



Universiteit
Leiden

The Netherlands

Choosing the right track: improving PTSD treatment outcomes for patients with childhood abuse-related posttraumatic stress disorder

Hoeboer, C.M.

Citation

Hoeboer, C. M. (2022, January 18). *Choosing the right track: improving PTSD treatment outcomes for patients with childhood abuse-related posttraumatic stress disorder*. Retrieved from <https://hdl.handle.net/1887/3249982>

Version: Publisher's Version


License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3249982>

Note: To cite this publication please use the final published version (if applicable).

Chapter 6

Personalization of treatment for patients with childhood abuse-related posttraumatic stress disorder



Published as: Hoeboer C.M., Oprel D.A.C., Kleine R.A.D., Schwartz B., Deisenhofer A.-K., Schoorl M., Van Der Does W.A.J., van Minnen A., & Lutz W. (2021). Personalization of Treatment for Patients with Childhood-Abuse-Related Posttraumatic Stress Disorder. *Journal of Clinical Medicine*, 10, 4522.

Abstract

Background: Differences in effectiveness among treatments for Posttraumatic Stress Disorder (PTSD) are typically small. Given the variation between patients in treatment response, personalization offers a new way to improve treatment outcome. The aim of this study was to identify predictors of psychotherapy outcome in PTSD and to combine these into a Personalized Advantage Index (PAI).

Methods: We used data from a recent randomized controlled trial comparing prolonged exposure (PE; $n = 48$), intensified PE (iPE; $n = 51$) and skills training (STAIR) followed by PE ($n = 50$) in 149 patients with Childhood Abuse-related PTSD (CA-PTSD). Outcome measures were clinician-assessed and self-reported PTSD symptoms. Predictors were identified in the exposure therapies (PE and iPE) and STAIR+PE separately using random forests and subsequent bootstrap procedure. Next, these predictors were used to calculate PAI and to determine optimal and suboptimal treatment in a leave-one-out cross-validation approach.

Results: More depressive symptoms, less social support, more axis-1 diagnoses and higher severity of childhood sexual abuse were predictors of worse treatment outcomes in PE and iPE. More emotion regulation difficulties, lower general health status and higher baseline PTSD symptoms were predictors of worse treatment outcomes in STAIR+PE. Randomization to optimal treatment based on these predictors resulted in more improvement than suboptimal treatment in clinician assessed (Cohens' $d = .55$) and self-reported PTSD symptoms (Cohens' $d = .47$).

Conclusion: Personalization based on PAI is a promising tool to improve therapy outcome in patients with CA-PTSD. Further studies are needed to replicate findings in prospective studies.

Keywords: Posttraumatic stress disorder, STAIR+PE, prolonged exposure therapy, personalized advantage index, predictors treatment outcome.

Introduction

Despite the well-established effectiveness of treatments for PTSD such as trauma-focused cognitive behavioral therapy (TF-CBT; Mavranouzouli et al., 2020), meta-analyses showed that about half of the patients do not benefit (enough) from treatment or dropout prematurely (Bradley et al., 2005; Ehring et al., 2014; Lewis et al., 2020; Watkins et al., 2018). For the past decades, research has focused on developing new treatments (e.g., Wilson, Becker, & Tinker, 1995) or adapting already existing ones (e.g., Cloitre et al., 2002; Hendriks et al., 2018). This has led to new effective treatments, but failed to improve treatment outcome (Mavranouzouli et al., 2020; Zhou et al., 2020). Given these alternative treatment options, *personalization* offers a new approach towards improving PTSD treatment outcome. The basic idea is that patients might respond differently to two distinct treatments. Hence, investigating which patients are most likely to benefit from one treatment compared to another may improve individual patient outcomes (Seidler & Wagner, 2006). Clinicians already use personalization to some degree on an intuitive level since treatments indications are often based on patient characteristics (e.g., Becker et al., 2004). However, intuition is prone to biases and this approach is unsystematic and not evidence-based (Perlis, 2016; Waller, 2009). In contrast, personalization based on statistical algorithms might result in systematic and empirically derived treatment recommendations.

