EUROPEAN JOURNAL OF PSYCHOTRAUMATOLOGY
2021, VOL. 12, 1958471
https://doi.org/10.1080/20008198.2021.1958471

EUROPEAN JOURNAL OF PSYCHO-TRAUMATOLOGY
THE OFFICIAL JOURNAL OF THE EUROPEAN SOCIETY FOR TRAUMATIC STRESS STUDIES

Taylor & Francis
Taylor & Francis Group

CLINICAL RESEARCH ARTICLE

OPEN ACCESS Check for updates

Predicting outcome of daycare cognitive behavioural therapy in a naturalistic sample of patients with PTSD: a machine learning approach

Heiner Stuke[a], Nikola Schoofs[a], Helen Johanssen[a], Felix Bermpohl[a], Dominik Ülsmann[b], Olaf Schulte-Herbrüggen[a,b] and Kathlen Priebe[a]

[a]Department of Psychiatry and Psychotherapy, Charité – Universitätsmedizin Berlin, Berlin, Germany; [b]Friedrich Von Bodelschwingh-Clinic for Psychiatry, Psychotherapy and Psychosomatics, Berlin, Germany

ABSTRACT

Background: Identifying predictors for treatment outcome in patients with posttraumatic stress disorder (PTSD) is important in order to provide an effective treatment, but robust and replicated treatment outcome predictors are not available up to now.

Objectives: We investigated predictors of treatment outcome in a naturalistic sample of patients with PTSD admitted to an 8-week daycare cognitive behavioural therapy programme following a wide range of traumatic events.

Method: We used machine learning (linear and non-linear regressors and cross-validation) to predict outcome at discharge for 116 patients and sustained treatment effects 6 months after discharge for 52 patients who had a follow-up assessment. Predictions were based on a wide selection of demographic and clinical assessments including age, gender, comorbid psychiatric disorders, trauma history, posttraumatic symptoms, posttraumatic cognitions, depressive symptoms, general psychopathology and psychosocial functioning.

Results: We found that demographic and clinical variables significantly, but only modestly predicted PTSD treatment outcome at discharge (r = 0.21, p = .021 for the best model) and follow-up (r = 0.31, p = .026). Among the included variables, more severe posttraumatic cognitions were negatively associated with treatment outcome. Early response in PTSD symptomatology (percentage change of symptom scores after 4 weeks of treatment) allowed more accurate predictions of outcome at discharge (r = 0.56, p < .001) and follow-up (r = 0.43, p = .001).

Conclusion: Our results underscore the importance of early treatment response for short- and long-term treatment success. Nevertheless, it remains an unresolved challenge to identify variables that can robustly predict outcome before the initiation of treatment.

ARTICLE HISTORY
Received 19 May 2020
Revised 28 June 2021
Accepted 11 July 2021

KEYWORDS
PTSD; behavioural therapy; outcome prediction; individualized treatment

PALABRAS CLAVE
TEPT; terapia conductual; predicción de resultados; tratamiento individualizado

关键词
PTSD; 行为疗法; 结果预测; 个体化治疗

HIGHLIGHTS

- Psychotherapy can improve PTSD, but many patients do not respond adequately.
- We found that clinical variables (high posttraumatic cognitions but low re-experiencing symptoms) modestly predicted poor response to CBT.
- Early therapy response more accurately predicted final outcome.

Predecir el resultado de la terapia cognitiva conductual en un hospital de día en una muestra naturalista de pacientes con TEPT: Un enfoque de aprendizaje automático

Antecedentes: La identificación de los predictores para el resultado de tratamiento en pacientes con trastorno de estrés postraumático (TEPT) es importante para proporcionar un tratamiento eficaz, pero hasta ahora no se dispone de predictores de respuesta de tratamiento robustos y replicables.

Objetivos: Investigamos los predictores de resultado de tratamiento en una muestra naturalista de pacientes con TEPT ingresados a un programa de tratamiento cognitivo conductual tipo hospital de día de ocho semanas, después de una amplia gama de eventos traumáticos.

Método: Utilizamos el aprendizaje automático (regresores lineales y no lineales y validación cruzada) para predecir el resultado al alta para 116 pacientes y los efectos sostenidos del tratamiento a los seis meses del alta para 52 pacientes que tuvieron una evaluación de seguimiento. Las predicciones se basaron en una amplia selección de evaluaciones demográficas y clínicas que incluyen edad, género, trastornos psiquiátricos comórbidos, antecedentes de trauma, síntomas postraumáticos, cogniciones postraumáticas, síntomas depresivos, psicopatología general y funcionamiento psicosocial.

Resultados: Encontramos que las variables clínicas y demográficas predijeron de manera significativa, pero solo modestamente, el resultado del tratamiento del TEPT al momento del alta (r = 0.21, p= .21 para el mejor modelo) y el seguimiento (r = 0.31, p = .026). Entre las variables incluidas, las cogniciones postraumáticas más severas se asociaron negativamente con el resultado del tratamiento. La respuesta temprana en la sintomatología de TEPT (cambio porcentual del puntaje en los síntomas después de cuatro semanas de tratamiento) permitió predicciones más precisas de los resultados al alta (r = 0.56, p < .001) y el seguimiento (r = 0.43, p = .001).

CONTACT Heiner Stuke heiner.stuke@charite.de Department of Psychiatry and Psychotherapy, Charité-Universitätsmedizin Berlin, Charitéplatz 1, Berlin, 10117, Germany

Supplemental data for this article can be accessed here.

Conclusiones: Nuestros resultados subrayan la importancia de una respuesta temprana al tratamiento para el éxito del tratamiento a corto y largo plazo. No obstante, sigue siendo un desafío sin resolver identificar variables que puedan predecir de manera sólida el resultado antes del inicio del tratamiento.

患者自然样本中日托认知行为疗法的预测结果:机器学习方法

背景:确定创伤后应激障碍 (PTSD) 患者治疗结果的预测因子对于提供有效治疗很重要, 但目前还没有可靠且可重复的治疗结果预测因子。
目的:我们在经历一系列创伤事件后接受八周日托认知行为治疗计划的 PTSD 患者自然样本中, 考查了治疗结果的预测因素。
方法:我们使用机器学习 (线性和非线性回归因子以及交叉验证) 来预测 116 名患者的出院结果和 52名接受了随访评估的患者出院后 6个月的持续治疗效果。预测基于广泛选择的人口统计学和临床评估, 包括年龄, 性别, 并发精神障碍, 创伤史, 创伤后症状, 创伤后认知, 抑郁症状, 一般精神病和社会心理功能。
结果:我们发现人口统计学和临床变量显著但仅中度预测出院时 (最佳模型 $r = 0.21$, $p = .021$) 和随访时 ($r = 0.31$, $p = .026$) 的 PTSD 治疗结果。在纳入的变量中, 更严重的创伤后认知与治疗结果呈负相关。PTSD 症状学的早期反应 (治疗 4 周后症状评分变化的百分比) 允许更准确地预测出院时 ($r = 0.56$, $p < .001$) 和随访时的结果 ($r = 0.43$, $p = .001$) 。
结论:我们的结果强调了早期治疗反应对于短期和长期治疗成功的重要性。然而, 确定治疗开始前可以稳健预测结果的变量仍然是一个未解决的挑战。

1. Introduction

With an estimated cross-national lifetime prevalence of 3.9%, posttraumatic stress disorder (PTSD) is a frequent mental disorder, often with significant impact on quality of life (Koenen et al., 2017; Yehuda et al., 2015). Although several psychotherapeutic treatments for PTSD have shown general efficacy, a substantial proportion of patients fails to respond to treatment (Bradley, Greene, Russ, Dutra, & Westen, 2005; Cusack et al., 2016; Watts et al., 2013). For instance, using author-defined criteria for clinically meaningful improvement, Bradley et al. reported that only 44% of all those who entered treatment and 54% of those who completed treatment were classified as improved at the end of the treatment (Bradley et al., 2005). Recently published treatment guidelines from the American Psychological Association (Courtois et al., 2017), the US Department of Veterans Affairs (Ostacher & Cifu, 2019) and the International Society for Traumatic Stress Studies (Berliner et al., 2019) strongly recommend the use of prolonged exposure (PE), cognitive processing therapy (CPT) and trauma-focused cognitive-behavioural therapy (TF-CBT) for the treatment of PTSD. However, there is scant evidence regarding the prediction of response and non-response and to guide the decision on which treatment to recommend for which patient. In order to tailor treatments to specific characteristics of patients, it is thus crucial to identify predictors of treatment success, such as demographic, biological and clinical characteristics of a patient.

