scores ($t(964.49) = 2.536$, $p = .01$) and TFP type (76.36% vs. 70.56% received CPT). Almost one quarter of patients were women and slightly more than one quarter were people of color. Slightly more than half of the patients focused on combat trauma during CPT or PE. Multiple traumas were documented in over half of the sample.

Distribution of Outcomes

Eight hundred and eight-nine (51.2%) of the 1,735 patients dropped out. The odds of dropout did not differ for PE and CPT (54.5% vs. 50.2%, $OR = 1.189$, 95% CI [0.957,

1.478], $p = 0.119$). Seventy-two (8.5%) of the 846 who completed a TFP did so in fewer than 8 sessions. A larger proportion of PE ($n = 27$ (6.2%)) than CPT ($n = 45$ (3.5%)) patients were early completers ($\chi^2(1) = 6.044$, $p = .014$). The mean number of sessions among the remaining 774 therapy completers was 11.37 ($SD = 2.43$). Among these 774 patients, those completing CPT had on average 0.74 more sessions than those patients completing PE ($M(SD) = 11.54(2.28)$ vs. $10.80(2.82)$, $t(772) = 3.53$, $p = .004$).

Baseline PCL-5 scores were available for 1,423 patients. The mean for baseline PCL-5 scores was 50.69 ($SD = 14.0$). Of the 1,273 patients who had both baseline and final session PCL-5 values, 547 (43.0%) had 10-point reduction in scores, meeting criterion for clinically meaningful improvement; 376 (29.5%) had five-point reduction in PCL-5 scores and achieved a final PCL-5 score ≤ 32 , meeting criteria for recovery. The odds of clinically meaningful improvement and recovery were similar for CPT and PE ($ps > 0.30$).

Preliminary Analyses

Dropout and clinical outcomes—The mean for PCL-5 change was 2.78 ($SD = 10.58$) among those who dropped out and 15.78 ($SD = 17.39$) among those who completed a TFP ($p < .001$). Patients who dropped out were much less likely to meet criteria for clinically meaningful improvement ($OR = 0.188$, 95% CI [0.147–0.241], $p < 0.001$) or recovery ($OR = 0.146$, 95% CI [0.107–0.196], $p < 0.001$) than those who completed a TFP.

Therapist delivery and clinic organizational variables and patient outcomes—Table 2 presents the odds ratios for each therapist-level predictor and study outcomes. Session spacing intensity and consistency were associated with dropout. That is, the odds of dropout increased as therapists' mean and average standard deviation of days between their patients' TFP sessions increased. The odds of dropout also increased as caseload increased but decreased as the implementation climate and implementation leadership improved. The odds of dropout marginally decreased as adherence improved.

Session spacing intensity and implementation leadership were associated with clinically meaningful improvement. Specifically, the odds of clinically meaningful improvement decreased as therapists' mean of days between sessions increased but increased as implementation leadership scores improved. None of these predictors were associated with recovery.

Therapist Effects for Patient Dropout

Table 3 presents variance components for the random effects and the odds ratios for the fixed effects for patient case-mix variables and TFP type in the base model for dropout. Among the patient characteristics meeting model inclusion criterion, younger age, lower education, and beginning TFP in the early phase of the pandemic increased the odds of dropout. TEs accounted for 5.812% of the unexplained variance adjusting for case-mix and TFP type. The average dropout rate for the 45 therapists in the top performing quartile was 27.0% while the average dropout rate for the 45 therapists in the bottom performing quartile was 78.8%. The average of the individual model estimated odds ratios for dropout among the therapists in the best performing quartile was 0.75 ($SD = 0.10$, range 0.52 – 0.87). For the therapists in

bottom performing quartile, the average of the model estimated odds ratios for dropout was 1.37 ($SD = 0.25$, range 1.15 – 2.29).

Figure 1A presents the TEs without case-mix adjustment (random effects), with adjustment for case mix and TFP type (base model) and when each therapist-level factor was added to the base model; Figure 1B shows the corresponding adjusted odds ratios for each predictor. The full output for each model is presented in Supplemental Material 2. Adding case-mix and TFP type reduced the TE by 10.6%, from 6.503% to 5.812%. The individual predictor associated with the largest reduction in TEs was session spacing intensity (therapists' mean of days between sessions) followed by implementation climate scores and session spacing consistency (therapists' standard deviation of days between sessions), all of which were significant fixed effects in the respective multilevel models ($p < .001$, see Supplemental Material 2). Including therapists' caseload, implementation leadership and adherence also led to small reductions in TEs, although the respective fixed effects were not significant. The TE was reduced to 2.031% and non-significant ($p = .140$) when therapists' session spacing intensity, adherence, caseload and implementation climate scores were simultaneously added to the base model. This was also the model with the lowest AIC (base model AIC: 2,340.546; base model + therapy delivery and clinic organizational variables AIC: 2,310.609). We did not include session spacing consistency or implementation leadership in this larger model because they were strongly correlated ($r > .71$) with session spacing intensity and implementation climate, respectively. We did not include PTSD specialty care involvement in this larger model because its inclusion did not alter the magnitude of TEs for dropout.

Therapist Effects for Clinical Effectiveness

Clinically meaningful improvement—Table 4 presents the variance components for the random effects and the odds ratios for the fixed effects for patient case-mix variables meeting criterion for inclusion in the base model for clinically meaningful improvement. The odds of clinically meaningful improvement were higher for patients who were retired ($p = .002$) and had high baseline psychiatric severity ($p < .001$). Controlling for case-mix, the TE, 2.032% of the unexplained variance, was not significant ($p = .237$).

The TE for clinically meaningful improvement was not significant and remained close to 2% of the unexplained variance when we added each of the therapy delivery and clinical organizational variables to the base model (data not shown). In post hoc analyses to understand why TEs would explain variation in patient dropout but not clinically meaningful improvement, we added dropout to the base model. The respective AIC values for the base model with and without dropout were 1,601.495 and 1,720.18, showing that the addition of dropout improved model fit. The TE for clinically meaningful improvement increased to 5.023% ($p = .025$), with an accompanying large fixed effect for dropout ($OR = 0.161$, 95% CI [0.112, 0.212], $p < 0.001$).

Among patients who dropped out, the estimated proportion of unexplained variance in clinically meaningful improvement explained by therapists was 2.830% ($p = .272$) whereas among patients who completed a TFP, it was 8.446% ($p = .005$), adjusting for case mix. Restricting the sample to those patients who completed a TFP, the average rate of clinically

meaningful improvement among the 38 therapists in the highest performing quartile was 63.1% while the average rate of clinically meaningful improvement among the 38 therapists in the lowest performing quartile was 24.0%. Among patients who completed a TFP, the average of the individual estimated odds ratios for clinically meaningful improvement were 1.36 ($SD = 0.23$, range 1.15 – 2.16) for therapists in the highest performing quartile and 0.74 ($SD = 0.10$, range 0.48 – 0.85) for therapists in the lowest performing quartile.