Treatment personalization of PTSD has received little attention compared to other fields (e.g. medicine). There have been three studies that investigated treatment personalization in patients with PTSD. Two studies used a Personalized advantage Index (PAI) which indicates relative preference for one treatment compared to another based on a combination of predictors or moderators of treatment outcome (Deisenhofer et al., 2018; Keefe et al., 2018). Both studies found that the PAI approach led to relevant treatment recommendations with medium effect sizes. Deisenhofer et al. (2018) compared trauma-focused cognitive behavioral therapy (TF-CBT) with eye movement desensitization and reprocessing (EMDR) and used depressive symptoms as outcome. They found that age, employment status, gender, and functional impairment were predictors of outcome in TF-CBT and baseline depressive symptoms and prescribed antidepressant medication were predictors of outcome in EMDR. Keefe et al. (2018) compared Prolonged Exposure (PE) with Cognitive Processing Therapy (CPT) and used drop-out as outcome. They assessed moderators of treatment outcome rather than predictors in the two treatment separately and found that childhood physical abuse, current relationship conflict, anger and being a racial minority moderated treatment outcome. The third study used generated modifiers (Petkova, Park, Ciarleglio, Ogden, & Tarpey, 2020), a composite moderator indicating differential treatment outcome in a support condition followed by PE (support+PE), skills training (STAIR) and skills training followed by exposure (STAIR+PE) in patients with Childhood Abuse-related PTSD (CA-PTSD; Cloitre et al., 2016). They used clinician-assessed PTSD symptoms as outcome. They found that the combination of symptom burden and emotion regulation might be relevant for personalization, but did not evaluate whether this led to relevant treatment recommendations (Cloitre et al., 2016).

To summarize, personalization offers a promising approach for PTSD treatment, but so far no study evaluated its relevance for treatment recommendations using PTSD symptoms as outcome while this is the primary focus of treatment. Furthermore, most studies only assessed a limited number of potential predictors, which does not capture the heterogeneous symptom representation of patients with PTSD. In the current study, we aimed to develop and evaluate treatment personalization in patients with CA-PTSD using PAI based on a broad range of patient characteristics including both self-reported and clinician assessed characteristics. We used a sample of 149 patients randomized to an exposure only condition (PE and intensified PE (iPE)) or STAIR+PE. Our first aim was to identify which patient characteristics were predictors of treatment outcome in the exposure only conditions (PE and iPE) and STAIR+PE separately. Our second aim was to calculate the PAI based on these predictors and evaluate whether optimal treatment according to the PAI resulted in better treatment outcome compared to suboptimal treatment.

Methods

This study used the data of a randomized controlled trial investigating three psychotherapies of CA-PTSD (Opel et al.; Opel et al., 2018). A total of 149 patients were recruited in two outpatient mental health services in The Hague and Rotterdam, the Netherlands. These patients were randomized to PE ($n = 48$), intensified PE ($n = 51$) or STAIR+PE ($n = 50$).

Participants

Inclusion criteria of the original study sample included: age between 18 and 65 years; PTSD diagnosis according to the DSM-5 established with the Clinician Administered PTSD Scale for DSM-5 (CAPS-5; Boeschoten et al., 2018); at least moderate severity of PTSD symptoms (CAPS-5 score ≥ 26) and a specific memory of the traumatic event. Exclusion criteria included: ongoing compensation case or legal procedures about admission or stay in The Netherlands; pregnancy; severe non-suicidal self-injury, which required hospitalization during the past three months; severe suicidal behavior in the past three months; severe disorder in the use of alcohol or drugs in the past three months; cognitive impairment (estimated IQ < 70); changes in psychotropic medication in the two months prior to inclusion; and engagement in any current psychological treatment. See Table 1 for sample characteristics. The trial was approved by the Medical Ethics Committee of Leiden University Medical Center (NL57984.058.16).

Procedures

Written informed consent was obtained from all participants before the baseline assessment when patients received all relevant information and decided to participate. Patients were randomized in a 1:1:1 ratio to PE, iPE and STAIR+PE. Predictors were assessed during the baseline assessment (T0). PTSD symptoms were assessed at baseline (T0), after four weeks (T1), eight weeks (T2) and post-treatment after 16 weeks (T3). Clinical interviews were carried out by independent interviewers who were blind to the treatment condition of patients. The authors assert that all procedures contributing to this work comply with the

ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Treatment

PE included 16 weekly sessions of 90 minutes and consisted of a combination of imaginal exposure and exposure in vivo (Foa et al., 2007). iPE included 12 sessions, three times a week (4 weeks total), followed by two booster sessions after one and two months respectively. Treatment protocols of PE and iPE were identical. STAIR+PE included 16 weekly sessions of which the first half consisted of 60 minutes STAIR and the second half consisted of 90 minutes PE. STAIR sessions included skills training in emotion regulation and interpersonal functioning. PE sessions were similar to the PE and iPE conditions.

