Evaluation of a standardized posttraumatic stress disorder treatment framework in routine mental health care

Effectiveness and predictors of treatment outcome in a consecutive sample

Scharff, Frederik B.; Lau, Marianne E.; Gondan, Matthias; Folke, Sofie; Arendt, Ida Marie T.P.; Moeller, Stine B.

Published in:
Journal of Traumatic Stress

DOI:
10.1002/jts.22790

Publication date:
2022

Document version:
Accepted manuscript

Citation for published version (APA):
Evaluation of a standardized posttraumatic stress disorder treatment framework in routine mental health care: Effectiveness and predictors of treatment outcome in a consecutive sample

Frederik B. Scharff\(^1\), Marianne E. Lau\(^1\), Matthias Gondan\(^2\), Sofie Folke\(^3\), Ida-Marie T. P. Arendt\(^4\), & Stine B. Moeller\(^1,5\)

\(^1\)Stolpegaard Psychotherapy Centre, Mental Health Services, Capital Region of Denmark, Copenhagen, Denmark

\(^2\)Department of Psychology, University of Innsbruck, Innsbruck, Austria

\(^3\)Danish Veterans Centre, Danish Ministry of Defense, Ringsted, Denmark

\(^4\)Clinic for Affective Disorders, Mental Health Services, Capital Region of Denmark, Copenhagen, Denmark

\(^5\)Department of Psychology, University of Southern Denmark, Odense, Denmark

AUTHOR NOTE

Financial support was provided by The Danish Victims Fund (18-610-00018). The authors have no conflicts of interest to declare.

Correspondence concerning this article should be addressed to Frederik Bernt Scharff, Aertemarken 15A, 2860, Soeborg, Denmark, f.b.scharff@gmail.com

ABSTRACT

The primary aim of the present study was to evaluate the effectiveness of standardized care package (CP) treatment for posttraumatic stress disorder (PTSD) in a Danish sample of adult psychiatric outpatients (\(N = 948\)). Secondary aims were to identify baseline predictors of treatment outcome and investigate between-group differences in outcome with regard to sex and treatment modality (i.e., group vs. individual therapy). The naturalistic, nonrandomized study followed a pre–post design. Patient data from five psychiatric outpatient clinics were collected between March 2011 and November 2017. Data were drawn from self-report questionnaires (i.e., SCL-90-R, WHO-5, BHS) and therapist-reported measures (i.e., GAF-S, GAF-F) administered at baseline and posttreatment. Between-group effects for sex and therapy modality (group vs. individual) were analyzed using analyses of variance, and possible predictors of outcome were selected through LASSO regression and analyzed via
hierarchical regression. Pre–post effects were small to moderate, $d_s = 0.39–0.69$. No differences emerged regarding treatment modality, but women had significantly better outcomes than men. Aside from sex, only baseline symptom severity predicted outcomes. The effectiveness of the CP treatment was generally limited, indicating the need to implement improved therapeutic practices, such as the use of evidence-based treatments and to provide better training to mental health clinicians. The findings underscore the need for further comparisons of group and individual treatment modalities using evidence-based therapies as well as the need to investigate factors that may affect treatment outcome.

Evaluation of a standardized posttraumatic stress disorder treatment framework in routine mental health care: Effectiveness and predictors of treatment outcome in a consecutive sample

Posttraumatic stress disorder (PTSD) is a psychiatric disorder that may develop after exposure to one or more traumatic events. PTSD is often a protracted disorder with a high rate of chronicity, and approximately 40% of individuals with PTSD still exhibit symptoms 10 years after the initial onset of the disorder (Kessler et al., 1995). In addition, PTSD patients can experience a range of negative secondary effects, including significantly increased rates of somatic symptoms and medical conditions (Pacella et al., 2013), elevated suicide risk, comorbid anxiety and depressive disorders, substance abuse (Pietrzak et al., 2011), and severe sleep problems (Lamarche & De Koninck, 2007). Hence, improving treatments for PTSD patients should remain a priority for clinicians and researchers (Bradley et al., 2005).

The estimated lifetime prevalence of PTSD varies between countries, ranging from 2% to 12% (Atwoli et al., 2015). This variability is also seen between sexes, as women are approximately 2–3 times more likely to develop PTSD than men after trauma exposure (Atwoli et al., 2015). This difference may be due to differences in the types of traumatic events to which men and women are exposed (Olff et al., 2007); however, several studies have found that some sex differences persist after controlling for the influence of trauma type (Ditlevsen & Elklit, 2012).

Evidence of treatment effects

Several individual psychosocial interventions have been shown to be effective for the treatment of PTSD. The effect sizes for these interventions range from moderate to large (Benish et al., 2008; Bisson et al., 2013; Cusack et al., 2016). The first-choice interventions recommended by most meta-analyses and treatment guidelines are trauma-focused cognitive behavioral therapy (TF-CBT; Cohen et al., 2012), cognitive processing therapy (CPT)
(Resick et al., 2016), prolonged exposure (PE; Foa et al., 2019), and eye movement desensitization and reprocessing (EMDR; Shapiro & Laliotis, 2011). These interventions currently have the strongest evidence base (Bisson et al., 2013; National Institute for Health and Care Excellence (NICE), 2018), although the authors of two meta-analyses concluded that there was insufficient evidence to make conclusions regarding treatment superiority (Benish et al., 2008; Cusack et al., 2016).