Recovery—Table 4 also presents the variance components for the random effects and the odds ratios for the fixed effects for patient case-mix variables meeting criterion for inclusion in the base model for recovery. The odds of recovery were lower for Native American, Hawaiian, and Asian patients considered together ($p = .012$) and for Black patients ($p = .012$), and for those with high baseline psychiatric severity ($p < .001$); the odds of recovery were higher for patients who were retired from employment ($p = .001$). Adjusting for case mix, there was no difference between therapists in patient recovery. The magnitude of TEs remained close to zero when we added the therapy delivery and clinical organizational predictors to the base model. In post hoc analysis, the TE for recovery with dropout added to the base model was 2.403% of the unexplained variance and remained non-significant ($p = .301$). Among patients who completed a TFP, the TE for recovery adjusting for case mix was 4.677% and approached significance ($p = .074$), compared with 1.978% ($p = .429$) among patients who dropped out.

Discussion

This study quantified TEs for TFPs in routine care in VHA and identified potentially modifiable therapy delivery and clinic organizational factors that explained observed TEs. Our first hypothesis was that there would be variation between therapists in dropout and clinical effectiveness. We found that almost 6% of the unexplained variance in dropout was due to variation between therapists, controlling for case mix and TFP type. Were TEs equivalent to changes in R^2 , this would be considered a small to medium-size effect. However, TEs are the proportion of conditional variance (variance that is left unexplained) in the latent variable outcome given the predictors and we are not aware of any convention for small, medium and large TEs. If a large proportion of the conditional outcome rates fall between .20 and .80, as is the case here, then there will remain a large amount of unexplained variance in a well-fitting logistic regression model as the conditional standard deviation in the outcome given the predictors will be relatively large (0.40 or more) compared to the maximum conditional standard deviation (0.50). Interpretation of TEs is aided by examining the odds ratios from the exponentiated TE residuals. Here, the odds of dropout for patients seen by therapists in the lowest performing quartile were 1.83 times greater than for patients seen by therapists in the highest performing quartile. The fact that TEs of this magnitude can have meaningful implications for clinical care was further demonstrated by the observation that on average approximately one in four patients dropped out among therapists in the top while almost four in five patients dropped out among therapists in the bottom performing quartile.

We did not observe a significant TE for clinically meaningful improvement until dropout was taken into consideration in post hoc analysis. This is because there was little variation

in clinically meaningful improvement to explain among the 51.2% of the sample that dropped out. Among patients who dropped out, only 22.0% achieved clinically meaningful improvement. In contrast, among patients who completed a TFP, 60.6% experienced clinically meaningful improvement and therapists accounted for 8.446% of the conditional variance in this outcome. The nonsignificant TE for clinically meaningful improvement in the full sample can be understood as reflecting the TEs averaged across patients who dropped out and those who completed. Similarly, another study reported no TE for PHQ-9 improvement among patients who dropped out, compared with 11.2% among patients who completed CBT or counseling (Saxon et al., 2017). We had not hypothesized that the TE for clinical outcomes would depend on therapy completion. Therefore, we interpret the TE for clinically meaningful improvement among therapy completers with caution and recommend that researchers take dropout into consideration when designing TE studies for clinical outcomes, particularly if dropout rates are likely to be large.

Dropout was also strongly related to recovery, with only 10.9% experiencing recovery among patients who dropped out and 45.6% experiencing recovery among those who completed at least 8 sessions or were early completers. However, even after dropout was added as a predictor, we did not observe a TE for recovery. This may be related to the fact that recovery was less common than expected from efficacy studies in active duty and veteran samples (Steenkamp et al., 2015, 2020). In our sample, less than 30% of patients met study criteria for recovery. We operationalized clinical outcomes based on consensus-based recommendations (International Society for Traumatic Stress Studies, 2022; National Center for PTSD, 2022). However, a recent study based on samples of male veterans suggests that these recommendations under-estimate the magnitude of within-person change in PCL-5 scores needed to identify change that is not due to measurement error (Marx et al., 2022). If confirmed in other studies based on more diverse samples, then the proportions of TFP patients meeting criteria for clinically meaningful improvement and recovery would have been even smaller. In post hoc simulation studies, where we varied the overall outcome rate and TE using therapist and patient sample sizes comparable to those here, we found that multilevel models were often unable to calculate variances for TEs ranging from 3% to 11% and outcomes rates 30%. We also observed that they often resulted in zero variance estimates. Others have found that multilevel models may underestimate TEs (Capanu et al., 2013; Tuerlinckx et al., 2006) or that very large samples are needed to identify TEs using multilevel modeling (Schiefele et al., 2017). The occurrence of these issues varies with the sample size, outcome rate, TE magnitude, and method for fitting the models. Thus, alternative analytic methods for evaluating TEs in situations that are less than optimal for multilevel modeling need to be developed.

Our second hypothesis was that session spacing intensity and adherence would at least partially explain TEs. On average, therapists provided TFP sessions with considerably longer ($M = 11.49$; $SD = 3.03$) between-session intervals than the once or twice a week interval prescribed in the treatment manuals and tested in most RCTs. TFP treatment protocols do not take into consideration the effect of patient and/or therapist scheduling constraints or unpredictability, logistics and preferences on session spacing in routine practice. While therapists' TFP session spacing intensity and consistency were less than specified in treatment manuals, the pattern of findings involving the session spacing

measures is consistent with RCTs showing that dropout rates are lower when a more intensive TFP course is prescribed (Foa et al., 2018; Galovski et al., 2022; Levinson et al., 2022; Sciarrino et al., 2020) and PTSD symptoms improvement is inversely related to TFP session frequency and consistency (Gutner et al., 2016).

As expected, session spacing intensity and to a lesser extent session spacing consistency partially explained the TE for dropout. This finding is novel and supports the development of interventions to reduce the variation between therapists in session spacing. Such interventions could include scheduling grids that facilitate weekly sessions over an episode of care or audit and feedback of days between session. At minimum, therapists and their patients should be informed about the contribution of session spacing to outcomes so that they can be intentional about scheduling and attendance. The fact that TFP dropout rates among active duty and military veterans are generally worse in routine care than in efficacy studies (Maguen et al., 2019 vs. Steenkamp et al., 2020) suggests that dedicating resources to scheduling sessions according to protocol and outreach to patients who miss sessions, both of which are more likely in clinical trials, may improve the intensity and consistency with which TFP sessions occur.