Measures

Outcome measures

PTSD symptom severity measured with the CAPS-5 (Weathers et al., 2013a) was the primary outcome of this study. The CAPS-5 includes 20 items on a 5-point Likert-scale resulting in a total score between 0 and 80 (Cronbach's α current study = .75).

Self-reported PTSD symptom severity measured with the PTSD checklist for DSM-5 (PCL-5; Blevins et al., 2015) was the secondary outcome of this study. The PCL-5 includes 20 items on a 5-point Likert-scale resulting in a total score between 0 and 80 (Cronbach's α current study = .89).

Predictor variables

Patient expectancies

Patients' expectancies of the treatments were indicated by two predictors: total score of the Expectancy of burden (Cronbach's α current study = .91) and Credibility questionnaire (Cronbach's α current study = .90) as used in previous studies (e.g., de Bont et al., 2013). See Table 1 for additional information about predictors.

Demographics

Demographic predictors included age, gender, cultural background education and employment.

Social support

Social support was indicated by the total score of the social support survey from the Medical Outcome Study (MOS; Kempen, 1992; Cronbach's α current study = .97).

Trauma background

We included four subscale scores of the Childhood Trauma Questionnaire (CTQ; Bernstein & Fink, 1998) as indicators of childhood trauma background: childhood emotional abuse (Cronbach's α current study = .86), emotional neglect (Cronbach's α current study = .86), physical abuse (Cronbach's α current study = .88) and sexual abuse (Cronbach's α current study = .88).

Table 1. Descriptive information about potential predictors for exposure therapies and STAIR+PE.

Predictors¹	Possible range of predictor scores min-max	Exposure therapies (<i>n</i> = 99) Mean (SD) or %	STAIR+PE (<i>n</i> = 50) Mean (SD) or %
Patient expectancies			
Expected burden	0-10	5.98 (2.56)	6.73 (2.37)
Credibility	0-10	6.75 (1.89)	6.72 (1.74)
Demographics			
Age, y		36.76 (11.47)	37.07 (12.39)
Gender, female		75.76	78.00
Cultural background, western		39.39	52.00
Education, high		21.21	18.00
Employment, yes		40.40	34.00
Social support			
MOS total score	1-5	3.41 (1.10)	3.32 (1.04)
Trauma background			
CTQ childhood emotional abuse	5-25	17.06 (6.04)	17.54 (6.21)
CTQ childhood emotional neglect	5-25	17.74 (5.08)	19.84 (5.38)
CTQ childhood physical abuse	5-25	13.09 (6.97)	14.42 (6.36)
CTQ childhood sexual abuse	5-25	15.48 (7.12)	15.62 (7.68)
General health status			
EQ-5D-5L general health status	0-100	55.56 (26.31)	58.18 (20.03)
Self-reported psychiatric symptoms			
BDI total score	0-63	33.63 (10.06)	34.88 (11.15)
PTCI total score	33-231	133.26 (36.40)	149.64 (31.64)
IIP total score	0-4	1.65 (.62)	1.70 (.50)
RSES total score	0-30	12.52 (5.84)	11.32 (6.14)
DERS total score	36-180	115.63 (21.27)	117.46 (20.46)
SDQ-5 total score	5-25	6.78 (2.93)	7.64 (3.11)
Psychotropic medication		49.49	44.00
Clinician-assessed psychiatric symptoms and disorders			
Any SCID-2 personality disorder		59.60	62.00
DSP-I total score	0-36	1.78 (3.20)	3.22 (5.65)
Axis-1 MINI diagnoses, excluding PTSD		2.99	3.38
CAPS-5 baseline total score	0-80	40.28 (8.73)	43.56 (10.46)