In the last decade, only one meta-analysis has compared the efficacy of group-based and individual treatment for PTSD. Schwartze et al. (2019) analyzed the efficacy of group therapy for PTSD compared with an active control group consisting of treatment as usual (TAU) or nonspecific interventions. The findings indicated that group-based interventions may be a reasonably effective intervention for PTSD, as the authors did not observe any significant differences between group versus individual treatment. However, in a recent RCT that was not included in the meta-analysis, Lamp et al. (2019) found that individual CPT was significantly more effective than group CPT. These findings replicate prior evidence of the superiority of individual CPT over group CPT for PTSD (Jeffreys et al., 2014; Resick et al., 2017).

Moderators of treatment modality outcome

Modest evidence exists regarding the factors that moderate the relative effect of group versus individual PTSD treatment. To our knowledge, only one study has directly analyzed moderators of group versus individual treatment modality. In this study, Resick et al. (2017) found that younger patients showed significantly better outcomes in individual versus group treatment. This indicates that patient characteristics can affect how much patients benefit from either group or individual therapy. However, more research is needed to elucidate which patient characteristics influence the effect of group and individual therapy.

Predictors of treatment dropout

Multiple studies have found evidence of the difficulties in retaining patients in PTSD treatment. Meta-analytic studies have reported dropout rates ranging from 0% to 54% (Bradley et al., 2005; Schottenbauer et al., 2008). When comparing dropout rates for trauma-focused versus non–trauma-focused therapy, Imel et al. (2013) found no differences between the two but reported a mean dropout rate of 18%, with substantial variation between studies. The authors also found a 12% higher dropout rate for group treatment compared to individual treatment. This effect of group therapy on dropout has since been replicated (Berke et al., 2019). In addition, both age, trauma type, and therapy type have been found to predict PTSD treatment dropout (Berke et al., 2019; Kehle-Forbes et al., 2016).

Predictors of treatment outcome

The current PTSD treatment literature includes mixed findings regarding predictors of PTSD treatment outcome, with significant effects observed for a broad range of demographic,
therapeutic, patient-centered, and trauma-specific variables (Ehlers et al., 2013; Fletcher et al., 2017; Imel et al., 2013; Minnen et al., 2002). However, few variables have been consistently replicated across studies and interventions. Thus, to examine whether non-intervention-specific predictors of PTSD treatment outcomes exist, there is a need for research that investigates the potential predictors of treatment across a range of interventions.

The current study

In 2010, Care Package treatments (CPs) were implemented in Denmark for all nonpsychotic adult psychiatric outpatients (Association of Danish Regions, 2011). The CP treatments are standardized resource-restricted courses of treatment, each tailored to a specific diagnosis. The CPs provide a quantitative framework that standardizes the treatment elements offered to all patients as well as the amount of clinician resources allocated to these elements. Treatment elements in the CPs include but are not limited to diagnostic assessments, psychotherapy, psychoeducation, counseling on complicating conditions, and psychopharmacological consultations (see Table 1 for a full summary). However, the CP framework does not mandate specific psychotherapeutic or pharmacological interventions, although it recommends the use of evidence-based interventions. In addition, the CPs do not contain provisions for therapist training or assessments of therapist fidelity, allowing each outpatient clinic to manage these processes as they see fit. Therefore, a variety of psychotherapeutic interventions are provided under the CP framework. This context provides an opportunity to study patient and treatment variables that affect PTSD treatment outcome in a naturalistic psychiatric setting across and psychotherapeutic interventions with similarly allocated clinician resources. Therefore, the present study aims were to (a) evaluate the effectiveness of the standardized and resource-restricted CP PTSD treatment for adult psychiatric outpatients with PTSD, (b) investigate predictors of treatment outcome and dropout, and (c) investigate differences in treatment outcome and dropout between therapy modalities (i.e., individual vs. group therapy).

METHOD

Participants and procedure

The current study used a naturalistic pre–post design and was based on data collected at five outpatient clinics in the Mental Health Sector in the Capital Region of Denmark (MHS-CRD). Patients were assessed at baseline and the end of treatment (EOT). Data were collected using an internet-based monitoring system (IMS) developed by the MHS-CRD. The IMS was developed as a tool to collect patient data for use by clinicians, and all included measures were selected before the inception of this study based on their utility for and prior use in the Danish psychiatric system.

This article is protected by copyright. All rights reserved.
Adult patients with PTSD were recruited consecutively from five public psychiatric outpatient clinics in the MHS-CRD. Inclusion was ongoing over a 6.5-year period from March 2011 to November 2017. The sample consisted of 948 treatment-seeking adult patients (i.e., 18 years of age and older) referred for psychiatric treatment by their general practitioner or clinicians from the MHS-CRD (see Table 2 for demographic data). To be eligible for the study, patients were required to (a) have PTSD according to the diagnostic criteria in the International Classification of Diseases (10th rev.; ICD-10; World Health Organization [WHO], 2004) and (b) be allocated to the current PTSD CP-treatment. Exclusion criteria for outpatient treatment were intellectual disability, current psychosis, acute suicidal risk, and ongoing substance abuse.

At baseline, all patients were assessed through two semistructured psychiatric diagnostic interviews at the outpatient clinic, one conducted by a clinical psychologist and one by a psychiatrist. During the interview, psychiatric diagnoses according to ICD-10 criteria (WHO, 2004) were established. Demographic data were collected at baseline (see Table 2). In addition, general clinical data and Global Assessment of Functioning (GAF) scores (Pedersen & Karterud, 2012) were collected by the interviewing clinicians. All patient data were registered in the IMS by the clinicians. At the end of CP treatment, treatment-related data, including treatment modality, number of sessions, length of sessions, and treatment site were registered in the IMS along with a GAF score. The GAF was scored by the clinician overseeing the patient’s treatment, who was often a different clinician from the baseline interviewer. Patients completed a battery of self-report questionnaires at pre- and posttreatment. After the baseline assessment, patients were allocated to either group or individual therapy for PTSD based on clinician evaluation and available resources. The study was approved by the Danish National Scientific Ethics Committee.