In a recent review on predictors of the success of psychological therapies for PTSD in 25 randomized controlled trials (RCTs), the authors concluded that 'associations were neither consistent nor strong' (Lewis, Roberts, Andrew, Starling, & Bisson, 2020). The most consistent associations (i.e. at least two studies reporting an effect in the same direction) were found for adherence to homework and experience of a more recent trauma as predictors for better treatment outcome and a comorbid diagnosis of depression as predictor for a worse treatment outcome. However, also for these predictors, the evidence was not unequivocal. For instance, for a more recent trauma, there were in addition to the two studies that reported an association, three other studies that reported no such association. Overall, most of the examined predictors (e.g. age, gender, employment and marital status, income) were not significantly associated with treatment outcome in the majority of studies. Results of studies on naturalistic (non-RCT) studies have to our knowledge not been synthesized in a systematic fashion so far. The results of these studies, however, largely confirm the results obtained from RCTs. For instance, van Minnen et al. showed no predictive value of demographic variables, clinical characteristics (depression and general anxiety, personality traits and disorders), trauma characteristics, and certain trauma-related feelings such as anger, guilt, and shame (van Minnen, Arntz, & Keijsers, 2002) on treatment outcome of patients with PTSD after mixed traumas. The null result for depression and dissociation was later replicated (Hagenaars, van Minnen, & Hoogduin, 2010) and an additional study in veterans with military-related PTSD did not find chronicity, alcohol use, and anxiety or depression severity as significant predictors for treatment outcome (Richardson, Elhai, & Sarreen, 2011). A relatively large study on treatment outcome in 330 patients with PTSD after mixed traumas found social problems and multiple traumas to be associated with

worse treatment outcome whilst a missing relationship, a comorbid mood disorder, suicide attempts, a history of substance dependence and months since trauma were not significantly associated with treatment outcome (Ehlers et al., 2013). One study identified anger, alcohol use, and depression as predictors of worse treatment outcome (Forbes, Creamer, Hawthorne, Allen, & McHugh, 2003), but the finding of anger was not replicated by a later study (Clifton, Feeny, & Zoellner, 2017). Finally, it was shown that a predictor index composed of a variety of assessments of medical doctors, psychologists and social workers could significantly predict treatment outcome, whereas the only single item significantly related to (worse) treatment outcome was unemployment (Sonne et al., 2016).

In summary, it must be stated that robust and replicated pre-treatment predictors of therapeutic success are not yet available. Since associations between trauma-related (negative) cognitions and PTSD severity as well as between the improvement of these cognitions and the improvement of PTSD symptoms have been clearly shown (Brown, Belli, Asnaani, & Foa, 2019), it is conceivable that the level of the trauma-related cognitions at the start of therapy has an impact on the outcome of the therapy. Similarly, the centrality of the traumatic event for the identity of the patient (Berntsen & Rubin, 2006) has recently attracted increasing attention as an influencing factor both on the severity of PTSD symptoms and on the positive processing of traumatic experiences (posttraumatic growth) (Groleau, Calhoun, Cann, & Tedeschi, 2013). To the best of our knowledge, however, the utility of the centrality of the event for predicting treatment outcome has not yet been investigated. The influence of rumination, which is ascribed a relevant role in the development and maintenance of PTSD (Moulds, Bisby, Wild, & Bryant, 2020), on the therapeutic outcome has also not been tested. Moreover, it could be shown for various mental disorders and corresponding interventions that a good early response to the therapy predicts a favourable outcome at the end of the treatment (e.g. (Lewis, Simons, & Kim, 2012; Lutz, Stulz, & Köck, 2009; Schindler, Hiller, & Witthöft, 2013; Van et al., 2008) for psychotherapy in depression and (Henkel et al., 2009; Katz, Meyers, Prakash, Gaynor, & Houston, 2009; Koran et al., 1995; Lin, Park, & McIntyre, 2019; Mulder, Joyce, Frampton, Luty, & Sullivan, 2006; Nierenberg et al., 1995; Stamm et al., 2014; Szegedi et al., 2003) for pharmacotherapy in depression). In the only study known to us on the predictive value of early response for the therapeutic outcome in PTSD, response after 2 weeks was related to a good treatment outcome in pharmacotherapy (sertraline) but not significantly in PE (Graham et al., 2018). In the present study, in addition to previously studied potential predictors such as demographics, symptom severity, and trauma characteristics, we tested the value of trauma-related cognitions, the centrality of event, rumination and the early treatment response for predicting the outcome of PTSD therapy.

Machine learning has mainly been used in PTSD research to study risk factors for the development of PTSD (Ramos-Lima, Waikamp, Antonelli-Salgado, Passos, & Freitas, 2020). Only few studies have employed machine learning algorithm to identify predictors of treatment outcome. In a sample of PTSD inpatients, (Herzog et al., 2021) found that a higher age, a wish to retire or being retired, the total number of comorbid diagnoses and more severe depressive symptoms were negative predictors of treatment outcome. Higher posttraumatic symptoms were positively associated with treatment outcome.

In the present study, we used machine learning to predict PTSD treatment outcome. The approach consisted of three different components: First, we used cross-validation, in which a predictive model is trained on one group and tested on another group of the patients. In this way, it can be determined directly how well the outcome of patients who were not part of the training process can be predicted based on the relationships learned in the training process with the other patients (single-case predictions). Therewith, cross-validation aims to increase the generalizability of the results across the individual data set (Kearns & Ron, 1999; Thompson, 1994). Second, we used feature reduction techniques to calculate a smaller number of components capturing much of the variance of a large number of predictors. This makes it possible to examine a higher number of individual and potentially correlated predictors (e.g. individual items in questionnaires). Third, in addition to conventional linear regression models, we also used a non-linear regression method. In linear regression, the dependent variable is modelled as a weighted sum of the predictors, so only additive relationships between independent and dependent variables can be captured. Nonlinear regression (e.g. tree-based methods) can also detect when a relationship between independent and dependent variables is not proportional and can thus help to identify predictive patterns that would have been overlooked with purely linear methods.

By these means, we aimed to test the predictive value of pre-treatment characteristics and early treatment response for the outcome of a daycare cognitive-behavioural therapy programme in patients with PTSD. In addition, we examined the predictors of long-term therapeutic outcome after 6 months in a subgroup of patients with follow-up assessment.

2. Methods

2.1. Treatment and data collection

Data were collected as part of routine clinical monitoring in a Berlin day clinic for the treatment of PTSD. The cognitive behavioural therapy daycare programme ran from 8:00 am to 3:30 pm from Monday to Friday over the course of an average of 8.59 treatment weeks (SD = 1.4) in the included sample. The treatment programme included four weekly sessions of trauma-focused individual therapy, as well as daily trauma-focused group therapy and was conducted by mental health professionals (psychologists, psychiatrists, nurses).

Treatment followed the CPT manual (Resick, Monson, & Chard, 2016) and was supplemented by behavioural experiments and subsequent monitoring of the reduction of avoidance and safety-seeking behaviour (Ehlers, Clark, Hackmann, McManus, & Fennell, 2005).

Patients were asked to write a statement on their beliefs why the traumatic event has happened, and how it has affected their beliefs about self, others and the world, especially regarding safety, trust, control/power, esteem, and intimacy. This „Impact Statement" was used to identify dysfunctional trauma-related beliefs („stuck points") which were subsequently challenged by socratic dialogue. The worksheet „Challenging Questions" was introduced in order to support the patients in changing their beliefs. In addition, patients were to write a narrative on their most distressing traumatic event and to read the written account to the therapist and daily to themselves. After working on the most distressing traumatic event, other traumatic events could be addressed if necessary.

Group therapy consolidated and extended these topics and comprised groups on psycho-education, on thought and behavioural analysis, on challenging dysfunctional beliefs, on planning and discussing the results of behavioural experiments and on monitoring the reduction of avoidance and safety-seeking behaviours.

Patients were admitted to the programme after a preliminary outpatient session that was led by a trained and experienced clinician (psychologist or psychiatrist). During this, the indication for a trauma-focused therapy was evaluated by a semi-structured clinical interview based on the diagnostic criteria of PTSD according to ICD-10. Moreover, posttraumatic symptomatology was confirmed by the Davidson Trauma Scale (DTS; (Davidson et al., 1997)), where patients had to have a sum score in the DTS of 40 or more in addition to the clinician confirmed PTSD diagnosis. This cut-off was previously identified as the score with the highest efficiency for the diagnosis (Davidson et al., 1997). The indication for trauma-focused treatment was hence established by clinical assessment in combination with diagnostic criteria and sufficient symptom severity in the DTS.