STAIR+PE = Skills Training in Affective and Interpersonal Regulation + Prolonged Exposure, Min: minimum, max: maximum, CAPS-5: Clinician Administered PTSD Scale, SCID II: Structured Clinical Interview for DSM-IV axis-II personality disorders, DSP-I: Dissociatief Subtype van PTSS interview, CTQ: Childhood Trauma Questionnaire, DERS: Difficulties in Emotion Regulation Scale, BDI: Beck Depression Inventory-II, PTCI: The posttraumatic cognitions inventory, SDQ-5: Somatoform Dissociation Questionnaire-5, IIP: Inventory of Interpersonal Problems, MOS: Medical Outcomes Study, RSES: Rosenberg Self-Esteem Scale, EQ-5D-5L: EuroQoL 5 Dimensions 5 Levels, SD = standard deviation, y = year, *n* = sample size, MINI = Mini-International Neuropsychiatric Interview.

¹Higher scores on predictors indicate higher symptom severity. Exceptions: for social support higher scores indicate more social support, for EQ-5D-5L general health status higher scores indicate better health status and for the CTQ higher scores indicate more severe childhood maltreatment.

General health status

General health status was measured with the visual analogue scale of the EuroQoL 5 Dimensions 5 Levels (EQ-5D-5L; Brooks, 1996; Le et al., 2013)

Self-reported psychiatric symptoms

Depressive symptoms were indicated by the Beck Depression Inventory (BDI; Beck et al., 1996; Cronbach's α current study = .87). Posttraumatic cognitions were indicated by the Posttraumatic Cognitions Inventory (PTCI; Foa et al., 1999; Cronbach's α current study = .94). Interpersonal problems were indicated by the Inventory of Interpersonal Problems (IIP; Barkham et al., 1996; Cronbach's α current study = .87). Self-esteem was indicated by the Rosenberg Self-esteem Scale (RSES; Schmitt & Allik, 2005; Cronbach's α current study = .87). Emotion regulation difficulties were indicated by the Difficulties in Emotion Regulation Scale (DERS; Lee et al., 2016b; Cronbach's α current study = .90). Somatoform dissociation was indicated by the screener version of the Somatoform Dissociation Questionnaire (SDQ-5; Nijenhuis, Spinhoven, VanDyck, VanderHart, & Vanderlinden, 1996; Cronbach's α current study = .71). The use of psychotropic medication was determined using a self-report question.

Clinician-assessed psychiatric symptoms and disorders

Meeting criteria for at least one personality disorders was assessed with the clinical interview for DSM-IV personality disorders (SCID-2; Weertman et al., 2003). Number of DSM-IV defined Axis-1 disorders (excluding PTSD) was assessed with the Mini International Neuropsychiatric Interview (MINI; Sheehan et al., 1998). Dissociation was indicated by Dissociative subtype of PTSD Interview (DSP-I; Eidhof et al., 2019; Cronbach's α current study = .78). PTSD symptom severity at baseline was assessed with the CAPS-5.

Statistical analysis

Outcome

Estimated change in CAPS-5 and PCL-5 scores from baseline to post-treatment in the exposure conditions ($n = 99$) and STAIR+PE ($n = 50$) were outcome variables in the analyses with higher scores indicating larger symptom decrease. They were separately estimated by subtracting the predicted post-treatment score from the baseline score per individual using all available measurements per outcome from baseline to post-treatment in a linear mixed effect model with R package lme4 (Bates et al., 2015). This model included random intercepts and random slopes. This method provides a more reliable indicator of treatment outcome compared to only using observed post-treatment scores (Raudenbush & Bryk, 2002)

Initial predictor selection with Boruta

Predictors of treatment outcome for the exposure conditions (PE and iPE) and STAIR+PE were selected separately out of the total number of potential predictors ($k = 24$) using R package Boruta (Kursa & Rudnicki, 2010). The Boruta algorithm determines the relevance of

predictors by comparing their performance with ‘shadow’ predictors, which are created by randomly shuffling the values of the original predictors. A random forest classifier is performed by developing multiple trees on different bagging samples of the dataset. Importance of shadow and original variables is calculated with Z-scores by dividing the average loss of accuracy of classification caused by random permutations of the variable between samples by its standard deviation. The original variable is a relevant predictor during a round when its Z-score is higher than the maximum shadow variable’s Z-score. This is stored as a hit in a vector. When the number of hits from a predictor is significantly higher or lower than the best shadow variable, the variable is deemed important or unimportant respectively. Unimportant variables are deleted from the dataset. The procedure repeats for maximum 1000 iterations or until all variables are categorized.