The hypothesis that therapist adherence would partially account for TEs was not strongly supported. In preliminary analyses, therapists' mean adherence scores were marginally related to the odds of dropout, but not to clinically meaningful improvement or recovery. The TE for dropout decreased slightly, from 5.812% to 5.530%, when adherence was added to the base model, suggesting it is a potentially modifiable factor that could have an incremental effect on patient retention in therapy. The relationship between adherence and patient outcomes is complex, as demonstrated by mixed results from prior studies that examined the relationship between therapist adherence and outcomes in other clinical contexts (Webb et al., 2010). However, contrary to some concerns and qualitative findings that high or "rigid" levels of adherence could reduce patient retention in TFPs (Doran & DeViva, 2018), our findings suggest that adherence has a neutral to positive association with TFP completion. We also note that because this study was designed to examine therapists' contribution to outcomes, we created adherence scores for each therapist which represented their average level of adherence. However, both therapists and patients contribute to adherence (Wiltsey Stirman et al., 2022). Patient factors that may be related to adherence include understanding, motivation, degree of improvement, and culture. Additionally, therapist competence, rather than adherence, may be more strongly associated with patient outcomes (Keefe et al., 2022; Marques et al., 2019) and is not discernable through documentation using templates.

In exploratory analyses, we examined whether clinic organizational factors were associated with patient outcomes and accounted for TEs. Therapists' involvement in PTSD specialty care clinics was not associated with the odds of dropout, clinically meaningful improvement or recovery and did not explain the TE for dropout. Instead, what mattered was the perceived support for TFP delivery that therapists received in their clinics. Specifically, therapists' perception of the implementation climate (i.e., extent to which clinic policies, practices, procedures and expectations are perceived as supporting TFP delivery) was associated with dropout while their perceptions of implementation leadership (the extent to which leaders

are perceived as proactive, knowledgeable, supportive and perseverant in regard to TFT implementation) was associated with dropout and clinically meaningful improvement. The TE for dropout were reduced by 23.1% (from 5.812% to 4.469%) when implementation climate scores were included in the multilevel model. Prior research has shown that therapists are more likely to use an evidence-based treatment if their work environment and leadership support its use (Williams et al., 2018). A novel and important finding in this study is that these clinic organizational factors can also affect TFP outcomes and partially explain variation between therapists in dropout. Both implementation climate and leadership are modifiable through training in leadership and organizational change and implementation support (Aarons et al., 2015; Williams et al., 2020; Worley et al., 2022).

While caseload was not associated with dropout in the multi-variable model adjusting for case mix, we find it interesting that it was associated with dropout in bivariate analysis. Despite its importance to clinic design, there is limited research on the contribution of caseload size to psychotherapy outcomes. We conjecture that a large caseload may make it difficult for therapists to put time and energy into patient engagement. Prior research has shown that therapists of TFP completers are more likely to provide patient-centered, flexible care and join with their patients in the mission of completing treatment (Kehle-Forbes et al., 2022). This type of TFP delivery may require additional time and effort that is not available to those with a higher caseload. Research on caseload and therapy delivery is needed to establish productivity expectations that optimize therapists' ability to retain patients in TFPs. This is critical because, as shown here, patients who drop out of therapy generally do not experience a clinically meaningful reduction in symptoms.

We are not aware of prior studies that examined the role of therapy delivery and clinic organizational factors in explaining variation between therapists in patient dropout or clinical outcomes. When therapists' session intensity, adherence, caseload and implementation climate scores were simultaneously added to the base model for dropout, the TE was reduced to 2.032% and non-significant. Systematic variation between therapists in TFP dropout was explained by this combination of therapy delivery and clinic organizational variables. These findings provide potentially actionable targets for interventions to improve TFP retention rates among low performing therapists.

Limitations

Limitations associated with the study design include lack of measurement of therapist competence, which, along with adherence, is a key component of fidelity (Waltz et al., 1993). We would have needed to audio-record therapy sessions for each therapist to evaluate competence, something that was beyond the scope of this work. Adherence ratings were made based on ratings of clinician documentation using templated checklists of essential therapy elements. This is a potential limitation, although recent research demonstrated that therapists in routine care settings can accurately report on their use of evidence-based treatment elements, including CPT (Gumport et al., 2020, 2021). This study would have also been improved by assessment of the therapeutic alliance, which is understudied in the context of TFPs (cf. Keefe et al., 2022). An understanding of whether the therapeutic alliance explains differences between therapists in patient outcomes has the potential to

inform research, therapist training and practice. Another design weakness is that we did not assess the reasons for the larger than prescribed number of days between TFP sessions. A better understanding of what contributes to the decrease in therapy session frequency and consistency in routine care compared with what is specified in the treatment manuals would help in overcoming obstacles to ensuring that patients receive an adequately intense therapy course.

Our sampling strategy was designed to obtain a nationally representative sample of TFP therapists. However, because we recruited therapist from strata defined by the type of therapy they provided (CPT, PE, both) and the census region rather than the clinic in which they worked, the ratio of therapists to clinics (generally one or two to one) in our sample did not allow for examination of clinic effects or the aggregation of the organizational measures to the clinic level. Indeed, because many of the therapists in the study were the only participant from their clinic, separating the clinic effects from the therapist effects was not possible. Instead, the implementation climate and leadership measures reflected therapists' individual perceptions and, along with workload, may have varied by therapist even when they were in the same clinic. A next step in this line of research would be to sample a sufficient number of therapists per clinic to also examine clinic effects for patient outcomes, something that has not been previously investigated, and to evaluate whether specific clinic policies and procedures explain any clinic effects. We also acknowledge that we did not evaluate how and why clinic organizational factors explained therapist effects for dropout. A better understanding the mechanisms by which clinic structures and processes affect therapists would improve our ability to design clinics that support the delivery of high-quality mental health care.

Another issue related to our therapist sample is that 166 (46.4%) of the 358 therapists who consented were excluded because they provided TFPs to zero to two patients during the 12 months following consent, despite providing TFPs more regularly in the prior two years. It is unknown whether this change in practice was specific to the timeframe for this study (discussed below) or reflective of a pattern where therapists either stop providing TFPs or provide them infrequently over time. Regardless, these findings do not pertain to therapists who provide TFPs to very few patients over a 12-month period.