The treatment administered in the present study complied with official Danish recommendations for PTSD CP-treatment (Association of Danish Regions, 2017). From 2011 to 2012, patients referred to PTSD treatment were allocated to adjustment disorder (AD) CP treatment (Association of Danish Regions, 2012), as the PTSD CP treatment was still under development (see Table 1 for an overview of the CPs). The contents of the CP treatments for AD and PTSD were similar, with PTSD patients receiving 29 hr of treatment and AD patients receiving 27 hr. Hours were calculated as therapist hours per patient, and, as such, there was a marked difference in hours of therapy attendance between group and individual modalities. Patients in individual therapy received 15 hr of therapy; patients in group therapy received 60 hr, which was spent in groups of eight patients led by two therapists, equaling 15 therapist hours per patient. The patients included in the present sample received a wide range of specific psychotherapeutic interventions, with 28 different treatments identified throughout the sample. Treatment was managed by an interdisciplinary team of psychologists, psychiatrists, physiotherapists, psychiatric nurses, and social counselors.

This article is protected by copyright. All rights reserved.
Measures

**Psychiatric symptoms**

The Symptom Checklist-90-Revised (SCL-90-R; Derogatis, 1994) is a 90-item, self-report measure that is used to assess general psychiatric symptoms. Respondents rate responses on a 5-point Likert scale ranging from 0 (not at all) to 4 (extremely), with higher scores indicating more severe symptoms. A total score, called the General Severity Index (GSI), is calculated as the mean of all items. In the present sample, Cronbach’s alpha for the GSI reliability was .96.

**PTSD symptoms**

The PTSD Checklist–Civilian Version (PCL–C; Weathers et al., 2013) was only added to the assessment battery late in the data collection. Thus, we did not have data on this measure for the full sample. As such, PTSD symptom severity was measured using the SCL-90-R Crime-Related PTSD (CR-PTSD; Saunders et al., 1990) subscale. The CR-PTSD is a unidimensional measure designed to assess PTSD symptoms based on the scoring of 28 SCL-90-R items. The CR-PTSD has shown excellent internal consistency (Cronbach’s α = .93) as well as good validity for women (r = .59); however, the same has not been observed for men (Carlozzi & Long, 2008). To ascertain the convergent validity of the CR-PTSD for both sexes in our sample, we performed a correlational analysis between baseline PCL-C scores and CR-PTSD scores for both women (n = 248) and men (n = 55). We also performed partial correlations between PCL-C EOT scores and CR-PTSD EOT scores, controlling for baseline scores for women (n = 72) and men (n = 16). At baseline, the CR-PTSD showed acceptable convergence with the PCL-C at both baseline and EOT for women (baseline: r = .66, EOT: r = .88) and men (baseline: r = .78, EOT: r = .95). This indicates that the CR-PTSD was an acceptable measure of PTSD symptom severity for the current study. Internal consistency for the CR-PTSD was excellent in the present sample, Cronbach’s α = .91.

The PCL-C (Weathers et al., 1993), a widely used and validated measure, was administered to a small subsample at baseline (n = 303) and EOT (n = 88), as it was not available in the IMS at the start of the data collection. The PCL-C is a 17-item, self-report questionnaire that is used to assess PTSD symptoms according to the criteria in the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM-IV*; American Psychiatric Association, 1994). Respondents score items on a 5-point Likert scale ranging from 1 (not at all) to 5 (extremely), and total PTSD symptom severity scores are calculated by summing all item scores (range: 17–85). In the present sample, the PCL-C demonstrated good internal consistency, Cronbach’s α = .84.

**Anger and depression**
Anger was measured using the SCL-6 Hostility subscale which is based on the scores of six SCL-90-R items. The SCL-6 has demonstrated acceptable construct validity (Olsen et al., 2004). The internal consistency was good in the present sample, Cronbach’s α = .83. Depression was measured using the SCL-D13, which is based on the scores of 13 SCL-90-R items. The SCL-D13 demonstrated acceptable construct validity (Olsen et al., 2004). In the current sample, Cronbach’s alpha was .87.

**Psychological well-being**

The WHO-5 Well-Being Index (WHO-5; Bech, 2004) is a five-item, self-report measure of psychological well-being. Respondents score each item on a 6-point Likert scale ranging from 0 (at no time) to 5 (all the time) with respect to their experiences over the past 2 weeks. The total score is calculated by summing all scores and multiplying by 4, with higher scores indicating higher levels of well-being (range: 0–100). The measure has shown excellent validity (McDowell, 2010). In the present sample, the WHO-5 demonstrated good internal consistency, Cronbach’s α = .85.

**Hopelessness and suicide risk**

The Beck Hopelessness Scale (BHS; Beck et al., 1974) is a 20-item, self-report measure used to screen for hopelessness and suicide risk. Items are rated as “true” (1) or “false” (0), and scores are calculated by summing the item scores, with higher scores indicating higher levels of hopelessness (range: 0–20). The BHS has historically demonstrated good internal reliability (e.g., Cronbach’s α = .83; Steed, 2001) and validity (Beck et al., 1974).