Since data were collected as part of the routine clinical monitoring included in the treatment, patients did not sign a consent statement. Data collection and analysis was approved by the ethical commission of the Charité Berlin university hospital. Every patient was asked to complete assessments at admission (baseline), after 4 weeks, at discharge, and 6 months after discharge (follow-up).

2.2. Predictors

Assessments at admission included the following demographic and clinical variables:

Demography: Patients provided information on age and gender in a self-report questionnaire.

Comorbid psychiatric disorder: Co-occurring substance use disorders and co-occurring affective disorders were assessed by trained and experienced clinicians (psychologists and psychiatrists) according to ICD-10.

Trauma characteristics: The Life-Event-Checklist of the Posttraumatic Diagnostic Scale (PDS) (Foa, Cashman, Jaycox, & Perry, 1997) was used to assess trauma history as well as characteristics and time of the index trauma.

Posttraumatic symptoms: The Davidson Trauma Scale (DTS) (Davidson et al., 1997) for DSM-IV served as the main PTSD symptom severity measure. It is a 17-item self-report measure assessing the 17 DSM-IV symptoms of PTSD. Items are rated on 5-point frequency (0 = 'not at all' to 4 = 'every day') and severity scales (0 = 'not at all distressing' to 4 = 'extremely distressing'). In addition, frequency of the 17 DSM-IV symptoms was assessed with the PDS.

Posttraumatic cognitions: The Posttraumatic Cognitions Inventory (PTCI) (Foa, Ehlers, Clark, Tolin, & Orsillo, 1999), measures trauma-related thoughts and beliefs with 48 items. The short version of the centrality of event scale (CES-7) (Berntsen & Rubin, 2006) was used to measure how central the trauma is to the patient's identity and life story.

Rumination: The Perseverative Thinking Questionnaire (PTQ) (Ehring et al., 2011) measures repetitive negative thinking on 15 items.

Depressive symptoms: The 21 items of the Beck Depression Inventory II (BDI) (Beck, Steer & Brown, 1996) were included for measuring depressive symptoms.

General psychopathology: The short version of the Brief Symptom Inventory (BSI-18) (Rath & Fox, 2018)

was used to assess psychological distress (somatization, depression and anxiety).

Psychosocial functioning: The 'Index zur Messung von Einschränkungen der Teilhabe' (IMET, english: 'Index for the Assessment of Participation Impairments') (Deck, Mittag, Hüppe, Muche-Borowski, & Raspe, 2011) measures disease-related functioning and inabilities in several areas of life such as work, housework or social relationships with 9 items. Additionally, we included the items of the PDS, which ask how the PTSD symptoms effect several areas of life.

We used these demographic data, comorbidities, and the total scores of the self-assessment scales described above (in case of the PDS, the 17 items of the PDS that assess the severity of different PTSD symptoms were summed up to a total score) to predict treatment outcome. Hence, the analysis comprised 12 predictors (sex, age, two comorbidity variables, PDS, DTS, PTCI, CES, PTQ, BDI, BSI, and IMET total scores). In an exploratory analysis, we aimed to predict treatment outcome based on single items instead of total scores. This was motivated by the assumption that, for example, individual posttraumatic cognitions, as recorded by the PTCI, in combination with individual PTSD symptoms, as assessed by DTS and PDS, might enable predictions about therapy results and that this potential possibly might be overlooked if only the total scores were investigated. However, this analysis is based on a high number of predictors compared to the sample size and hence, its results require further confirmation. We therefore present details on this analysis and its results in the Supplementary Material.

In addition to these 'baseline models', we tested the predictive power of 'early response models', in which the percent symptom score reduction 4 weeks after admission (early response) was included as an additional predictor. This analysis examines whether the outcome at discharge can be predicted with the early response after 4 weeks. If such a prediction should be possible, it would be of great interest whether the prediction is based solely on the fact that the progress achieved up to week 4 is maintained until the end of therapy or whether progress up to week 4 also predicts the further course of therapy (week 4 until discharge). In order to answer this question, we also examined whether the early response is related to the further response (week 4 to discharge) by computing Pearson correlations between the two.

2.3. Treatment outcome

The DTS served as the main PTSD symptom severity measure, which was assessed at baseline, after 4 weeks of therapy, at discharge and 6 months after the end of therapy. Treatment outcome was defined as the percentage change in the DTS total score (severity and frequency of 17 PTSD symptoms) between admission and discharge, early response as the percentage change between admission and the intermediate assessment at week 4. We chose percentage of change (instead of the absolute value of symptom score improvement) as our outcome measure, because it takes differences of pre-treatment symptom severity into account (Hiller, Schindler, & Lambert, 2012). Using absolute change relies on the assumption that the expected treatment effect is independent of baseline severity, which is not true for many disorders and treatments and might lead to poorer statistical fit and greater measurement error (Karin, Dear, Heller, Gandy, & Titov, 2018). The usage of percentage instead of absolute improvement was particularly important in our study, because we observed a significant correlation between symptom severity (DTS total score) at admission and absolute improvement (DTS total score at discharge minus DTS total score at admission): $r = -0.241$, $p = .009$. Standardizing the improvement with respect to the baseline severity (i.e. calculating the percentage change) eliminated the correlation between baseline severity and outcome measure: $r = 0.020$, $p = .830$. Hence, the treatment outcome was calculated as (DTS total score at discharge – DTS total score at admission)/DTS total score at admission. Here, negative values indicate better treatment outcome.

2.4. Imputation of missing values

There were varying degrees of missing data, with some patients having entirely missing scales and some patients only having missing single items of otherwise completed scales. Since we used individual items of the scales as predictors in some analyses, patients were excluded in whom scales were completely missing, because the basis for an imputation of missing values was not available in this case (e.g. it would be difficult to impute the severity of single posttraumatic cognitions assessed by the PTCI without having any assessments on posttraumatic cognitions for the patient in question). We imputed missing single items of a scale if less than 20% of the items of the scale were missing by putting in the mean value of available items for the same scale and patient. This applied for 10 patients with imputation of single item values.

2.5. Data preprocessing, regression models, and cross-validation

When predictive models with a large number of potential predictors are fitted, spurious relationships can be found (related to the statistical issues of multiple testing and model overfitting). To avoid such false-positive findings, we chose

to probe predictive performance in a cross-validation approach, in which training and test set are strictly separated (the analysis procedure is shown in Figure 1). Specifically, we used leave-one-out cross-validation, in which relationships between predictors and dependent variable are learned in a training set, which comprises all but one (left out) patient. The model which is fitted to the patients in the training set is then applied to predict treatment outcome in the left out (test) patient and this prediction is compared with the true outcome of the test patient. This procedure is repeated for all patients (all patients are sequentially left out and the predictions are trained on the other patients). By these means, it can be tested, which predictors allow reliable predictions of treatment outcome in patients, who were not part of the training process and the generalizability of results can be improved (Kearns & Ron, 1999; Thompson, 1994).

Since the examined predictors consisted of strongly correlated variables, we used a feature reduction technique that attempts to reduce the predictor variables to a small number of independent components. For this purpose, we applied a principal component analysis (PCA), which aims at capturing the variance of the original features with a linear combination of a pre-specified smaller number of components (Burges, 2010). We determined the optimal number of components of the PCA in a hyperparameter optimization with 10-fold cross-validation in the training set, where the number of components (with candidate numbers from 1 to 10) with the lowest squared error in the training set was selected. The resulting components for each participant were then used (in the case of the early response models supplemented by the early response) as input for the regression models.