Further predictor selection using bootstrap procedure

After identifying predictors of treatment outcome with the Boruta algorithm, we performed a bootstrapped model using R package *bootStepAIC* (Rizopoulos, 2009) and selected the variables of the model with the best model fit. Since the aim of Boruta is to identify all variables which have *any* relevance under some circumstances, further selection ensured that we did not overfit the data. Furthermore, since the PAI is calculated using a linear combination of variables, the bootstrapped AIC approach ensured that we included the best combination of variables to predict outcome in a linear manner.

Personalized advantage index

With the final set of predictors, we calculated the predicted outcome of all patients using a regression model with a leave-one-out cross-validation approach (predicted outcome per patient was based on a training set including all other patients). Treatment outcome of the treatment that patients did not receive was predicted by the model based on the patients of the other condition (so every patient had two predictions in total: one for exposure therapies and one for STAIR+PE). PAI was calculated by subtracting predicted outcome in the STAIR+PE condition from predicted outcome in exposure conditions and indicated relative advantage of exposure conditions over STAIR+PE. When patients had been randomized to their recommended treatment, we defined them as having received optimal treatment versus suboptimal, when they had been randomized to their non-recommended treatment.

Results

The average estimated change in CAPS-5 scores from baseline to post-treatment was not different in the exposure conditions ($M = 21.38$; $SD = 7.90$) compared to STAIR+PE ($M = 20.13$; $SD = 6.75$), while estimated change in PCL-5 scores from baseline to post-treatment was significantly larger in the exposure conditions ($M = 25.82$; $SD = 10.14$) compared to STAIR+PE ($M = 20.16$; $SD = 9.29$).

Variable selection for exposure therapies and STAIR+PE

Figure 1 depict the results of Boruta for exposure conditions and STAIR+PE. Variables dropped in the subsequent bootstrap procedure can be found in the Appendix. For the

6

6



6

Table 2.a Final prediction models of exposure therapies and STAIR+PE with estimated change in CAPS-5 score baseline to post-treatment as outcome variable.

Exposure therapies	Estimate	Std. Error	t-value	p
BDI	-.24	.07	-3.40	< .001
MOS	2.23	.62	3.63	< .001
Axis-1 MINI diagnoses	-.89	.37	-2.42	.02
CTQ sexual abuse	-.18	.09	-2.04	.04
STAIR+PE				
EQ-5D-5L	.07	.04	1.97	.05
DERS	-.10	.04	-2.54	.01
CAPS-5 baseline	-.26	.08	-3.19	.003

Note that prediction models of all individuals differed slightly due to the cross-validation approach.

Table 2.b Final prediction models of exposure therapies and STAIR+PE with estimated change in PCL-5 score baseline to post-treatment as outcome variable.

Exposure therapies	Estimate	Std. Error	t-value	p
BDI	-.26	.10	-2.65	.01
MOS	2.59	.89	2.90	.005
STAIR+PE				
EQ-5D-5L	.11	.06	1.78	.08
DERS	-.16	.06	-2.75	.009

Note that prediction models of all individuals differed slightly due to the cross-validation approach.

STAIR+PE = Skills Training in Affective and Interpersonal Regulation + Prolonged Exposure, CAPS-5: Clinician Administered PTSD Scale, CTQ: Childhood Trauma Questionnaire, DERS: Difficulties in Emotion Regulation Scale, BDI: Beck Depression Inventory-II, MOS: Medical Outcomes Study, EQ-5D-5L: EuroQoL 5 Dimensions 5 Levels, MINI = Mini-International Neuropsychiatric Interview.