**Global functioning**

The GAF (Pedersen, & Karterud, 2012) is a measure of overall psychosocial impairment as experienced over the previous month. The measure is scored on two dimensions based on a clinician’s evaluation of the patient’s symptoms and social and occupational functioning. The two scores range from 1 to 100, with lower scores GAF-S scores indicating higher levels of symptom severity (GAF-S) and lower GAF-F scores indicating more impaired social and occupational functioning. The GAF has demonstrated satisfactory validity (Pedersen & Karterud, 2012); however, results concerning the reliability of the GAF have been mixed. Some studies have found the GAF to have insufficient reliability for use in routine clinical care (Groffenboer et al., 2010; Vatnaland et al., 2007), whereas others have found the GAF to have acceptable reliability (Pedersen et al., 2007; Sonesson et al., 2010). Findings from a study by Söderbern et al. (2005) may bridge this discrepancy, as the authors found that although the GAF did not have sufficient reliability to measure change at the individual level, it was sufficient to measure change at the group level. Thus, we decided to include the GAF data in this study.
**Remission**

PTSD diagnosis remission was defined as scoring under the cutoff for caseness (i.e., 0.89) on the CR-PTSD SCL-90-R scale at EOT (Saunders et al., 1990). Remission of general psychopathology was defined as scoring under the cutoff for caseness on the GSI (i.e., 1.08 for women, 0.87 for men; Olsen et al., 2006).

**Treatment attendance and dropout**

Attendance was measured as the total number of hours patients attended therapy during the treatment period compared with the standard for individual (15 hr) and group (60 hr, groups of eight patients and two therapists) treatment. Dropout was registered by clinicians when patients stopped attending therapy without prior agreement with the clinician before attending at least 50% of the planned sessions.

**Data analysis**

Data were analyzed using R (Version 3.6.1; R core team, 2019). Multiple imputation was conducted to handle missing data at baseline and EOT (Sinharay et al., 2001; see Figure 1 for an overview of missingness). This procedure assumes that data are missing at random (Sinharay et al., 2001). A total of 20 imputed data sets were derived. In each data set, missing values were replaced by values generated using predictive mean matching (PMM), which picked the observed value closest to the value generated by a linear regression model. Pre–post outcomes were examined using a Welch paired-samples t test. Symptom reduction was expressed as the average pre–post change. Standardized effects (i.e., Cohen’s $d$; Cohen, 2013) were calculated for all outcomes.

Between-group differences in outcomes for sex (i.e., men and women) and treatment modality (i.e., individual and group therapy) were calculated using analyses of covariance (ANCOVAs) with baseline severity as a covariate. Differences in demographic and clinical variables for the two groups were summarized using standard descriptive statistics. Due to the large number of different treatments (i.e., 28 different interventions), we determined that testing the effect of specific treatments against each other would not provide meaningful findings.

The potential predictors of outcome were selected using a least absolute shrinkage and selection operator regression analysis (LASSO; Friedman et al., 2010), which reduces the overfitting of models in broad datasets through regularization. The 31 variables included in the LASSO analysis were: age at intake, therapeutic modality (group and individual therapy), earlier psychiatric treatment, symptom duration, treatment attendance, number of psychiatrist consultations, educational attainment, employment status, type of employment, number of children, number of children living at home, marital status, cohabitation status, comorbid...
diagnosis, antidepressant use, use of other psychopharmacological drugs, drug abuse, alcohol abuse, clinician-rated suicidality, clinician-rated aggressiveness, SCL-90 Hostility score, SCL-90 Depression score, SCL-90 GSI score, GAF-F score, GAF-S score, CR-PTSD score, BHS score, WHO-5 score, childhood sexual abuse history, childhood placement in foster care, and treatment site. However, due to the multiple imputation procedure performed on the dataset, we were unable to use a standard LASSO approach, as the set of candidates selected would vary across the imputed datasets. Therefore, we used the LASSO method on the nonimputed dataset for variable selection and analyzed the selected candidates in the imputed data through hierarchical regression models.

The same procedure was used to identify predictors of dropout. We used a binomial version of the LASSO regression (Friedman et al., 2010) and analyzed the selected predictors through logistic regression. Sex and childhood sexual abuse history were analyzed as potential moderators through a regression model with an interaction term included.

To test whether the study was adequately powered, we conducted a post hoc power analysis. Using the following formula, we calculated the average difference in outcomes from pre- to posttreatment that would be observable with 90% power:

$$diff = SD \cdot \sqrt{(4 \cdot (z_{1-\alpha/2} + z_{1-\beta})^2)/N}$$

For the primary outcome on the CR-PTSD scale, the observable difference was $0.70 \cdot \sqrt{42/719} = 0.17$, and the observed difference in our study was 0.33, indicating that we had adequate power to detect differences in the primary outcome. Further analysis showed that for the remaining outcomes, observable versus observed differences were similarly acceptable, GSI: 0.15 versus 0.28; GAF-S: 1.7 versus 6.6; GAF-F: 3.6 versus 6.5; WHO-5 = 4.0 versus 7.4. The moderator analysis was tested using the same formula, using the residual standard deviation of the analysis. The calculation showed that for the GSI and CR-PTSD, there was sufficient power to detect differences of 0.04 between group and individual therapy for men and women, which was adequate compared to the observed differences.