To find possible linear and non-linear predictive patterns, we used linear and non-linear regression models. Linear combinations of predictors were assessed using ordinary least squares linear regression. Non-linear regression was performed with an ADAboost regressor (Drucker, 1997) as implemented in the scikit-learn toolbox for Python. In short, the predictions of this regressor are based on a combination of predictions from 'weak' regressors

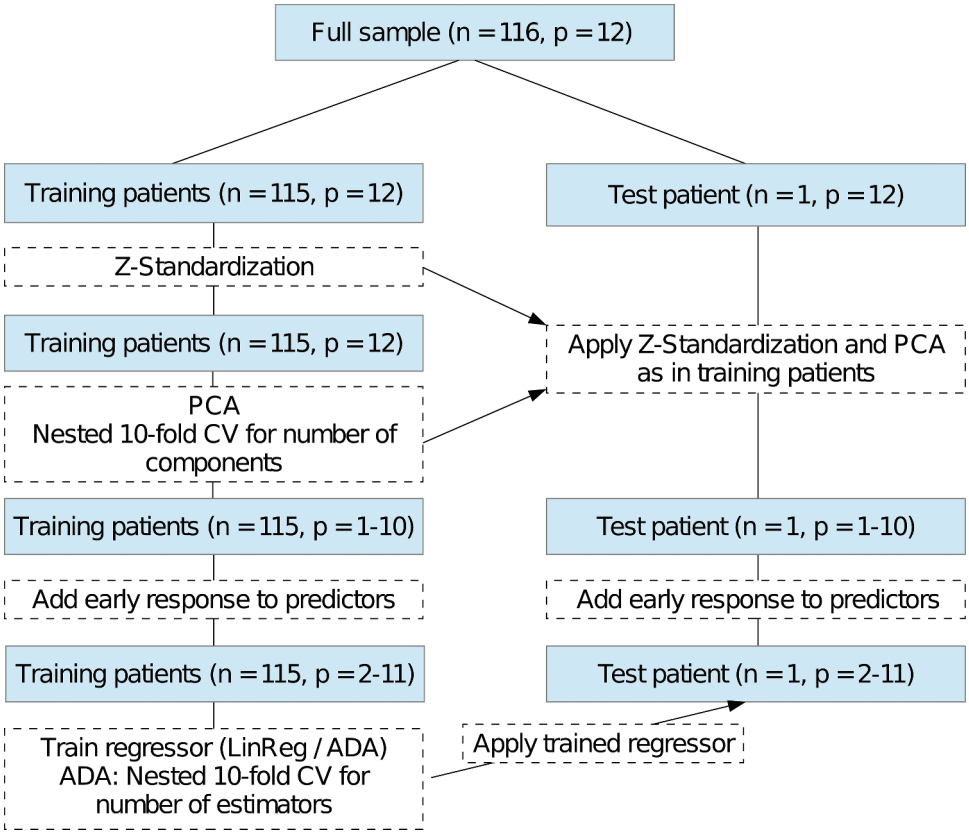

Figure 1. Schematic representation of the analysis process in one fold of the leave-one-out cross-validation. After splitting the data in training set and test patients, features were z-standardized and a principal component analysis (PCA) was performed to represent the variance of the features with a smaller number of components. The optimal number of components was chosen in a nested 10-fold cross-validation within the training set with a candidate grid of 1–10 components and selection of the number of components with the lowest squared error in the test set of the nested CV. In the early response models, early response was added as an additional predictor. After this, the regressor was trained in the training set. In the test patient, the same transformations as in the training set were applied (and early response was added in case of early response models). Then the regressor trained in the training set was used to predict treatment outcome in the test patient. The procedure was repeated for all patients (all patients were sequentially left out and the predictions were trained on the other patients). Finally, the predicted treatment outcomes of the test patients were compared with their true treatment outcomes. *n*: sample size; *p*: number of predictors.

(typically decision trees), which are weighted in such a way that 'weak regressors', which contributed significantly to the prediction of the target in the training set, are given a higher impact. ADAboost can learn linear and non-linear relationships between predicting and dependent variables and deals well with not normally distributed data. The crucial hyperparameter of the ADAboost regressor, the number of estimators, was also optimized with 10-fold cross-validation in the training set (with a candidate grid of 2, 5, 10, 20, 40 estimators), the other parameters were set to the default values of the scikit-learn implementation.

In the described manner, we predicted firstly the therapy outcome (percentage symptom score change from admission to discharge, n = 116) and secondly the sustained effect (percentage symptom score change from admission to follow-up, n = 52). In each case, Pearson correlations between predicted and true outcome of the test patients were computed to assess the predictive power of the included predictors. This analysis tests if relationships between the true outcome and the predicted outcome of test patients (who were not themselves used to learn relationships between predictors and outcome) were significantly better than chance and how much variance in true outcomes was explained by predicted outcomes. In case of significant overall predictions, we investigated univariate associations of the predictors to the outcome (Pearson correlation coefficients).

3. Results

3.1. Patient characteristics

In total, 209 patients were consecutively admitted to the therapeutic programme from April 2015 to June 2019. Of this initial sample, 75 participants could not be included in our study because of either not meeting diagnostic criteria for PTSD (i.e. having a DTS sum score of less than 40, n = 10) or having missing predictor values (baseline questionnaires or early response assessments after 4 weeks, n = 65).

Of the included 134 patients, 116 completed the therapy. The 18 non-completers and the 116 completers did not differ significantly in sex (non-completers 4 males and 14 females, completers 34 males and 82 females, chi-square test p = .734), age (mean for non-completers = 39.8, SD = 12.2, mean for completers = 41.6, SD = 11.1, two-sample t-test t = 0.625, p = .533), number of distinct trauma types (according to the life-event-checklist of the PDS, median for non-completers = 4, median for completers = 4, Mann–Whitney U test, U = 937, p = .242), baseline symptom severity of PTSD (DTS sum score, mean for non-completers = 82.7, SD = 24.9, mean for completers = 82.5, SD = 20.0, two-sample t-test t = −0.049, p = .961) or depression (BDI sum score, mean for non-completers = 29.4, SD = 12.6, mean for completers = 29.9, SD = 10.3, two-sample t-test t = 0.198, p = .843).

The descriptive statistics of the included sample are summarized in Table 1. Patients reported on average 4.4 (SD = 2.2) different trauma types. The index trauma was in 41.4% of the cases related to sexual assault (childhood: n = 36; adulthood: n = 6; age unknown: n = 6), in 28.4% to physical assault (childhood: n = 12, adulthood: n = 13; age unknown: n = 8), in 17.2% to armed assault or combat experience or terror attack (n = 20), in 10.3% to accidents (n = 12), and in 2.6% to other incidents (witness of suicide or death; n = 3). The time between the index trauma and the start of therapy was 1–3 months (2 patients), 3–6 months (9 patients), 6 months – 3 years (20 patients), 3 years – 5 years (15 patients), more than 5 years (70 patients). The majority of patients had received previous psychosocial treatment, 25% outpatient only (n = 29), 9% inpatient only (n = 10), 22% outpatient and inpatient (n = 26), 37% no previous treatment (n = 43), 7% missing information on previous treatments (n = 8).

Fifty-two of the 116 patients also showed up for a long-term follow-up with symptom assessment 6 months after discharge. Patients with and without follow-up assessment differed significantly in symptom severity of PTSD (DTS sum score) at admission (t = 2.803, p = .006), but not at discharge (t = 0.606, p = .546), and not in BDI scores, age or in the number in distinct trauma types according to the PDS (Table 1).

3.2. Treatment outcome in the whole sample

In the whole sample, the average DTS sum score was 82.5 (SD = 20.0) at admission, 80.5 (SD = 22.3) after 4 weeks and 61.9 (SD = 28.5) at discharge. This corresponds to a treatment effect size (Cohen's d with

Table 1. Patient characteristics.

	Main sample (n = 116)	Sample with follow-up assessment (n = 52)
Measure	Absolute numbers	Absolute numbers
Sex	82 females, 34 males	38 females, 14 males
Measure	Mean (SD)	Mean (SD)
Age	41.6 (11.1)	42.4 (12.0)
BDI total score at admission	29.9 (10.3)	28.6 (10.3)
BDI total score at discharge	21.3 (13.0)	21.1 (13.6)
DTS total score at admission[1]	82.5 (20.0)	76.9 (18.7)
DTS total score at discharge	61.9 (28.5)	60.1 (30.1)
Number of distinct trauma types[2]	4.4 (2.2)	4.0 (2.2)

DTS: Davidson Trauma Scale; BDI: Beck Depression Inventory
1 Difference between sample with (n = 52) and without (n = 64) follow-up assessment is significant (two-sample t-test)
2 According to the life-event-checklist of the Posttraumatic Diagnostic Scale, which queries twelve trauma types

pooled variances) of d = 0.84 (95% CI [0.57 1.11], significant admission-to-discharge difference with t = 8.6, p < .001, paired t-test). At discharge, 49 out of 116 participants had a symptom improvement of 30% or more and 24 of those had an improvement of more than 50%.

The distribution of DTS scores at the assessment time points is shown in Figure 2.

3.3. Treatment outcome and sustained effects in the sample with follow-up

In the subsample with follow-up assessment, average DTS sum score was 76.8 (SD = 18.7) at admission, 60.1 (SD = 30.1) at discharge, and 54.8 (SD = 30.9) at follow-up corresponding to a treatment effect size of d = 0.67 (95% CI [0.27 1.07], t = 4.9, p < .001) at discharge and d = 0.89 (95% CI [0.46 1.27], t = 5.4, p < .001) at follow-up. At discharge, 23 out of 52 participants had a symptom improvement of 30% or more and 11 of those had an improvement of more than 50%. At follow-up, 24 out of 52 participants had a symptom improvement of 30% or more and 16 of those had an improvement of more than 50%.