A final set of limitations stem from factors beyond investigators' control. The COVID-19 pandemic began during data collection and disrupted routine care. In particular, therapists provided TFPs to fewer patients than expected based on data collected to develop the recruitment targets. Second, the pandemic was associated with a reduction in the number PCL-5 scores documented. Therapists in this study documented an initial PCL-5 scores for 85.5% of patients in the dropout sample who initiated a TFP before March, 2020 and to 74.2% of patients who initiated a TFP after March, 2020. It is unknown whether patients without PCL-5 change scores differed from those with PCL-5 change scores in terms of clinical outcomes and therefore whether their exclusion from the analyses of clinically meaningful improvement and recovery biased our results. Third, as reported here, the odds of dropout increased for patients who began a TFP during the early phase of the pandemic (Table 3), likely due to the switch from in-person to telemedicine during the therapy course. As a result, fewer of the enrolled therapists met our inclusion criteria and the number of

patients per therapist, particularly the number of patients with PCL-5 change scores, was less than expected. There was, however, no difference in clinical effectiveness by pandemic period. Unfortunately, the numbers of patients per therapist who had initial and final TFP sessions in each pandemic period were not large enough to compare TEs before to during the COVID-19 pandemic.

Conclusions

In routine VHA care, some therapists were more effective than others at retaining patients in TFPs. Variation between therapists for dropout from these manualized treatments was seen even after controlling for demographic and clinical case-mix variables associated with dropout. Because patients who dropped out from a TFP generally did not demonstrate clinically meaningful improvement or recovery, identification of modifiable factors that enhance therapists' ability to reduce dropout is critical. We found that therapy delivery and clinic organizational factors explained TEs for dropout. Taken together, findings support the testing of implementation interventions that target clinic support for TFP delivery, therapists' caseload and their ability to deliver the essential intervention components at the intensity and consistency prescribed in treatment manuals. While such interventions are not likely to eliminate dropout, they have the potential to reduce variation between therapists in dropout so that more patients stay engaged long enough to experience clinical benefit.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was funded by the US Department of Veterans Affairs (VA), Health Services Research and Development (HSR&D) (IIR 17-178). The content is solely the responsibility of the authors and does not necessarily represent the views of the VA or HSR&D.

This study was conducted, and the analyses were performed as specified in the grant-funded protocol. It was not formally preregistered. The dataset has not been fully deidentified and, due to privacy and ethical concerns, is not available.

References

- Aarons GA, Ehrhart MG, & Farahnak LR (2014). The implementation leadership scale (ILS): Development of a brief measure of unit level implementation leadership. *Implementation Science*, 9(1). 10.1186/1748-5908-9-45
- Aarons GA, Ehrhart MG, Farahnak LR, & Hurlburt MS (2015). Leadership and organizational change for implementation (LOCI): A randomized mixed method pilot study of a leadership and organization development intervention for evidence-based practice implementation. *Implementation Science*, 10(1). 10.1186/s13012-014-0192-y
- Appelbaum M, Cooper H, Kline RB, Mayo-Wilson E, Nezu AM, & Rao SM (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, 73(1), 3–25. 10.1037/amp0000191 [PubMed: 29345484]
- Baldwin SA, & Imel ZE (2013). Therapist effects: Findings and methods. In Lambert MJ (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 258–297). Wiley.

- Barber JP, Triffleman E, & Marmar C (2007). Considerations in treatment integrity: Implications and recommendations for PTSD research. *Journal of Traumatic Stress*, 20(5), 793–805. 10.1002/jts.20295 [PubMed: 1795529]
- Blevins CA, Weathers FW, Davis MT, Witte TK, & Domino JL (2015). The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): Development and initial psychometric evaluation. *Journal of Traumatic Stress*, 28(6), 489–498. 10.1002/jts.22059 [PubMed: 26606250]
- Bovin MJ, Marx BP, Weathers FW, Gallagher MW, Rodriguez P, Schnurr PP, & Keane TM (2016). Psychometric properties of the PTSD Checklist for Diagnostic and Statistical Manual of Mental Disorders–Fifth Edition (PCL-5) in veterans. *Psychological Assessment*, 28(11), 1379–1391. 10.1037/pas0000254 [PubMed: 26653052]
- Capanu M, Gönen M, & Begg CB (2013). An assessment of estimation methods for generalized linear mixed models with binary outcomes. *Statistics in Medicine*, 32(26), 4550–4566. 10.1002/sim.5866 [PubMed: 23839712]
- Charlson ME, Pompei P, Ales KL, & MacKenzie CR (1987). A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *Journal of Chronic Diseases*, 40(5), 373–383. 10.1016/0021-9681(87)90171-8 [PubMed: 3558716]
- Cook JM, Dinnen S, Thompson R, Ruzek J, Coyne JC, & Schnurr PP (2015). A quantitative test of an implementation framework in 38 VA residential PTSD programs. *Administration and Policy in Mental Health and Mental Health Services Research*, 42(4), 462–473. 10.1007/s10488-014-0590-0 [PubMed: 25199813]
- Cooper AA, Kline AC, Baier AL, & Feeny NC (2018). Rethinking Research on Prediction and Prevention of Psychotherapy Dropout: A Mechanism-Oriented Approach. *Behavior Modification*, 014544551879225. 10.1177/0145445518792251
- Dennhag I, Gibbons MBC, Barber JP, Gallop R, & Crits-Christoph P (2012). How many treatment sessions and patients are needed to create a stable score of adherence and competence in the treatment of cocaine dependence? *Psychotherapy Research*, 22(4), 475–488. 10.1080/10503307.2012.674790 [PubMed: 22449079]
- Doran JM, & DeViva J (2018). A naturalistic evaluation of evidence-based treatment for veterans with PTSD. *Traumatology*, 24(3), 157–167. 10.1037/trm0000140
- Eftekhari A, Crowley JJ, Mackintosh M-A, & Rosen CS (2020). Predicting treatment dropout among veterans receiving prolonged exposure therapy. *Psychological Trauma: Theory, Research, Practice, and Policy*, 12(4), 405–412. 10.1037/tra0000484 [PubMed: 31318247]
- Ehrhart MG, Aarons GA, & Farahnak LR (2014). Assessing the organizational context for EBP implementation: The development and validity testing of the Implementation Climate Scale (ICS). *Implementation Science*, 9(1). 10.1186/s13012-014-0157-1
- Farmer CC, Mitchell KS, Parker-Guilbert K, & Galovski TE (2017). Fidelity to the cognitive processing therapy protocol: Evaluation of critical elements. *Behavior Therapy*, 48(2), 195–206. 10.1016/j.beth.2016.02.009 [PubMed: 28270330]
- Fernandez E, Salem D, Swift JK, & Ramtahal N (2015). Meta-analysis of dropout from cognitive behavioral therapy: Magnitude, timing, and moderators. *Journal of Consulting and Clinical Psychology*, 83(6), 1108–1122. 10.1037/ccp0000044 [PubMed: 26302248]
- Finley EP, Garcia HA, Ketchum NS, McGeary DD, McGeary CA, Stirman SW, & Peterson AL (2015). Utilization of evidence-based psychotherapies in Veterans Affairs posttraumatic stress disorder outpatient clinics. *Psychological Services*, 12(1), 73–82. 10.1037/ser0000014 [PubMed: 25419915]
- Foa EB, McLean CP, Zang Y, Rosenfield D, Yadin E, Yarvis JS, Mintz J, Young-McCaughan S, Borah EV, Dondanville KA, Fina BA, Hall-Clark BN, Lichner T, Litz BT, Roache J, Wright EC, Peterson AL, & for the STRONG STAR Consortium. (2018). Effect of Prolonged Exposure Therapy Delivered Over 2 Weeks vs 8 Weeks vs Present-Centered Therapy on PTSD Symptom Severity in Military Personnel: A Randomized Clinical Trial. *JAMA*, 319(4), 354. 10.1001/jama.2017.21242 [PubMed: 29362795]
- Galovski TE, Werner KB, Weaver TL, Morris KL, Dondanville KA, Nanney J, Wamser-Nanney R, McGlinchey G, Fortier CB, & Iverson KM (2022). Massed cognitive processing therapy for posttraumatic stress disorder in women survivors of intimate partner violence. *Psychological*