RESULTS

Baseline data and patient flow

In total, data from 948 patients diagnosed with PTSD in five Danish public outpatient clinics were included. From this sample, full baseline data were collected for 719 patients, treatment data were collected for 633 patients, dropout data were collected from 730 patients, and full data from baseline to EOT were collected from 251 patients (see Figure 1).
The mean patient age was 38 years ($SD = 11.4$), and 76.8 were women. Of participants, 79.8% had received earlier psychotherapeutic treatment (see Table 2 for further descriptive, baseline symptom, and sociodemographic data). Regarding therapeutic modality, 45.8% of patients were assigned to individual therapy and 54.2% to group therapy. The sex distribution between group and individual therapy was skewed: 62.1% ($n = 90$) of men were allocated to individual therapy compared to 39.9% ($n = 195$). At intake, 651 (90.5%) of 719 patients with baseline data on the CR-PTSD scale scored over the clinical cutoff point of 0.89, indicating PTSD caseness (Saunders et al., 1990). The frequency of caseness at baseline did not differ by sex or therapeutic modality. No differences in outcome were found between the patients that scored above the cutoff for caseness and those who did not. Baseline means scores on the five outcome measures (i.e., CR-PTSD, GSI, GAF, and WHO-5) did not differ significantly between the sexes. Regarding treatment modality, baseline mean scores did not differ significantly on the GAF-F, GAF-S, or WHO-5 variables between patients who engaged in individual versus group therapy; however, scores differed significantly on the GSI and CR-PTSD scales, $p = .008$. Patients allocated to individual therapy (GSI: $M = 1.81$, $SD = 0.64$; CR-PTSD: $M = 1.94$, $SD = 0.70$) had baseline scores indicating more severe psychiatric and PTSD symptoms than patients allocated to group therapy (GSI: $M = 1.69$, $SD = 0.61$; CR-PTSD: $M = 1.79$, $SD = 0.66$). We found no differences in baseline scores between dropouts and treatment completers.

Effectiveness of the time-restricted PTSD CP treatment

Among all participants ($N = 948$), PTSD CP treatment led to a significant improvement between pre- and posttreatment on all outcome variables (See Table 3). There was a significant improvement from baseline to EOT in mean GSI score, $d = 0.42$; CR-PTSD score, $d = 0.45$; GAF-F score, $d = 0.69$; GAF-S score, $d = 0.48$; and WHO-5 score, $d = 0.39$. Thus, the analyses indicated that patients experienced small-to-moderate improvements across all outcomes. For the CR-PTSD scale, 18.8% of patients who scored above the caseness cutoff at baseline experienced remission; for the GSI, the remission rate was 22.2%.

Baseline data as predictors and moderators of treatment outcome and dropout

The baseline variables selected by the LASSO model in the unimputed dataset ($n = 251$; see Table 4 for predictor variables and coefficients selected by the LASSO for treatment outcomes and dropout) were analyzed through hierarchical regression in the imputed dataset ($n = 948$). For most outcomes, baseline score for the investigated variable was the only meaningful predictor of variance. Including the additional variables selected by the LASSO analysis (see Table 4) to the regression models did not increase the explanatory value of the models in most cases. CR-PTSD scores at EOT were significantly predicted by CR-PTSD scores at baseline, $B = 0.71$, $t(53.5) = 13.40$, $p < .001$, $R^2 = .38$, indicating that 38% of the variance in CR-PTSD score was explained by the model. EOT GSI and WHO-5 scores were

---

This article is protected by copyright. All rights reserved.
also predicted by their corresponding baseline scores, GSI: $B = 0.69$, $t(50.2) = 13.64$, $p < .001$, $R^2 = .38$; WHO-5: $B = 0.432$, $t(48.6) = 5.96$, $p < .001$, $R^2 = .26$. Only EOT GAF scores predicted by variables other than baseline GAF scores. Baseline GAF-S scores predicted EOT GAF-S scores, $B = 0.30$, $t(144.4) = 5.80$, $p < .001$, $R^2 = .09$, but the addition of baseline WHO-5 score to the model, $B = 0.15$, $t(85.5) = 5.17$, $p < .001$, increased the explanatory value of the model, $\Delta R^2 = .06$. Baseline GAF-F score did not significantly predict EOT GAF-F score, $p = .132$. The only meaningful and significant predictor of EOT GAF-F score was baseline WHO-5 score, $b = 0.17$, $t(75.4) = 7.25$, $p < .001$, $R^2 = .10$.

The results of ANCOVAs ($n = 948$) conducted to investigate the role of sex as a predictor indicated that there was a significant effect of sex on multiple outcomes. For the CR-PTSD, $F(2, 76.6) = 3.68$, $p < .001$, a post hoc analysis showed that men ($M = 1.78$, $SD = 0.81$, 95% CI [1.61, 1.94]) scored significantly higher than women at EOT ($M = 1.46$, $SD = 0.76$, 95% CI [1.38, 1.56]), $p < .001$, after controlling for baseline CR-PTSD score, indicating more severe PTSD symptoms for men compared to women. For the GSI, $F(2, 44.1) = 0.22$, $p = .004$, men ($M = 1.65$, $SD = 0.72$, 95% CI [1.52, 1.79]) also had significantly more severe symptoms than women ($M = 1.40$, $SD = 0.68$, 95% CI [1.34, 1.50]) at EOT, $p = .002$. Thus, sex emerged as a predictor of treatment outcome with regard to these variables. In the regression analysis, none of the variables selected by the LASSO as potential predictors of dropout (see Table 4) significantly predicted dropout. We analyzed the interaction between sex and therapy modality, which showed that men engaged in individual treatment scored significantly higher than women on both the CR-PTSD ($M = 1.77$, $SD = 0.81$ vs. $M = 1.47$, $SD = 0.79$) and the GSI ($M = 1.61$, $SD = 0.73$ vs. $M = 1.41$, $SD = 0.71$) after controlling for baseline scores. Childhood sexual abuse history did not significantly moderate the outcome of treatment modality on any measures.