3.4. Predictors of treatment outcome and sustained effects

All investigated models and their performance in predicting treatment outcome and sustained effects are summarized in Table 2.

It can be seen that without early response, treatment outcome at discharge and follow-up could be predicted significantly better than chance with linear models, whilst ADABoost tended to perform slightly worse and predictions did not reach the significance threshold. However, the explained variance in treatment outcome remained rather small (r = 0.214, p = .021 at discharge and r = 0.309, p = .026 at follow-up). The used predictors and their univariate correlations with treatment outcome are summarized in Table 3 for total scores and linear models. Univariate correlations of total scores revealed that severe posttraumatic cognitions, higher centrality of event, and higher depressive symptoms were associated with a somewhat poorer outcome (please note that a negative correlation indicates better outcome with higher predictor values).

Inclusion of early (week 4) treatment response as an additional predictor allowed more robust predictions of treatment outcome at discharge and (to a lower degree) at follow-up (Table 2), demonstrating that the progress in the first 4 weeks is predictive for the outcome at discharge. However, the further progress after 4 weeks (i.e. the percentage change in symptom severity from week 4 to discharge) was not related to the early treatment outcome: r = −0.035, p = .710. This suggests that predictions of outcome at discharge were based on improvements in the early therapy phase, whilst improvements in the later therapy phase were independent from early response.

4. Discussion

In the present study, we aimed to predict the outcome of a trauma-focused cognitive-behavioural therapy for PTSD based on a variety of pre-treatment demographic and clinical variables as well as early response after 4 weeks. With linear models and total scores of pre-treatment variables only, treatment outcome at discharge and 6 months later could be predicted significantly better than chance. Although there was no single exceptionally strong predictor, a high degree of posttraumatic cognitions was particularly associated with a poor therapy outcome. However, the predictive power was rather limited (r around 0.2–0.3), which is in the lower range reported by other studies using machine learning to predict outcome in psychotherapy (Herzog et al., 2021; Hilbert et al., 2020) and arguably below a threshold for clinical utility.

We found that high levels of posttraumatic cognitions were associated with less improvement. This is in line with the finding by Moser et al. (Moser, Cahill, & Foa, 2010), but contrary to previous studies that found no impact of pre-treatment cognitions on outcome ((Jun, Zoellner, & Feeny, 2013; Lindebo Knutsen, Sachser, Holt, Goldbeck, & Jensen, 2020) for PE (Moser et al., 2010)). A recent meta-analysis on the association of cognitions and PTSD also highlighted the role of self-appraisals (Gomez de La Cuesta, Schweizer, Diehle, Young, & Meiser-Stedman, 2019). Beliefs about oneself as diminished, defeated and worthless have been included as one of the three domains of disturbances in self-organization in the new ICD-11 complex PTSD diagnosis. It has been shown that complex presentations of PTSD are related to worse treatment outcome (Karatzias et al., 2019). Treatment outcomes may be improved if these cognitions are addressed more successfully.

While many previous studies have identified the severity of PTSD symptoms as a predictor for treatment outcome, conflicting results have been reported regarding the direction of this effect. In our study, greater PTSD symptoms at baseline were associated with larger absolute improvements, consistent with the results of several studies (e.g. (Forbes et al., 2003; Herzog et al., 2021; Richardson et al., 2011)). However, it is contrary to other studies (e.g. (Lindebo Knutsen et al., 2020; Litz et al., 2019; van Minnen et al., 2002)), which found that patients with higher baseline PTSD scores improved less than those with lower baseline scores. One explanation for positive relationships between baseline symptoms and absolute improvements might be that greater baseline

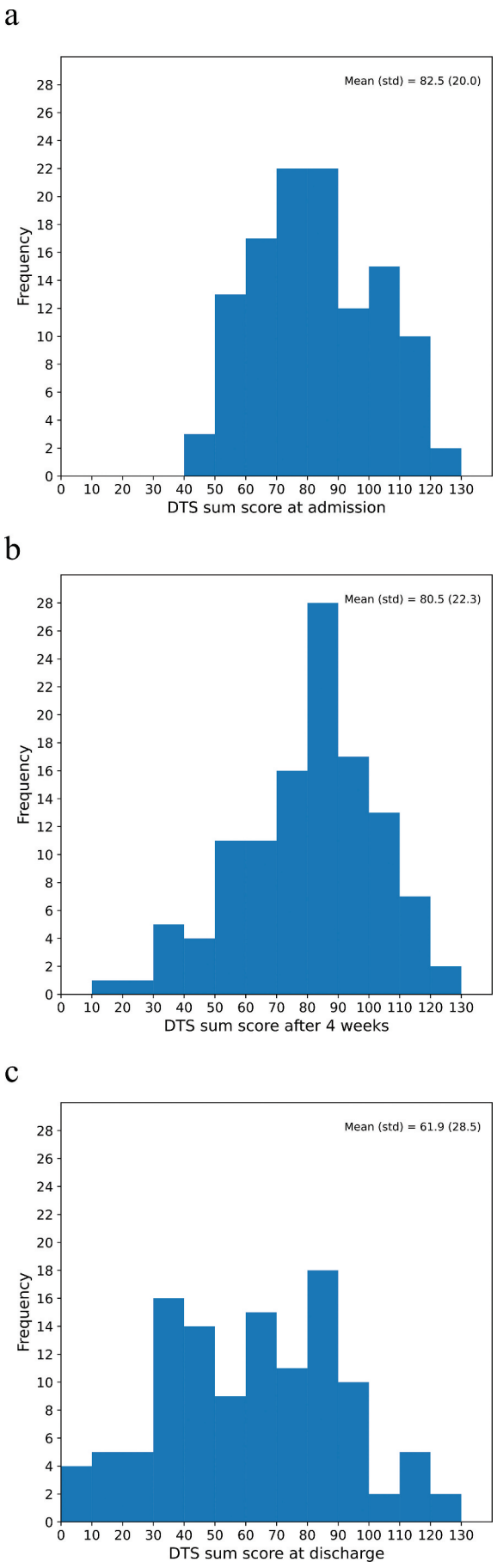

Figure 2. Distribution, mean values and standard deviation of Davidson Trauma Scale (DTS) sum scores at admission (a), after 4 weeks of treatment (b), and at discharge (c).

Table 2. Predictive performance of predicting outcome at discharge and follow-up.

Outcome at discharge (n = 116)		
Predictors	Regressor	r and p in CV
Total scores	Linear model	r = 0.214, p = .021
	ADABoost	r = 0.162, p = .081
Total scores and week 4 response	Linear model	r = 0.560, p < .001
	ADABoost	r = 0.471, p < .001

Outcome at follow-up (n = 52)		
Predictors	Regressor	r and p in CV
Total scores	Linear model	r = 0.309, p = .026
	ADABoost	r = 0.256, p = .067
Total scores and week 4 response	Linear model	r = 0.433, p = .001
	ADABoost	r = 0.197, p = .162

r and p in CV: Pearson correlation coefficient and its significance between true outcome of test patients and their predicted outcomes in the cross-validation

Table 3. Included predictors using total scores and linear regression and their correlation with treatment outcome.

Predictor	Correlation with outcome at discharge	Correlation with outcome at follow-up
Posttraumatic Cognitions Inventory total score	0.277	0.242
Centrality Of Event total score	0.202	0.141
Beck Depression Inventory total score	0.201	0.307
Gender (0: male; 1: female)	0.173	−0.195
Brief Symptom Inventory total score	0.151	0.171
Posttraumatic Diagnostic Scale total score	0.136	−0.022
Comorbid affective disorder (0: no; 1: yes)	0.089	0.426
Index for the Assessment of Participation Impairments total score	0.073	0.189
Perseverative Thinking Questionnaire total score	0.066	0.026
Davidson Trauma Scale total score	0.020	−0.008
Age	−0.126	0.205
Comorbid substance use disorder (0: no; 1: yes)	−0.173	−0.066

severity allows greater improvements. Other authors have discussed the possibility of over-reporting symptoms in self-assessments at the beginning of treatments (Forbes et al., 2003; Karatzias et al., 2007). In our study, we corrected for baseline symptom severity by using the percentage instead of absolute improvement as our outcome variable.