- Trauma: Theory, Research, Practice, and Policy, 14(5), 769–779. 10.1037/tra0001100 [PubMed: 34472941]
- Goetter EM, Bui E, Ojserkis RA, Zakarian RJ, Brendel RW, & Simon NM (2015). A systematic review of dropout from psychotherapy for posttraumatic stress disorder among Iraq and Afghanistan Combat Veterans: Dropout from PTSD treatment in OEF/OIF veterans. *Journal of Traumatic Stress, 28*(5), 401–409. 10.1002/jts.22038 [PubMed: 26375387]
- Gumport NB, Hernandez S, Colloway A, Vogt D, Dean K, Young SJ, Johnson C, Song J, DeRubeis R, Park AL, Margues L, Creed TA, & Stirman SW (2021, December). An evaluation of clinician-rated measures of fidelity to cognitive processing therapy and cognitive behavior therapy in routine mental health care settings. 14th Annual Conference on the Science of Dissemination and Implementation in Health, Online.
- Gumport NB, Yu SH, Mullin AC, Mirzadegan IA, & Harvey AG (2020). The Validation of a Provider-Reported Fidelity Measure for the Transdiagnostic Sleep and Circadian Intervention in a Community Mental Health Setting. *Behavior Therapy, 51*(5), 800–813. 10.1016/j.beth.2019.11.006 [PubMed: 32800307]
- Gutner CA, Suvak MK, Sloan DM, & Resick PA (2016). Does timing matter? Examining the impact of session timing on outcome. *Journal of Consulting and Clinical Psychology, 84*(12), 1108–1115. 10.1037/ccp0000120 [PubMed: 27213491]
- Hale AC, Bohnert KM, Ganoczy D, & Sripada RK (2019). Predictors of Treatment Adequacy During Evidence-Based Psychotherapy for PTSD. *Psychiatric Services, 70*(5), 367–373. 10.1176/appi.ps.201800361 [PubMed: 30784379]
- Hamblen JL, Norman SB, Sonis JH, Phelps AJ, Bisson JI, Nunes VD, Megnin-Viggars O, Forbes D, Riggs DS, & Schnurr PP (2019). A guide to guidelines for the treatment of posttraumatic stress disorder in adults: An update. *Psychotherapy, 56*(3), 359–373. 10.1037/pst0000231 [PubMed: 31282712]
- Holder N, Holliday R, Williams R, Mullen K, & Surfis A (2018). A preliminary examination of the role of psychotherapist fidelity on outcomes of cognitive processing therapy during an RCT for military sexual trauma-related PTSD. *Cognitive Behaviour Therapy, 47*(1), 76–89. 10.1080/16506073.2017.1357750 [PubMed: 28793834]
- International Society for Traumatic Stress Studies. (2022). PTSD Checklist for the DSM-5 (PCL-5). <https://istss.org/clinical-resources/assessing-trauma/ptsd-checklist-dsm-5>
- Jacobson NS, & Truax P (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12–19. 10.1037/0022-006X.59.1.12 [PubMed: 2002127]
- Johns RG, Barkham M, Kellett S, & Saxon D (2019). A systematic review of therapist effects: A critical narrative update and refinement to review. *Clinical Psychology Review, 67*, 78–93. 10.1016/j.cpr.2018.08.004 [PubMed: 30442478]
- 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. *Journal of Traumatic Stress, 23*(6), 663–673. 10.1002/jts.20588 [PubMed: 21171126]
- Keefe JR, Hernandez S, Johaneck C, Landy MSH, Sijercic I, Shnaider P, Wagner AC, Lane JEM, Monson CM, & Stirman SW (2022). Competence in Delivering Cognitive Processing Therapy and the Therapeutic Alliance Both Predict PTSD Symptom Outcomes. *Behavior Therapy, 53*(5), 763–775. 10.1016/j.beth.2021.12.003 [PubMed: 35987537]
- Kehle-Forbes SM, Ackland PE, Spont MR, Meis LA, Orazem RJ, Lyon A, Valenstein-Mah HR, Schnurr PP, Zickmund SL, Foa EB, Chard KM, Alpert E, & Polusny MA (2022). Divergent experiences of U.S. veterans who did and did not complete trauma-focused therapies for PTSD: A national qualitative study of treatment dropout. *Behaviour Research and Therapy, 154*, 104123. 10.1016/j.brat.2022.104123 [PubMed: 35644083]
- Kehle-Forbes SM, Meis LA, Spont MR, & Polusny MA (2016). Treatment initiation and dropout from prolonged exposure and cognitive processing therapy in a VA outpatient clinic. *Psychological Trauma: Theory, Research, Practice, and Policy, 8*(1), 107–114. 10.1037/tra0000065 [PubMed: 26121175]