Personalized advantage index

The PAI was calculated based on the final models using leave-one-out cross-validation. For the CAPS-5, the average error of the predictions (difference between predicted score based on final models and estimated outcome) was 5.09 ($SD = 7.57$) in the exposure conditions and 4.06 ($SD = 7.25$) in the STAIR+PE conditions. Half of the patients ($n = 75$; 50%) were randomized to their optimal treatment, while $n = 74$ (50%) were not. Patients randomized to their optimal treatment improved more on the CAPS-5 from baseline to post-treatment ($M_{\text{improvement}} = 22.96$; $SD_{\text{improvement}} = 6.99$), compared to patients randomized to their suboptimal treatment ($M_{\text{improvement}} = 18.94$; $SD_{\text{improvement}} = 7.57$; $F(1,147) = 11.36$, $p < .001$). The standardized mean difference between optimal and suboptimal treatments corresponded to a medium effect size (Cohen's $d = .55$ [.23, .88]). For the PCL-5, the average error of the predictions was 7.09 ($SD = 6.16$) in the exposure conditions and 7.24 ($SD = 4.74$) in the STAIR+PE condition. Based on the PCL data, a little more than over half of the patients ($n = 94$; 63%) were randomized to their optimal treatment, while $n = 55$ (37%) were not.

Patients randomized to their optimal treatment improved more on the PCL-5 ($M_{\text{improvement}} = 25.65$; $SD_{\text{improvement}} = 10.04$) compared to patients randomized to their suboptimal treatment ($M_{\text{improvement}} = 20.96$; $SD_{\text{improvement}} = 9.84$; $F(1,147) = 7.67$, $p = .006$). The standardized mean difference between optimal and suboptimal treatments corresponded to a medium effect size (Cohen's $d = .47$ [.13, .81]). Figure 2 depicts the distribution of estimated change in PCL-5 and CAPS-5 scores from baseline to post-treatment for patients randomized to their optimal versus suboptimal treatment.

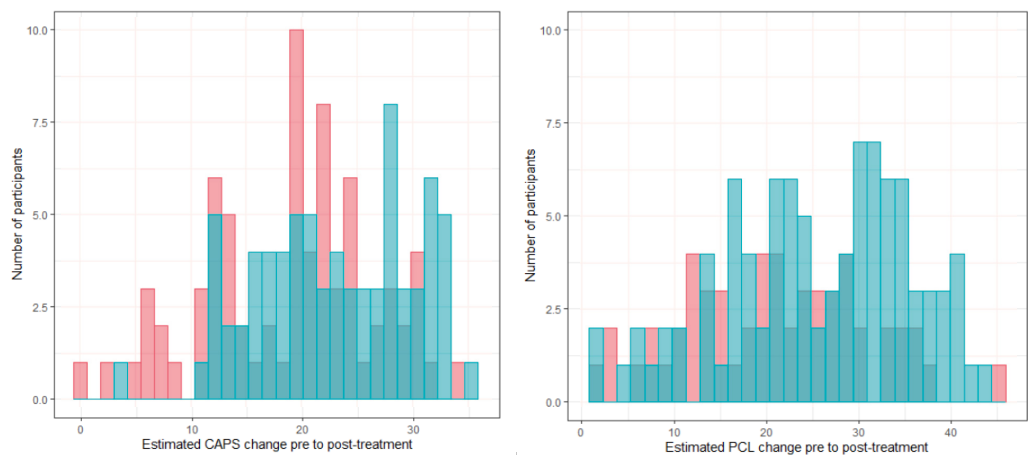


Figure 2. Estimated change in CAPS-5 (left) and PCL-5 (right) scores from baseline to post-treatment for patients randomized in their optimal (blue) and suboptimal (red) treatment condition.

Discussion

This study aimed to identify characteristics of patients with CA-PTSD which predicted treatment outcome in exposure conditions and STAIR+PE, and to evaluate the relevance of the PAI for differential treatment outcome based on the combination of these predictors. Predictors were different in the two conditions, which implies that personalized treatment recommendations have clinical potential. We found that more severe depressive symptoms and less social support were related to worse treatment outcome in the exposure conditions for both clinician-assessed and self-reported PTSD symptoms. For clinician-assessed PTSD symptoms, we also found that more axis-1 diagnoses and more severe childhood sexual abuse were related to worse treatment outcome. For the STAIR+PE condition, we found that more severe emotion regulation difficulties and lower general health status were related to worse treatment outcome for clinician-assessed and self-reported PTSD symptoms. For clinician-assessed PTSD symptoms, we also found that more severe baseline PTSD symptoms were related to worse treatment outcome. Patients randomized to their optimal treatment based on the PAI improved significantly more with medium effect sizes in clinician assessed and self-reported PTSD symptoms compared to patients randomized to their suboptimal

treatment. About half of the patients were randomized to their suboptimal treatment, implying that these patients could have benefitted from randomization based on baseline predictors.