No relevant improvements in predictions could be achieved in our study when a more advanced non-linear regression method was used instead of a conventional linear regression. On the contrary, the results for predicting the sustained response after 6 months tended to be worse, which in our data set is probably due to the small number of patients with follow-up assessment. Nevertheless, the crucial step in trying to improve predictions of the success of PTSD therapies before therapy

start appears to be the identification of more powerful predictive variables. Here, a combined approach might be fruitful, which incorporates certain clinical features such as treatment outcome expectation (Constantino, Visla, Coyne, & Boswell, 2018) as well as biological measures such as cue-elicited responses (Norrholm et al., 2016) and glucocorticoid sensitivity measures (Yehuda et al., 2014). Ideally, future studies would not only examine the main effects of these predictors on therapy success, but also investigate interaction effects with the form of therapy: While it would already be helpful to be able to predict the success (or failure) of a specific therapy before its start, an investigation of interaction effects (i.e. to look for variables, which predict a good treatment outcome with one but a bad treatment outcome with another possible therapy form) would more directly support treatment recommendations (see (Deisenhofer et al., 2018)) for first attempts in this research direction showing that guidance of treatment in the context of two highly effective treatment according to patient characteristics might improve the therapy response).

We found that including early treatment response at week 4 as an additional predictor allowed substantially more accurate predictions of symptom improvement at discharge. This finding fits into mounting evidence, which identifies the early phase of therapy as prognosis-deciding in various psychiatric disorders (Henkel et al., 2009; Katz et al., 2009; Koran et al., 1995; Lewis et al., 2012; Lin et al., 2019; Lutz et al., 2009; Mulder et al., 2006; Nierenberg et al., 1995; Schindler et al., 2013; Stamm et al., 2014; Szegedi et al., 2003; Van et al., 2008). However, additional analyses revealed that improvements *after* the week 4 assessment were independent from improvements up to the week 4 assessment (early response). Hence, while it is possible to determine to a certain extent after 4 weeks whether a patient will be better at the end of the therapy, no reliable statement can be made about the exact extent of the improvements to be expected in the further course. For the future, it would be important to clarify whether a successful prediction of the outcome at discharge by early response is based on the fact that there is actually an association between early response and response in the late course of therapy, or whether (as in our case) the early therapeutic success only persists until the end of the therapy. Only in the first case would it be possible to use the early therapy response to plan and modify the later therapy phases.

A strength of our study is the application cross-validation scheme, which helps to reduce spurious findings of relationships due to multiple potential predictors. In addition, the study's naturalistic design suggests a high generalizability to similar patient populations.

A limitation of our study is that because all measures were gathered as part of routine clinical practice, we used a self-report measure (DTS) as our primary treatment outcome. However, usage of self-report measures like DTS or PTSD Checklist (PCL) for treatment response is common in research on predictors of PTSD psychotherapy (e.g. (Gros, Yoder, Tuerk, Lozano, & Acierno, 2011; Richardson et al., 2011; van Minnen et al., 2002)). A second limitation, also due to the naturalistic character of the study, is the lack of a control group, which makes us unable to determine whether changes were partly due to time or to non-specific treatment components, and the lack of a formal control of therapy adherence. Despite having included many potential predictors, we might have not assessed some other important variables. For instance, previous studies have found employment-related issues as well as social problems to be associated with treatment outcome (Ehlers et al., 2013; Herzog et al., 2021; Sonne et al., 2016). The rather small sample size of the study is another limitation. In addition, a relatively large number of patients could not be included in the analyses because of missing predictors or because they discontinued therapy. We would argue that the missing predictors at baseline were primarily due to technical causes that were not systematically related to patient characteristics. We were also able to show that the completers and non-completers of the therapy did not differ significantly in key variables such as age, gender, symptom severity and depressive symptoms at baseline. Nevertheless, we cannot rule out a certain bias in the selection of the included patients with certainty.

We found large pre-to-post and large pre-to-follow up improvements in PTSD symptoms and our study therefore adds evidence that trauma-focused treatment can be effectively implemented in routine clinical care for patients with a wide range of traumatic events. Other studies have shown medium (Herzog et al., 2021) to large (Kratzer et al., 2019; Steil, Dyer, Priebe, Kleindienst, & Bohus, 2011) pre-to-post effect sizes in inpatient naturalistic PTSD samples. However, our treatment programme, which mainly followed the cognitive processing therapy manual, did not perform as well as CPT has in randomized controlled trials (Asmundson et al., 2019; Cusack et al., 2016; Watts et al., 2013). This might be due to the naturalistic sample with consecutively admitted patients who had received previous treatments and still met PTSD diagnosis. We do not know how the addition of behavioural experiments and the monitoring of the reduction of avoidance and seeking-safety behaviour has affected treatment outcome in our study.

In summary, our results show the possibility of predicting the success of PTSD psychotherapy before starting therapy and that these predictions can be improved by monitoring the early response to therapy. Due to the small predictive power of pre-treatment characteristic seen in our and also in prior work (Gros et al., 2011; Haagen, Smid, Knipscheer, & Kleber, 2015; Richardson et al., 2011; van Minnen et al., 2002) we would argue that no patients (regardless of age, PTSD severity, comorbid disorders, etc.) should be excluded

from therapy per se as a favourable outcome is conceivable regardless of these patient characteristics. Before prediction can be used clinically, the accuracy of the predictions still needs to be increased, especially through research into more powerful predictors.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to containing information that could compromise the privacy of research participants.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Heiner Stuke is participant in the Charité Clinician Scientist Programme funded by the Charité – Universitätsmedizin Berlin and the Berlin Institute of Health;Berlin Institute of Health [Junior Clinician Scientist Grant]. We acknowledge support from the German Research Foundation (DFG) and the Open Access Publication Fund of Charité – Universitätsmedizin Berlin.

References

Asmundson, G. J. G., Thorisdottir, A. S., Roden-Foreman, J. W., Baird, S. O., Witcraft, S. M., Stein, A. T., … Powers, M. B. (2019). A meta-analytic review of cognitive processing therapy for adults with posttraumatic stress disorder. *Cognitive Behaviour Therapy*, *48*(1), 1–14. doi:10.1080/16506073.2018.1522371

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). Beck depression inventory-II. *San Antonio*, *78*(2), 490–498. doi:10.1037/t00742-000

Berliner, L., Bisson, J., Cloitre, M., Forbes, D., Goldbeck, L., Jensen, T., … Shapiro, F. (2019). Position paper on complex PTSD in children and adolescents. *International Society for Traumatic Stress Studies Guidelines*. https://istss.org/getattachment/Treating-Trauma/New-ISTSS-Prevention-and-Treatment-Guidelines/ISTSS_CPTSD-Position-Paper-(Child_Adol)_FNL.pdf.aspx

Berntsen, D., & Rubin, D. C. (2006). The centrality of event scale: A measure of integrating a trauma into one's identity and its relation to post-traumatic stress disorder symptoms. *Behaviour Research and Therapy*, *44*(2), 219–231. doi:10.1016/j.brat.2005.01.009

Bradley, R., Greene, J., Russ, E., Dutra, L., & Westen, D. (2005). A multidimensional meta-analysis of psychotherapy for PTSD. *American Journal of Psychiatry*, *162*(2), 214–227. doi:10.1176/appi.ajp.162.2.214

Brown, L., Belli, G., Asnaani, A., & Foa, E. (2019). A review of the role of negative cognitions about oneself, others, and the world in the treatment of PTSD. *Cognitive Therapy and Research*, *43*(1), 143–173. doi:10.1007/s10608-018-9938-1

Burges, C. J. (2010). *Dimension reduction: A guided tour*. Hanover, MA: Now Publishers Inc.