- Kitchiner NJ, Lewis C, Roberts NP, & Bisson JI (2019). Active duty and ex-serving military personnel with post-traumatic stress disorder treated with psychological therapies: Systematic review and meta-analysis. *European Journal of Psychotraumatology*, 10(1), 1684226. 10.1080/20008198.2019.1684226 [PubMed: 31762951]
- Kroenke K, Spitzer RL, & Williams JBW (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9), 606–613. 10.1046/j.1525-1497.2001.016009606.x [PubMed: 11556941]
- Laska KM, Smith TL, Wislocki AP, Minami T, & Wampold BE (2013). Uniformity of evidence-based treatments in practice? Therapist effects in the delivery of cognitive processing therapy for PTSD. *Journal of Counseling Psychology*, 60(1), 31–41. 10.1037/a0031294 [PubMed: 23356465]
- Levinson DB, Halverson TF, Wilson SM, & Fu R (2022). Less dropout from prolonged exposure sessions prescribed at least twice weekly: A meta-analysis and systematic review of randomized controlled trials. *Journal of Traumatic Stress*, 35(4), 1047–1059. 10.1002/jts.22822 [PubMed: 35278229]
- Maguen S, Holder N, Li Y, Madden E, Neylan TC, Seal KH, Lujan C, Patterson OV, DuVall SL, & Shiner B (2020). Factors associated with PTSD symptom improvement among Iraq and Afghanistan veterans receiving evidenced-based psychotherapy. *Journal of Affective Disorders*, 273, 1–7. 10.1016/j.jad.2020.04.039 [PubMed: 32421589]
- Maguen S, Li Y, Madden E, Seal KH, Neylan TC, Patterson OV, DuVall SL, Lujan C, & Shiner B (2019). Factors associated with completing evidence-based psychotherapy for PTSD among veterans in a national healthcare system. *Psychiatry Research*, 274, 112–128. 10.1016/j.psychres.2019.02.027 [PubMed: 30784780]
- Marques L, Valentine SE, Kaysen D, Mackintosh M-A, Dixon De Silva LE, Ahles EM, Youn SJ, Shtasel DL, Simon NM, & Wiltsey-Stirman S (2019). Provider fidelity and modifications to cognitive processing therapy in a diverse community health clinic: Associations with clinical change. *Journal of Consulting and Clinical Psychology*, 87(4), 357–369. 10.1037/ccp0000384 [PubMed: 30883163]
- Marx BP, Lee DJ, Norman SB, Bovin MJ, Sloan DM, Weathers FW, Keane TM, & Schnurr PP (2022). Reliable and clinically significant change in the clinician-administered PTSD Scale for DSM-5 and PTSD Checklist for DSM-5 among male veterans. *Psychological Assessment*, 34(2), 197–203. 10.1037/pas0001098 [PubMed: 34941354]
- Mickey RM, & Greenland S (1989). THE IMPACT OF CONFOUNDER SELECTION CRITERIA ON EFFECT ESTIMATION. *American Journal of Epidemiology*, 129(1), 125–137. 10.1093/oxfordjournals.aje.a115101 [PubMed: 2910056]
- Moullin JC, Ehrhart MG, & Aarons GA (2018). The Role of Leadership in Organizational Implementation and Sustainment in Service Agencies. *Research on Social Work Practice*, 28(5), 558–567. 10.1177/1049731517718361
- National Center for PTSD. (2022). PTSD Checklist for the DSM-5 (PCL-5). <https://www.ptsd.va.gov/professional/assessment/adult-sr/ptsd-checklist.asp>
- Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi J-C, Saunders LD, Beck CA, Feasby TE, & Ghali WA (2005). Coding Algorithms for Defining Comorbidities in ICD-9-CM and ICD-10 Administrative Data: *Medical Care*, 43(11), 1130–1139. 10.1097/01.mlr.0000182534.19832.83 [PubMed: 16224307]
- Roberge EM, Weinstein HR, & Bryan CJ (2022). Predicting response to cognitive processing therapy: Does trauma history matter? *Psychological Trauma: Theory, Research, Practice, and Policy*, 14(5), 871–882. 10.1037/tra0000530 [PubMed: 31738073]
- Saxon D, Firth N, & Barkham M (2017). The relationship between therapist effects and therapy delivery factors: Therapy modality, dosage, and non-completion. *Administration and Policy in Mental Health and Mental Health Services Research*, 44(5), 705–715. 10.1007/s10488-016-0750-5 [PubMed: 27424106]
- Sayer NA, Wiltsey-Stirman S, Rosen CS, Bernardy NC, Spont MR, Kehle-Forbes SM, Eftekhari A, Chard KM, & Nelson DB (2022). Investigation of Therapist Effects on Patient Engagement in Evidence-Based Psychotherapies for Posttraumatic Stress Disorder in the Veterans Health Administration. *Journal of Traumatic Stress*, 35(1), 66–77. 10.1002/jts.22679 [PubMed: 34048602]

- Schiefele A-K, Lutz W, Barkham M, Rubel J, Böhnke J, Delgadillo J, Kopta M, Schulte D, Saxon D, Nielsen SL, & Lambert MJ (2017). Reliability of therapist effects in practice-based psychotherapy research: A guide for the planning of future studies. *Administration and Policy in Mental Health and Mental Health Services Research*, 44(5), 598–613. 10.1007/s10488-016-0736-3 [PubMed: 27147427]
- Schnurr PP, Chard KM, Ruzek JI, Chow BK, Resick PA, Foa EB, Marx BP, Friedman MJ, Bovin MJ, Caudle KL, Castillo D, Curry KT, Hollifield M, Huang GD, Chee CL, Astin MC, Dickstein B, Renner K, Clancy CP, ... Shih M-C (2022). Comparison of Prolonged Exposure vs Cognitive Processing Therapy for Treatment of Posttraumatic Stress Disorder Among US Veterans: A Randomized Clinical Trial. *JAMA Network Open*, 5(1), e2136921. 10.1001/jamanetworkopen.2021.36921 [PubMed: 35044471]
- Sciarrino NA, Warnecke AJ, & Teng EJ (2020). A Systematic Review of Intensive Empirically Supported Treatments for Posttraumatic Stress Disorder. *Journal of Traumatic Stress*, 33(4), 443–454. 10.1002/jts.22556 [PubMed: 32598561]
- Shiner B, Levis M, Dufort VM, Patterson OV, Watts BV, DuVall SL, Russ CJ, & Maguen S (2022). Improvements to PTSD quality metrics with natural language processing. *Journal of Evaluation in Clinical Practice*, 28(4), 520–530. 10.1111/jep.13587 [PubMed: 34028937]
- Steenkamp MM, Litz BT, Hoge CW, & Marmar CR (2015). Psychotherapy for military-related PTSD: A review of randomized clinical trials. *JAMA*, 314(5), 489. 10.1001/jama.2015.8370 [PubMed: 26241600]
- Steenkamp MM, Litz BT, & Marmar CR (2020). First-line Psychotherapies for Military-Related PTSD. *JAMA*, 323(7), 656. 10.1001/jama.2019.20825 [PubMed: 31999301]
- Tuerlinckx F, Rijmen F, Verbeke G, & Boeck P (2006). Statistical inference in generalized linear mixed models: A review. *British Journal of Mathematical and Statistical Psychology*, 59(2), 225–255. 10.1348/000711005X79857 [PubMed: 17067411]
- Waltz J, Addis ME, Koerner K, & Jacobson NS (1993). Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. *Journal of Consulting and Clinical Psychology*, 61(4), 620–630. 10.1037/0022-006X.61.4.620 [PubMed: 8370857]
- Webb CA, DeRubeis RJ, & Barber JP (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, 78(2), 200–211. 10.1037/a0018912 [PubMed: 20350031]
- Williams NJ, Ehrhart MG, Aarons GA, Marcus SC, & Beidas RS (2018). Linking molar organizational climate and strategic implementation climate to clinicians' use of evidence-based psychotherapy techniques: Cross-sectional and lagged analyses from a 2-year observational study. *Implementation Science*, 13(1), 85. 10.1186/s13012-018-0781-2 [PubMed: 29940989]
- Williams NJ, Wolk CB, Becker-Haimes EM, & Beidas RS (2020). Testing a theory of strategic implementation leadership, implementation climate, and clinicians' use of evidence-based practice: A 5-year panel analysis. *Implementation Science*, 15(1), 10. 10.1186/s13012-020-0970-7 [PubMed: 32033575]
- Wiltsey Stirman S, La Bash H, Nelson D, Orazem R, Klein A, & Sayer NA (2022). Assessment of modifications to evidence-based psychotherapies using administrative and chart note data from the US department of veterans affairs health care system. *Frontiers in Public Health*, 10, 984505. 10.3389/fpubh.2022.984505 [PubMed: 36457312]
- Wise EA (2004). Methods for Analyzing Psychotherapy Outcomes: A Review of Clinical Significance, Reliable Change, and Recommendations for Future Directions. *Journal of Personality Assessment*, 82(1), 50–59. 10.1207/s15327752jpa8201_10 [PubMed: 14979834]
- Worley CB, Rosen CS, LoSavio ST, Aajmain ST, Stirman SW, & Sloan DM (2022). An examination of individual and organizational theory in a pilot virtual facilitated learning collaborative to implement written exposure therapy. *Psychological Services*. 10.1037/ser0000698