Differences in treatment outcome and dropout between therapy modalities

An analysis of correlations between therapy modalities and treatment outcomes ($n = 948$) demonstrated a small but significant effect of modality on GAF-F score, $F(2, 494) = 2.09$, $p = .033$. A post hoc analysis found that patients in group therapy ($M = 56.5$, $SD = 10.3$) scored significantly higher on the GAF-F than the patients in individual therapy ($M = 54.7$, $SD = 7.9$) after controlling for baseline scores, $p = .024$. The dropout rate for the unimputed sample ($n = 635$) was 25.4%. We did not find any significant differences in dropout rates between group and individual therapy.

DISCUSSION

The present study indicates that patients demonstrated significant improvements across both patient- and clinician-reported outcomes. Cohen’s $d$ effect sizes ranged from 0.39 to 0.69, which represent small-to-medium effects. The PTSD remission rate for the sample was
18.8%; this constitutes a relatively small treatment effect compared to the current evidence in the PTSD treatment literature, as existing studies on PTSD treatment in routine clinical care have reported large effect sizes ranging from 0.87 to 1.63 (Eftekhari et al., 2013; Ehlers et al., 2013). Meta-analyses of randomized controlled trials of PTSD treatments have reported Cohen’s $d$ effect sizes ranging from 1.14 to 1.66, as well as an average remission rate of 56% (Bradley et al., 2005; Cusack et al., 2016). As a result, while also considering potential confounding by regression to the mean, it is questionable whether the observed mean score improvements in the present study constitute meaningful clinical change.

Several factors may have contributed to the modest outcome. The investigated CP treatment was relatively heterogeneous, with a wide variety of psychotherapeutic interventions provided. The treatments spanned from first-line evidence-based interventions for PTSD to less evidence-based interventions or locally developed interventions. Due to the naturalistic nature of the data collection for this study, wherein the type of intervention provided was registered solely by the attending clinicians, 28 different types of psychotherapy were registered. In addition, no data were collected on therapist adherence to or use of treatment manuals, therapist training and supervision, or the actual content of the different interventions.

The heterogeneity of interventions may have reduced the treatment effect in multiple ways. First, it may reflect a tendency for clinics to have allowed therapists to practice their preferred interventions instead of implementing systematic training in evidence-based interventions. Second, several studies have found evidence of the superiority of specific trauma-focused interventions (Bisson et al., 2013; NICE, 2018). Thus, some of the treatments provided may have been inferior to others, reducing the average effect of the treatments. Based on our results and existing treatment guidelines, it is possible that implementing fewer evidence-based treatment options might increase treatment effects. A narrower range of treatment options would also allow for the implementation of standardized therapist training and supervision across the MHS-CRD, which could further increase the effectiveness of treatment.

The modest effects of the treatment may also be related to the fact that the sample was characterized by at least some degree of treatment resistance. Most participants (approximately 80%) had received prior psychotherapeutic interventions, and over half of the sample had experienced symptoms for more than 5 years (see Table 1). Thus, it may be unreasonable to expect high additional effects of further psychotherapeutic treatments. In addition, the sample reported a high level of general psychopathology, with a baseline mean SCL-90-R GSI score of 1.74; notably, Leichsenring et al. (2020) defined a mean score of 1.37 as indicative of severe symptoms. In addition, the average participant baseline CR-PTSD score was 1.86 compared to the caseness cutoff of 0.89 (Saunders et al., 1990). The
high level of symptom severity in the sample may have affected the outcome negatively, as baseline symptom severity has been shown to affect treatment outcome (Galovski et al., 2013; Minnen et al., 2002).

We found few differences in outcomes between group and individual therapy. Across all outcomes, the only observed difference was a small but significant difference on the GAF-F whereby patients in group therapy showed marginally larger degrees of improvement than patients in individual therapy. This result is in line with theories about the added social benefits of group therapy for PTSD (Schwartze et al., 2019). However, it is unclear whether these effects are clinically relevant due to the small effect size, general multiplicity of tests used, and the fact that the results regarding the interrater reliability of the GAF used in public health settings have been mixed, ranging from poor to good (Söderberg et al., 2005).

The lack of differences in outcome with regard to treatment modality is interesting, as it adds to a growing body of research on the efficacy of group therapy for PTSD compared to individual therapy. Schwartze et al. (2019) found that existing evidence was not sufficient to conclude whether group therapy for PTSD was comparable to bona fide individual therapy for PTSD. This is relevant because researchers and clinicians have proposed several different potential benefits of group therapy, not least of which is increased cost-effectiveness (Schwertze et al., 2019).

There are, however, several caveats regarding this finding. The small effect sizes and heterogeneity of treatments used may dilute real differences between treatment modalities to a degree where they are not statistically significant. In addition, the patients that were allocated to group therapy had significantly lower levels of symptom severity at baseline and were predominantly women. We controlled for this difference statistically, but it exemplifies one of the inherent problems of a nonrandomized study, as additional bias may have affected allocation to group or individual treatment. Patients with more severe psychopathology may have been systematically allocated to individual treatment. This potential for skewness in the allocation compromises direct comparison. Thus, further investigation is needed to ascertain whether group and individual treatment for PTSD are equally effective in a public health setting.

We found that the only significant predictor for most of the outcome measures was the corresponding baseline score. This replicates earlier findings (Fletcher et al., 2017; Minnen et al., 2002). Additionally, it underlines the lack of consistent results on general predictors of PTSD therapy outcomes, as previously described. It is possible that the search for predictors needs to be approached from a different direction. The lack of consistency in results when examining patient-level or trauma-specific variables may indicate that a shift of focus to a different type of variables as outcome predictors for PTSD is needed: Investigating therapy-
specific variables, mechanisms of change, and therapist variables, rather than patient variables, may hold more promise.