Clifton, E. G., Feeny, N. C., & Zoellner, L. A. (2017). Anger and guilt in treatment for chronic posttraumatic stress disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, *54*, 9–16. doi:10.1016/j.jbtep.2016.05.003

Constantino, M. J., Visla, A., Coyne, A. E., & Boswell, J. F. (2018). A meta-analysis of the association between patients' early treatment outcome expectation and their posttreatment outcomes. *Psychotherapy (Chic)*, *55*(4), 473–485. doi:10.1037/pst0000169

Courtois, C. A., Brown, L. S., Cook, J., Fairbank, J., Friedman, M., Gone, J., … Roberts, J. (2017). Clinical practice guideline for the treatment of posttraumatic stress disorder (PTSD). *American Psychological Association Guideline*. https://www.apa.org/ptsd-guideline/ptsd.pdf

Cusack, K., Jonas, D. E., Forneris, C. A., Wines, C., Sonis, J., Middleton, J. C., … Gaynes, B. N. (2016). Psychological treatments for adults with posttraumatic stress disorder: A systematic review and meta-analysis. *Clinical Psychology Review*, *43*, 128–141. doi:10.1016/j.cpr.2015.10.003

Davidson, J. R., Book, S. W., Colket, J. T., Tupler, L. A., Roth, S., David, D., … Feldman, M. E. (1997). Assessment of a new self-rating scale for post-traumatic stress disorder. *Psychological Medicine*, *27*(1), 153–160. doi:10.1017/S0033291796004229

Deck, R., Mittag, O., Hüppe, A., Muche-Borowski, C., & Raspe, H. (2011). IMET-Index zur Messung von Einschränkungen der Teilhabe. doi:10.23668/psycharchives.381

Deisenhofer, A. K., Delgadillo, J., Rubel, J. A., Böhnke, J. R., Zimmermann, D., Schwartz, B., & Lutz, W. (2018). Individual treatment selection for patients with posttraumatic stress disorder. *Depression and Anxiety*, *35*(6), 541–550. doi:10.1002/da.22755

Drucker, H. (1997). Improving regressors using boosting techniques. *International Conference on Machine Learning*. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.31.314&rep=rep1&type=pdf

Ehlers, A., Clark, D. M., Hackmann, A., McManus, F., & Fennell, M. (2005). Cognitive therapy for post-traumatic stress disorder: Development and evaluation. *Behaviour Research and Therapy*, *43*(4), 413–431. doi:10.1016/j.brat.2004.03.006

Ehlers, A., Grey, N., Wild, J., Stott, R., Liness, S., Deale, A., … Clark, D. M. (2013). Implementation of cognitive therapy for PTSD in routine clinical care: Effectiveness and moderators of outcome in a consecutive sample. *Behaviour Research and Therapy*, *51*(11), 742–752. doi:10.1016/j.brat.2013.08.006

Ehring, T., Zetsche, U., Weidacker, K., Wahl, K., Schonfeld, S., & Ehlers, A. (2011). The Perseverative Thinking Questionnaire (PTQ): Validation of a content-independent measure of repetitive negative thinking. *Journal of Behavior Therapy and Experimental Psychiatry*, *42*(2), 225–232. doi:10.1016/j.jbtep.2010.12.003

Foa, E. B., Cashman, L., Jaycox, L., & Perry, K. (1997). The validation of a self-report measure of posttraumatic stress disorder: The Posttraumatic Diagnostic Scale. *Psychological Assessment*, *9*(4), 445–451. doi:10.1037/1040-3590.9.4.445

Foa, E. B., Ehlers, A., Clark, D. M., Tolin, D. F., & Orsillo, S. M. (1999). The posttraumatic cognitions inventory (PTCI): Development and validation. *Psychological Assessment*, *11*(3), 303–314. doi:10.1037/1040-3590.11.3.303

Forbes, D., Creamer, M., Hawthorne, G., Allen, N., & McHugh, T. (2003). Comorbidity as a predictor of

symptom change after treatment in combat-related posttraumatic stress disorder. *The Journal of Nervous and Mental Disease*, *191*(2), 93–99. doi:10.1097/01.NMD.0000051903.60517.98

Gomez de La Cuesta, G., Schweizer, S., Diehle, J., Young, J., & Meiser-Stedman, R. (2019). The relationship between maladaptive appraisals and posttraumatic stress disorder: A meta-analysis. *European Journal of Psychotraumatology*, *10*(1), 1620084. doi:10.1080/20008198.2019.1620084

Graham, B., Garcia, N. M., Burton, M. S., Cooper, A. A., Roy-Byrne, P. P., Mavissakalian, M. R., … Zoellner, L. A. (2018). High expectancy and early response produce optimal effects in sertraline treatment for post-traumatic stress disorder. *The British Journal of Psychiatry*, *213*(6), 704–708. doi:10.1192/bjp.2018.211

Groleau, J., Calhoun, L., Cann, A., & Tedeschi, R. (2013). The role of centrality of events in posttraumatic distress and posttraumatic growth. *Psychological Trauma: Theory, Research, Practice, and Policy*, *5*(5), 477. doi:10.1037/a0028809

Gros, D. F., Yoder, M., Tuerk, P. W., Lozano, B. E., & Acierno, R. (2011). Exposure therapy for PTSD delivered to veterans via telehealth: Predictors of treatment completion and outcome and comparison to treatment delivered in person. *Behavior Therapy*, *42*(2), 276–283. doi:10.1016/j.beth.2010.07.005

Haagen, J. F., Smid, G. E., Knipscheer, J. W., & Kleber, R. J. (2015). The efficacy of recommended treatments for veterans with PTSD: A metaregression analysis. *Clinical Psychology Review*, *40*, 184–194. doi:10.1016/j.cpr.2015.06.008

Hagenaars, M. A., van Minnen, A., & Hoogduin, K. A. (2010). The impact of dissociation and depression on the efficacy of prolonged exposure treatment for PTSD. *Behaviour Research and Therapy*, *48*(1), 19–27. doi:10.1016/j.brat.2009.09.001

Henkel, V., Seemuller, F., Obermeier, M., Adli, M., Bauer, M., Mundt, C., … Riedel, M. (2009). Does early improvement triggered by antidepressants predict response/remission? Analysis of data from a naturalistic study on a large sample of inpatients with major depression. *Journal of Affective Disorders*, *115*(3), 439–449. doi:10.1016/j.jad.2008.10.011

Herzog P., Voderholzer U., Gärtner T., Osen B., Svitak M., Doerr R., … Brakemeier E. L.… (2021). Predictors of outcome during inpatient psychotherapy for posttraumatic stress disorder: A single-treatment, multi-site, practice-based study. *Psychotherapy Research*, *31*(4), 468–482. doi:10.1080/10503307.2020.1802081

Hilbert, K., Kunas, S. L., Lueken, U., Kathmann, N., Fydrich, T., & Fehm, L. (2020). Predicting cognitive behavioral therapy outcome in the outpatient sector based on clinical routine data: A machine learning approach. *Behaviour Research and Therapy*, *124*, 103530. doi:10.1016/j.brat.2019.103530

Hiller, W., Schindler, A. C., & Lambert, M. J. (2012). Defining response and remission in psychotherapy research: A comparison of the RCI and the method of percent improvement. *Psychotherapy Research*, *22*(1), 1–11. doi:10.1080/10503307.2011.616237

Jun, J. J., Zoellner, L. A., & Feeny, N. C. (2013). Sudden gains in prolonged exposure and sertraline for chronic PTSD. *Depression and Anxiety*, *30*(7), 607–613. doi:10.1002/da.22119

Karatzias, A., Power, K., McGoldrick, T., Brown, K., Buchanan, R., Sharp, D., & Swanson, V. (2007). Predicting treatment outcome on three measures for post-traumatic stress disorder. *European Archives of Psychiatry and Clinical Neuroscience*, *257*(1), 40–46. doi:10.1007/s00406-006-0682-2

Karatzias, T., Murphy, P., Cloitre, M., Bisson, J., Roberts, N., Shevlin, M., … Hutton, P. (2019). Psychological interventions for ICD-11 complex PTSD symptoms: Systematic review and meta-analysis. *Psychological Medicine*, *49*(11), 1761–1775. doi:10.1017/S0033291719000436

Karin, E., Dear, B. F., Heller, G. Z., Gandy, M., & Titov, N. (2018). Measurement of Symptom Change Following Web-Based Psychotherapy: Statistical Characteristics and Analytical Methods for Measuring and Interpreting Change. *JMIR Mental Health*, *5*(3), e10200. doi:10.2196/10200

Katz, M. M., Meyers, A. L., Prakash, A., Gaynor, P. J., & Houston, J. P. (2009). Early symptom change prediction of remission in depression treatment. *Psychopharmacol Bull*, *42*(1), 94–107.