Clinical predictors identified in the current study correspond well to the type of predictors found in previous personalization studies in patients with PTSD. Symptom burden, emotion regulation and social support are consistent indicators for personalization (Cloitre et al., 2016; Deisenhofer et al., 2018; Keefe et al., 2018). In contrast to previous studies, we did not identify demographics that predicted treatment outcome (Deisenhofer et al.; Keefe et al., 2018). This may be related to the larger number of clinical predictor candidates in our study which may be more important for treatment outcome than demographics. Predictors of the exposure conditions correspond to previously identified predictors of PTSD treatment in general. There is considerable evidence for the relationship between more severe depressive symptoms and worse treatment outcome and between less social support and worse treatment outcome of PTSD treatment (Barawi et al., 2020; Dewar et al., 2020). Predictors of STAIR+PE have not been frequently investigated, but the finding that more emotion regulation difficulties predicted worse treatment outcome in this condition seems to contradict a previous study which found that more emotion regulation difficulties relative to symptom burden was related to better outcome in STAIR+PE compared to PE (Cloitre et al., 2016). Since that study used a method which combined several moderators using a different comparator condition than our study (support+PE), results are difficult to compare. However, our finding is notable since STAIR+PE was specifically designed for patients with severe emotion regulation difficulties who might not be able to tolerate and benefit from PE (Cloitre et al., 2002). We found the opposite: more severe emotion regulation difficulties were related to *worse* outcomes in STAIR+PE specifically. Furthermore, many predictors often indicated as relevant for PTSD treatment outcome such as dissociation and personality disorders did not predict worse treatment outcomes for exposure conditions and STAIR+PE. This suggests that only a few predictors might have to be taken into account for relevant personalization recommendations.

Strengths and limitations

Strengths of the current study include the repeated measures of clinician-assessed and self-reported PTSD symptoms, the broad range of predictor candidates including patient expectancies and the robust predictor selection process and use of cross-validation techniques (James, Witten, Hastie, & Tibshirani, 2013). Notably, although we investigated a broad range of characteristics for personalization, the predictors were predominantly self-report questionnaires and similar to previous studies which focused on a limited set of predictors. Moreover, most predictors were consistent for self-reported and clinician-assessed PTSD symptoms, which implies that predictors are robust and not questionnaire specific.

An important limitation of the current study is the sample size, which did not allow for evaluation of the model using a 5-10 fold cross-validation or an holdout sample - a

statistically independent validation sample (see for example: Delgadillo & Duhne, 2020). A recent study showed that the evaluation in a holdout sample might lead to somewhat less optimistic results than the more traditional evaluation within one sample (Schwartz et al., 2020). Since there has been no external validation of personalization models in patients with PTSD yet, future personalization studies should focus on evaluating previously found models in independent samples. Additionally, the PAI was based on the linear combination of predictors in the current study. Some of the predictors identified in the Boruta algorithm but dropped during the bootstrap procedure (e.g. posttraumatic cognitions) might be relevant for treatment outcome in a non-linear manner. Future studies might evaluate how these predictors are related to treatment outcome.

Conclusions

The current study identified predictors of exposure therapies and STAIR+PE and showed that a combination of these predictors is relevant for differential treatment outcomes of patients with CA-PTSD. Future studies could evaluate previously found prediction models in independent samples and perform prospective studies in which patients are randomized based on personalized predictions or routine care (Delgadillo & Lutz, 2020). Notably, a first prospective randomized controlled trial found that treatment strategy recommendations improved treatment outcomes when therapists followed the recommended feedback (Lutz et al., 2021). If personalized predictions lead to significantly better treatment outcomes than routine care, the personalized predictions can be implemented into clinical practice using a system such as the Trier Treatment Navigator and to keep updating the predictions based on previous patients to further improve the prediction models (Lutz, Rubel, Schwartz, Schilling, & Deisenhofer, 2019). In conclusion, this study shows that tailored treatment indications based on a combination of predictors is a promising way to improve treatment outcome for patients with PTSD.