Women were found to have significantly better outcomes than men with regard to PTSD symptom severity and general psychological symptoms even after controlling for baseline severity scores. The effect of sex on PTSD treatment has been reported in some earlier studies (Békés et al., 2016; Eftekhar et al., 2013; Fletcher et al., 2017; Galovski et al., 2013), whereas other studies have found no effect of sex or even the opposite effect, where women had poorer treatment outcomes than men (Blain et al., 2010; Ehlers et al., 2013). The lack of consistency in these findings may be explained by differences in the rates of exposure to specific types of trauma, resulting in skewed sex distributions in some studies; for example, in mixed-trauma populations that contain a strong overrepresentation of men with military-related trauma exposure or women who have experienced sexual trauma (Blaine et al., 2010; Galovski et al., 2013). The present findings add to a small but growing evidence pool indicating that in civilian mixed-trauma PTSD populations, women seem to have better outcomes than men. This effect of sex could be the result of secondary benefits of therapy. Social support has been found to increase after PTSD treatment, and this increase in social support may be involved in the difference in outcomes between women and men. In a recent study, Békés et al. (2016) found that women reported a significantly larger increase in the frequency of supportive interactions with their partner and adaptive coping strategies, such as positive reappraisal and support-seeking, after treatment compared with men. Unfortunately, we did not have data on social support for the present study. Further analysis demonstrated that sex moderated outcomes with regard to treatment modality whereby women in individual therapy were found to have significantly better outcomes than men in individual therapy.

The naturalistic nature of the study contributed to several study strengths. First, it provided us with a broad range of variables to explore. This can be useful for creating hypotheses for further investigation through controlled experimentation. Second, the current study investigated psychotherapy as it occurred in the real world, giving the results a high degree of generalizability. The study design also had several limitations to discuss. First, with no randomization, the presence of an allocation bias to the different treatment modalities was likely. We statistically controlled for this to the degree possible, but the comparisons between group and individual therapy as well as between women and men must be considered explorative. Providing clinical recommendations based on explorative analyses is not advised. Second, due to the lack of a nonactive comparison group, we could not determine the degree to which the effects were due to the treatment versus regression to the mean. Third, the lack of a well-validated PTSD self-report measure was also a limitation because it might have impacted the findings regarding the effects of sex and treatment on PTSD symptoms. In addition, the use of the GAF, for which reliability may be suspect, constitutes a limitation as it might have led to an overestimation of the effect of the treatment. Fifth, we
were unable to measure the $p$ values of our LASSO model on the imputed dataset. As we ran a multiple regression model on the selected predictors, the predictive value of the model was reduced. Sixth, due to a lack of follow-up data, it is unclear whether the treatment effects persisted after EOT. The final and possibly most important limitation was that the dataset had large data loss both at baseline and EOT. We attempted to remedy this through multiple imputation procedures, but due to the lack of data at all time points, it was not possible to impute data for all cases.

We found that the effect of PTSD treatment in the MHS was relatively low compared to existing evidence in the field. Women had significantly better treatment outcomes than men, and no differences were found between group and individual therapy. These findings have several implications for both clinical practice and further research. The low treatment effects underline the importance of measuring the effectiveness of public mental health programs. They also indicate that some changes may improve the current treatment setup: Implementing evidence-based, first-line treatments in the mental health sector may be an important step toward optimizing treatment effects in a population of PTSD patients with severe symptom levels. In addition, ensuring that all clinicians are adequately trained in using these treatments is essential. Women and men may need different treatments, and men may need longer, more extensive CPs. However, although the naturalistic nature of the present study makes it inadvisable to formulate clinical guidelines based on these results, the results do indicate that it would be meaningful to use mixed-sex samples in future RCTs to investigate differences in treatment response. In addition, the present findings suggest the value in continuing to offer group-based modalities for CP PTSD treatments, as these are more cost-effective and could have extra benefits with regard to social functioning. The limitations of this study with regard to comparing group and individual therapy underline the need for new randomized controlled trials that compare group and individual interventions based on treatment principles derived from evidence-based interventions for PTSD.

**Open Practices Statement**

The study reported in this article was not formally preregistered. Neither the data nor the materials have been made available on a permanent third-party archive; requests for the materials should be sent via email to the lead author at f.b.scharff@gmail.com. The analyzed data are not under the authors' direct control; requests to access the data should be directed to the Danish National Archives.

**References**


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


Leichsenring, F., Jaeger, U., Masuhr, O., Dally, A., Dömpelmann, M., Fricke-Neef, C., Spitzer, C., & Steinert, C. (2020). To be or not to be improved: Patients’ perception of symptom improvement—linking the SCL-90-R to patient-rated global improvement in a large


This article is protected by copyright. All rights reserved.


This article is protected by copyright. All rights reserved.


https://www.ptsd.va.gov/professional/assessment/documents/PCL5_Standard_form.PDF


### TABLE 1

<table>
<thead>
<tr>
<th>Treatment element</th>
<th>AD</th>
<th>PTSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial psychiatric and somatic assessment</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Psychometric assessment</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Psychoeducation</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Counseling on complicating conditions (e.g., stress or drug abuse)</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Group psychotherapy or individual therapy</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Psychopharmacologic treatment</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Consultations with next of kin</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Continuity and coherence in ongoing treatment</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
This article is protected by copyright. All rights reserved.

### TABLE 2

Baseline sample characteristics of the sample.