Kearns, M., & Ron, D. (1999). Algorithmic stability and sanity-check bounds for leave-one-out cross-validation. *Neural Computation*, *11*(6), 1427–1453. doi:10.1162/089976699300016304

Koenen, K. C., Ratanatharathorn, A., Ng, L., McLaughlin, K. A., Bromet, E. J., Stein, D. J., … Kessler, R. C. (2017). Posttraumatic stress disorder in the World Mental Health Surveys. *Psychological Medicine*, *47*(13), 2260–2274. doi:10.1017/S0033291717000708

Koran, L. M., Hamilton, S. H., Hertzman, M., Meyers, B. S., Halaris, A. E., Tollefson, G. D., Downs, J. M., Folks, D. G., Jeste, D. V., Lazarus, L. W., & Satlin, A.… (1995). Predicting response to fluoxetine in geriatric patients with major depression. *Journal of Clinical Psychopharmacology*, *15*(6), 421–427. doi:10.1097/00004714-199512000-00006

Kratzer, L., Heinz, P., Schennach, R., Schiepek, G. K., Padberg, F., & Jobst, A. (2019). [Inpatient Treatment of Complex PTSD Following Childhood Abuse: Effectiveness and Predictors of Treatment Outcome]. *Psychother Psychosom Med Psychol*, *69*(3–4), 114–122. doi:10.1055/a-0591-3962

Lewis, C., Roberts, N. P., Andrew, M., Starling, E., & Bisson, J. I. (2020). Psychological therapies for post-traumatic stress disorder in adults: Systematic review and meta-analysis. *European Journal of Psychotraumatology*, *11*(1), 1729633. doi:10.1080/20008198.2020.1729633

Lewis, C. C., Simons, A. D., & Kim, H. K. (2012). The role of early symptom trajectories and pretreatment variables in predicting treatment response to cognitive behavioral therapy. *Journal of Consulting and Clinical Psychology*, *80*(4), 525–534. doi:10.1037/a0029131

Lin, C. H., Park, C., & McIntyre, R. S. (2019). Early improvement in HAMD-17 and HAMD-7 scores predict response and remission in depressed patients treated with fluoxetine or electroconvulsive therapy. *Journal of Affective Disorders*, *253*, 154–161. doi:10.1016/j.jad.2019.04.082

Lindebo Knutsen, M., Sachser, C., Holt, T., Goldbeck, L., & Jensen, T. K. (2020). Trajectories and possible predictors of treatment outcome for youth receiving trauma-focused cognitive behavioral therapy. *Psychological Trauma: Theory, Research, Practice, and Policy*, *12*(4), 336–346. doi:10.1037/tra0000482

Litz, B. T., Berke, D. S., Kline, N. K., Grimm, K., Rusowicz-Orazem, L., Resick, P. A., … Peterson, A. L. (2019). Patterns and predictors of change in trauma-focused treatments for war-related posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology*, *87*(11), 1019–1029. doi:10.1037/ccp0000426

Lutz, W., Stulz, N., & Köck, K. (2009). Patterns of early change and their relationship to outcome and follow-up among patients with major depressive disorders. *Journal of Affective Disorders*, *118*(1–3), 60–68. doi:10.1016/j.jad.2009.01.019

Moser, J. S., Cahill, S. P., & Foa, E. B. (2010). Evidence for poorer outcome in patients with severe negative trauma-related cognitions receiving prolonged exposure plus cognitive restructuring: Implications for treatment matching in posttraumatic stress disorder. *Journal of Nervous & Mental Disease*, *198*(1), 72–75. doi:10.1097/NMD.0b013e3181c81fac

Moulds, M. L., Bisby, M. A., Wild, J., & Bryant, R. A. (2020). Rumination in posttraumatic stress disorder: A systematic review. *Clinical Psychology Review*, *82*, 101910. doi:10.1016/j.cpr.2020.101910

Mulder, R. T., Joyce, P. R., Frampton, C. M., Luty, S. E., & Sullivan, P. F. (2006). Six months of treatment for depression: Outcome and predictors of the course of illness. *American Journal of Psychiatry*, *163*(1), 95–100. doi:10.1176/appi.ajp.163.1.95

Nierenberg, A. A., McLean, N. E., Alpert, J. E., Worthington, J. J., Rosenbaum, J. F., & Fava, M. (1995). Early nonresponse to fluoxetine as a predictor of poor 8-week outcome. *Am J Psychiatry*, *152*(10), 1500–1503. doi:10.1176/ajp.152.10.1500

Norrholm, S. D., Jovanovic, T., Gerardi, M., Breazeale, K. G., Price, M., Davis, M., . . . Rothbaum, B. O. (2016). Baseline psychophysiological and cortisol reactivity as a predictor of PTSD treatment outcome in virtual reality exposure therapy. *Behaviour Research and Therapy*, *82*, 28–37. doi:10.1016/j.brat.2016.05.002

Ostacher, M. J., & Cifu, A. S. (2019). Management of Posttraumatic Stress Disorder. *JAMA*, *321*(2), 200–201. doi:10.1001/jama.2018.19290

Ramos-Lima, L. F., Waikamp, V., Antonelli-Salgado, T., Passos, I. C., & Freitas, L. H. M. (2020). The use of machine learning techniques in trauma-related disorders: A systematic review. *Journal of Psychiatric Research*, *121*, 159–172. doi:10.1016/j.jpsychires.2019.12.001

Rath, J. F., & Fox, L. M. (2018). Brief Symptom Inventory. In J. S. Kreutzer, J. DeLuca, & B. Caplan (Eds.), *Encyclopedia of Clinical Neuropsychology* (pp. 633–636). Cham: Springer International Publishing.

Resick, P. A., Monson, C. M., & Chard, K. M. (2016). *Cognitive processing therapy for PTSD: A comprehensive manual*. New York, NY: Guilford Publications.

Richardson, J. D., Elhai, J. D., & Sarreen, J. (2011). Predictors of treatment response in Canadian combat and peacekeeping veterans with military-related posttraumatic stress disorder. *Journal of Nervous & Mental Disease*, *199*(9), 639–645. doi:10.1097/NMD.0b013e318229ce7b

Schindler, A., Hiller, W., & Witthöft, M. (2013). What predicts outcome, response, and drop-out in CBT of depressive adults? A naturalistic study. *Behavioural and Cognitive Psychotherapy*, *41*(3), 365–370. doi:10.1017/S1352465812001063

Sonne, C., Carlsson, J., Bech, P., Vindbjerg, E., Mortensen, E. L., & Elklit, A. (2016). Psychosocial predictors of treatment outcome for trauma-affected refugees. *European Journal of Psychotraumatology*, *7*(1), 30907. doi:10.3402/ejpt.v7.30907

Stamm, T. J., Becker, D., Sondergeld, L. M., Wiethoff, K., Hiemke, C., O'Malley, G., . . . Adli, M. (2014). Prediction of antidepressant response to venlafaxine by a combination of early response assessment and therapeutic drug monitoring. *Pharmacopsychiatry*, *47*(4–5), 174–179. doi:10.1055/s-0034-1383565

Steil, R., Dyer, A., Priebe, K., Kleindienst, N., & Bohus, M. (2011). Dialectical behavior therapy for posttraumatic stress disorder related to childhood sexual abuse: A pilot study of an intensive residential treatment program. *Journal of Traumatic Stress*, *24*(1), 102–106. doi:10.1002/jts.20617

Szegedi, A., Muller, M. J., Anghelescu, I., Klawe, C., Kohnen, R., & Benkert, O. (2003). Early improvement under mirtazapine and paroxetine predicts later stable response and remission with high sensitivity in patients with major depression. *The Journal of Clinical Psychiatry*, *64*(4), 413–420. doi:10.4088/JCP.v64n0410

Thompson, B. (1994). The pivotal role of replication in psychological research: Empirically evaluating the replicability of sample results. *Journal of Personality*, *62*(2), 157–176. doi:10.1111/j.1467-6494.1994.tb00289.x

Van, H. L., Schoevers, R. A., Kool, S., Hendriksen, M., Peen, J., & Dekker, J. (2008). Does early response predict outcome in psychotherapy and combined therapy for major depression? *Journal of Affective Disorders*, *105*(1–3), 261–265. doi:10.1016/j.jad.2007.04.016

van Minnen, A., Arntz, A., & Keijsers, G. P. (2002). Prolonged exposure in patients with chronic PTSD: Predictors of treatment outcome and dropout. *Behaviour Research and Therapy*, *40*(4), 439–457. doi:10.1016/S0005-7967(01)00024-9

Watts, B. V., Schnurr, P. P., Mayo, L., Young-Xu, Y., Weeks, W. B., & Friedman, M. J. (2013). Meta-analysis of the efficacy of treatments for posttraumatic stress disorder. *The Journal of Clinical Psychiatry*, *74*(6), e541–550. doi:10.4088/JCP.12r08225

Yehuda, R., Hoge, C. W., McFarlane, A. C., Vermetten, E., Lanius, R. A., Nievergelt, C. M., . . . Hyman, S. E. (2015). Post-traumatic stress disorder. *Nat Rev Dis Primers*, *1*, 15057. doi:10.1038/nrdp.2015.57

Yehuda, R., Pratchett, L. C., Elmes, M. W., Lehrner, A., Daskalakis, N. P., Koch, E., . . . Bierer, L. M. (2014). Glucocorticoid-related predictors and correlates of post-traumatic stress disorder treatment response in combat veterans. *Interface Focus*, *4*(5), 20140048. doi:10.1098/rsfs.2014.0048