Public Health Significance

In routine care, some therapists were more effective than others at retaining patients in trauma-focused psychotherapy. Specifically, approximately one in four patients dropped out among therapists in the best performing quartile while almost four in five patients dropped out among therapists in the worst performing quartile. Because patients who dropped out from a trauma-focused psychotherapy generally did not demonstrate clinically meaningful improvement or recovery, identification of modifiable factors that enhance therapists' ability to reduce dropout is critical. Our findings suggest that interventions to reduce the time between sessions and enhance clinic support for delivery of trauma-focused psychotherapy have the potential to reduce the difference in dropout rates between high and low performing therapists.

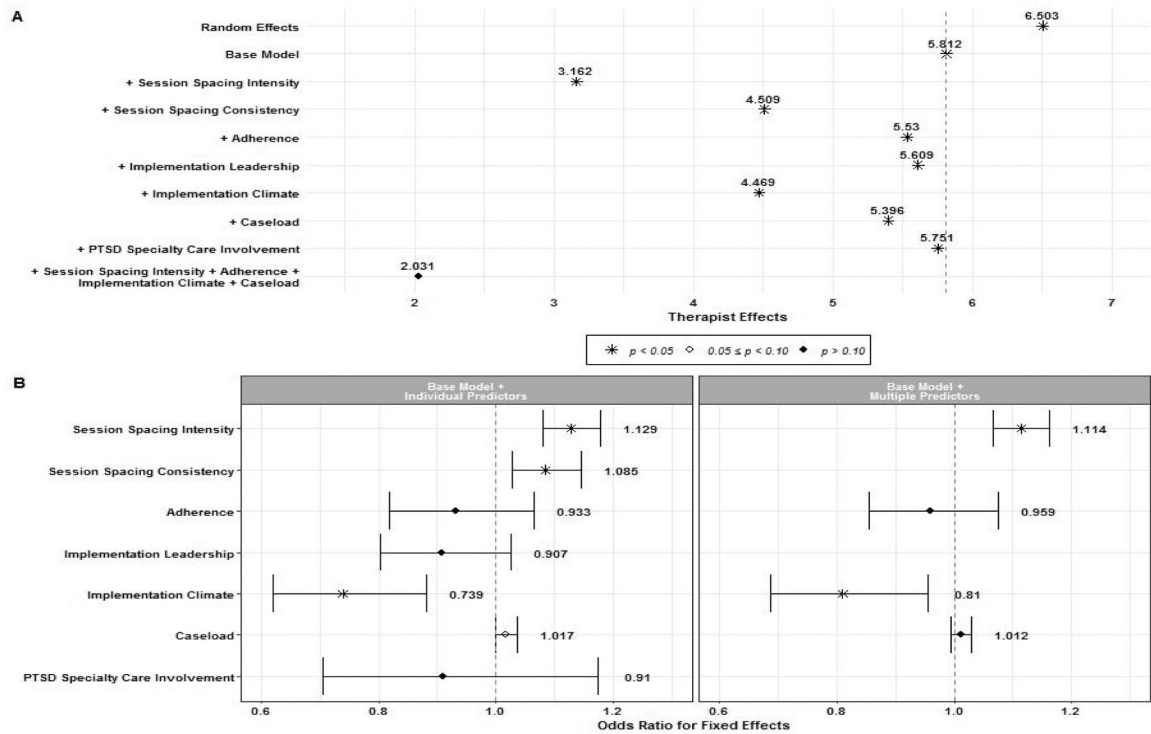


Figure 1. (A) Estimates for therapist effect variances in multilevel logistic regression models for dropout with the addition of the predictors listed on the Y-axis. The random effects model contained random effects for therapist and an intercept. The base model included patient case-mix variables and type of trauma focused psychotherapy. (B) Odds ratios with 95% confidence intervals for the fixed effects for therapist-level predictors from multilevel models for dropout. The left panel B presents fixed effects with each predictor added individually to the base model. The right panel B presents fixed effects with session spacing intensity, adherence, implementation climate scores and caseload added simultaneously to the base model.

Table 1

Characteristics of Patients in the Dropout Sample and Clinical Outcomes Subsample

	Dropout Sample	Clinical Outcomes Subsample
<i>N</i>	1,735	1,273
Age, <i>M (SD)</i>	46.69 (13.86)	46.75 (13.9)
Sex (% female)	24.7	24.2
Service era (%)		
Afghanistan or Iraq	37.8	38.3
Persian Gulf	38.2	36.9
Vietnam	13.5	13.4
Other	10.5	11.4
Race (%)		
Black	22.7	21.7
Hawaiian or Pacific Islander	1.6	1.7
Asian	1.4	1.3
Native American	0.6	0.6
White	65.8	67.2
Multiracial	1.0	1.0
missing	7.0	6.4
Ethnicity (% Hispanic ^a)	9.5	7.7
Current marital status (%)		
Married or partnered	67.7	67.4
Divorced or separated	19.9	20.1
Widowed	1.3	1.1
Never married and single	7.2	7.7
missing	3.9	3.7
Education (%)		
Less than high school	0.5	0.6
High school	19.7	18.6
Some college or trade school	30.3	30.2
College	16.2	15.6
> College	7.8	7.9
missing	25.5	25.7
Employment status (%)		
Employed	47.9	46.4
Unemployed	31.1	32.9
retired	13.5	13.4
missing	7.6	7.2
Homeless or unstable housing (%)	6.4	7.4
Index trauma (%)		
Combat	52.5	52.8
Other trauma	17.9	17.4