<table>
<thead>
<tr>
<th>Variable</th>
<th>n</th>
<th>%</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>948</td>
<td>38.0</td>
<td>11.4</td>
<td></td>
</tr>
<tr>
<td>Female sex</td>
<td>728</td>
<td>76.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohabiting with partner</td>
<td>365</td>
<td>38.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Educational attainment ≥ 10 years</td>
<td>797</td>
<td>84.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed or studying</td>
<td>218</td>
<td>22.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sick leave &gt; 3 months</td>
<td>311</td>
<td>32.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed</td>
<td>366</td>
<td>38.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior psychotherapy treatment</td>
<td>757</td>
<td>79.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Symptom duration c &lt; 5 years</td>
<td>526</td>
<td>55.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Childhood sexual abuse f,g</td>
<td>278</td>
<td>29.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychopharmacology g</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antidepressants</td>
<td>323</td>
<td>34.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>212</td>
<td>22.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAF-S score</td>
<td>948</td>
<td>47.0</td>
<td>8.4</td>
<td></td>
</tr>
<tr>
<td>GAF-F score</td>
<td>948</td>
<td>49.2</td>
<td>17.3</td>
<td></td>
</tr>
<tr>
<td>SCL-90 (GSI) score g</td>
<td>719</td>
<td>1.7</td>
<td>0.63</td>
<td></td>
</tr>
</tbody>
</table>

Note. AD = adjustment disorder; PTSD = posttraumatic stress disorder.
CR-PTSD score\textsuperscript{d,f} & 719 & 1.9 & 0.70 \\
WHO-5 score\textsuperscript{e,f} & 740 & 24.9 & 16.8 \\

\textbf{Note.} \(N = 948\). GAF = Global Assessment of Functioning; GAF-F = GAF Functioning scale; GAF-S = GAF-S = GAF Symptom scale; SCL-90 (GSI): Symptom Checklist–90–Revised General Severity Index; WHO-5 = WHO-5 Well-Being Index; CR-PTSD = SCL-90 Crime-Related Posttraumatic Stress Disorder subscale. \textsuperscript{a}10 (1.1\%) missing. \textsuperscript{b}53 (5.6\%) missing. \textsuperscript{c}6 (0.6\%) missing. \textsuperscript{d}229 (24.2\%) missing. \textsuperscript{e}208 (21.9\%) missing. \textsuperscript{f}Patient-registered. \textsuperscript{g}Therapist registered.

\begin{table}[ht]
\centering
\begin{tabular}{lcccccc}
\hline
\textbf{Measure} & \textbf{Baseline} & \textbf{Posttreatment} & \textbf{\(\Delta\)} & \textbf{95\% CI} & \textbf{\(t\)} & \textbf{\(df\)} \\
 & \textbf{\(M\)} & \textbf{SD} & \textbf{\(M\)} & \textbf{SD} & & \\
\hline
GAF-S & 47.03 & 6.60 & 53.63 & 9.97 & 6.61*** & [5.58, 7.51] & 10.91 & 50.64 \\
GSI & 1.74 & .63 & 1.46 & .69 & -0.28*** & [-0.36, -0.20] & -8.46 & 35.53 \\
CR-PTSD & 1.86 & .69 & 1.53 & .78 & -0.33*** & [-0.41, -0.22] & -8.60 & 35.64 \\
\hline
\end{tabular}
\caption{Baseline and posttreatment scores for all outcome measures}
\end{table}


\textsuperscript{***}p < .001.
TABLE 4

Variables selected by the LASSO as potential predictors of any treatment outcome or treatment dropout in the unimputed dataset

<table>
<thead>
<tr>
<th>Outcome</th>
<th>CR-PTSD</th>
<th>GSI</th>
<th>GAF-F</th>
<th>GAF-S</th>
<th>WHO-5</th>
<th>Dropout</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predictor</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>GAF-F</td>
<td>0.001</td>
<td>0.001</td>
<td>0.009</td>
<td>0.021</td>
<td>0.022</td>
<td>-0.003</td>
</tr>
<tr>
<td>GAF-S</td>
<td>-0.004</td>
<td>-0.004</td>
<td>0.047</td>
<td>0.233</td>
<td>0.127</td>
<td>-</td>
</tr>
<tr>
<td>WHO-5</td>
<td>-0.008</td>
<td>-0.009</td>
<td>0.117</td>
<td>0.128</td>
<td>0.471</td>
<td>0.004</td>
</tr>
<tr>
<td>BHS</td>
<td>0.025</td>
<td>0.025</td>
<td>-0.086</td>
<td>-0.098</td>
<td>-0.573</td>
<td>-</td>
</tr>
<tr>
<td>Hours\textsuperscript{a}</td>
<td>-0.001</td>
<td>-</td>
<td>0.052</td>
<td>0.054</td>
<td>0.039</td>
<td>-</td>
</tr>
<tr>
<td>CR-PTSD</td>
<td>0.09</td>
<td>0.009</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Modality\textsuperscript{b}</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>25.393</td>
</tr>
</tbody>
</table>

\textit{Note.} n = 251. \textsuperscript{a}“\textsuperscript{a} indicates the variable was not selected. GAF = Global Assessment of Functioning; GAF-F = GAF Functioning scale; GAF-S = GAF-S = GAF Symptom scale; SCL-90 (GSI): Symptom Checklist–90–Revised General Severity Index; WHO-5 = WHO-5 Well-Being Index; CR-PTSD = SCL-90 Crime-Related Posttraumatic Stress Disorder subscale.

\textsuperscript{a}Hours spent in therapy. \textsuperscript{b}Group or individual.

FIGURE 1

\textit{Participant flow}

This article is protected by copyright. All rights reserved.