	Dropout Sample	Clinical Outcomes Subsample
Military sexual trauma	18.8	19.0
Other sexual trauma	3.8	3.7
Multiple sources	3.9	4.1
missing	3.2	3.0
Multiple trauma history (%)	55.4	56.5
Childhood trauma history (%)	35.6	37.2
Charlson comorbidity index, <i>M (SD)</i>	0.49 (1.04)	0.52 (1.08)
PTSD service connection (%)	66.2	65.8
High baseline psychiatric severity (%)	53.0	63.6
Number psychiatric comorbidities in prior year, <i>M (SD)</i>	2.94 (1.32)	3.01 (1.34)
Psychiatric comorbidities in prior year (%)		
Trauma- and stress-related disorders	98.33	98.35
Depressive disorders	69.5	70.9
Anxiety disorders	43.9	44.2
Alcohol use disorders	19.4	20.1
Other substance-related and addictive disorders	13.5	15.2
Bipolar and related disorders	10.2	10.7
Schizophrenia and other psychotic disorders	1.3	1.4

Note. Service era, number psychiatric comorbidities and psychiatric comorbidities in prior year were not considered for case mix adjustment in the base model due to redundancy with other variables.

^aHispanic is the ethnicity term used in administrative databases in the Veterans Health Administration.

Table 2
Therapist-Level Predictors and Odds Ratios for Dropout, Clinically Meaningful Improvement and Recovery

Variable	Dropout			Clinically Meaningful Improvement			Recovery		
	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P
Session spacing intensity	1.127	[1.086, 1.171]	< . 001	0.942	[0.902, 0.984]	.008	0.985	[0.939, 1.031]	.526
Session spacing consistency	1.088	[1.042, 1.137]	< . 001	0.968	[0.918, 1.020]	.226	1.001	[0.945, 1.059]	.980
Adherence	0.991	[0.982, 1.001]	.068	1.003	[0.990, 1.016]	.695	1.01	[0.995, 1.024]	.196
Implementation leadership	0.911	[0.830, 1.000]	.049	1.121	[1.003, 1.253]	.045	1.097	[0.972, 1.239]	.136
Implementation climate	0.744	[0.647, 0.854]	< . 001	1.054	[0.897, 1.239]	.524	1.115	[0.936, 1.331]	.224
Caseload ^a	1.163	[1.032, 1.307]	.013	0.934	[0.810, 1.074]	.357	0.961	[0.817, 1.118]	.583
PTSD specialty care involvement	0.899	[0.742, 1.09]	.281	1.128	[0.899, 1.416]	.299	0.886	[0.694, 1.132]	.330

Note. PTSD = posttraumatic stress disorder.

^aPer eight additional patients per month. Bold indicates statistically significant effects

Table 3

Estimated Variance Components for Random Effects and Odds Ratios for Fixed Effects for Base Model for Dropout (N = 1,735)

Random Effects	Estimate	SE	Proportion of Variance	P
Therapist, σ^2_T	0.203	0.071	5.812%	< 0.001
Fixed Effects				
	OR	[95% CI]	p	
PE (ref.: CPT)	1.276	[0.991, 1.619]	.059	
Age (mean centered)	0.986	[0.978, 0.995]	.002	
Race (ref.: all other)				
Black	1.275	[0.992, 1.638]	.057	
Multiracial	2.122	[0.695, 6.480]	.187	
Ethnicity (ref.: Hispanic)				
not Hispanic	0.749	[0.522, 1.074]	.116	
missing	0.682	[0.376, 1.236]	.207	
Education (ref.: college)				
High school	1.530	[1.126, 2.079]	.007	
Some college or trade school	1.331	[1.012, 1.750]	.041	
missing	1.431	[1.070, 1.915]	.016	
Employment (ref.: all other)				
Retired	0.790	[0.557, 1.122]	.188	
Trauma history (ref.: no multiple trauma)				
Multiple trauma	1.155	[0.936, 1.426]	.180	
Index trauma (ref.: all other)				
Other or missing	1.159	[0.904, 1.485]	.245	
Pandemic period (ref.: pre-pandemic)				
Early pandemic	2.003	[1.521, 2.637]	< 0.001	
During pandemic	1.124	[0.886, 1.425]	.330	

Note. PE = prolonged exposure; CPT = cognitive processing therapy. Race reference group = White, Native American, Asian, Hawaiian, and missing. Employment reference group = employed, unemployed and missing. Index trauma reference group = combat military sexual trauma, other sexual trauma, and multiple sources. Bold indicates statistically significant effects.

Table 4

Estimated Variance Components for Random Effects and Odds Ratios for Fixed Effects for Base Models for Clinically Meaningful Improvement and Recovery (N = 1,273)

Random Effects	Clinically Meaningful Improvement			Recovery		
	Estimate	SE	Proportion of Variance	Estimate	SE	Proportion of Variance
Therapist, σ^2_T	0.068	0.112	2.032%	0.00	0.062	0.00%

	Fixed Effects			
	OR	[95% CI]	OR	[95% CI]
Age (mean centered)	--		1.001	[0.990, 1.011]
Race (ref.: all other)				
Native American, Hawaiian, Asian	0.571	[0.296, 1.101]	0.343	[0.149, 0.789] *
Black	--		0.665	[0.484, 0.915] *
Employment (ref.: all other)				
Retired	1.697	[1.213, 2.373] **	1.964	[1.306, 2.953] **
Childhood trauma	1.234	[0.968, 1.574]	--	
Index trauma (ref.: all other)				
Sexual trauma	1.448	[0.779, 2.690]	--	
PTSD service connection	0.823	[0.646, 1.049]	--	
High baseline psychiatric severity	1.596	[1.254, 2.031] ***	0.438	[0.341, 0.563] ***
Pandemic period (ref.: pre- and during pandemic)				
Early pandemic	--		0.809	[0.589, 1.110]
Psychiatric disorders other than PTSD	--		1.424	[0.711, 2.851]

Note. PTSD = posttraumatic stress disorder. Race reference group for clinically meaningful improvement = White, Black, Multiracial, and missing; for recovery = White, Multiracial and missing. Employment reference group = employed, unemployed and missing. Index trauma reference group = combat, other trauma, MST, multiple trauma and missing. Bold indicates statistically significant effects.

* $p < .05$.

** $p < .01$.

*** $p < .001$.