



## ORIGINAL ARTICLE

# Theranostic markers for personalized therapy of spider phobia: Methods of a bicentric external cross-validation machine learning approach

Hanna Schwarzmeier<sup>1</sup>  | Elisabeth Johanna Leehr<sup>2</sup> | Joscha Böhnlein<sup>2</sup>  |  
 Fabian Reinhard Seeger<sup>1</sup> | Kati Roesmann<sup>3,5</sup> | Bettina Gathmann<sup>4</sup> |  
 Martin J. Herrmann<sup>1</sup> | Niklas Siminski<sup>1</sup> | Markus Junghöfer<sup>3,5</sup> | Thomas Straube<sup>4,5</sup> |  
 Dominik Grotegerd<sup>2</sup> | Udo Dannlowski<sup>2,5</sup>

<sup>1</sup>Department of Psychiatry, Psychosomatics, and Psychotherapy, Center for Mental Health, University Hospital of Würzburg, Würzburg, Germany

<sup>2</sup>Department of Psychiatry and Psychotherapy, University of Münster, Münster, Germany

<sup>3</sup>Institute for Biomagnetism and Biosignalanalysis, University of Münster, Münster, Germany

<sup>4</sup>Institute of Medical Psychology and Systems Neuroscience, University of Münster, Münster, Germany

<sup>5</sup>Otto-Creutzfeld Center for Cognitive and Behavioral Neuroscience, University of Münster, Münster, Germany

## Correspondence

Hanna Schwarzmeier, Department of Psychiatry, Psychosomatics, and Psychotherapy, Center for Mental Health, University Hospital of Würzburg, Germany.  
 Email: schwarzmei\_h@ukw.de

## Funding information

Deutsche Forschungsgemeinschaft, Grant/Award Number: Projektnummer 44541416-TRR58; Interdisciplinary Center for Clinical Research (IZKF) of the Medical Faculty of Münster, Grant/Award Number: Dan3/012/17 to UD

## Abstract

**Objectives:** Embedded in the Collaborative Research Center “Fear, Anxiety, Anxiety Disorders” (CRC-TRR58), this bicentric clinical study aims at identifying biobehavioral markers of treatment (non-)response by applying machine learning methodology with an external cross-validation protocol. We hypothesize that *a priori* prediction of treatment (non-)response is possible in a second, independent sample based on multimodal markers.

**Methods:** One-session virtual reality exposure treatment (VRET) with patients with spider phobia was conducted on two sites. Clinical, neuroimaging, and genetic data were assessed at baseline, post-treatment and after 6 months. The primary and secondary outcomes defining treatment response are as follows: 30% reduction regarding the individual score in the Spider Phobia Questionnaire and 50% reduction regarding the individual distance in the behavioral avoidance test.

**Results:**  $N = 204$  patients have been included ( $n = 100$  in Würzburg,  $n = 104$  in Münster). Sample characteristics for both sites are comparable.

**Discussion:** This study will offer cross-validated theranostic markers for predicting the individual success of exposure-based therapy. Findings will support clinical decision-making on personalized therapy, bridge the gap between basic and clinical research, and bring stratified therapy into reach. The study is registered at ClinicalTrials.gov (ID: NCT03208400).

## KEYWORDS

machine learning, spider phobia, theranostic markers

Hanna Schwarzmeier and Elisabeth Johanna Leehr contributed equally to the manuscript and both should therefore be regarded as first authors

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2019 The Authors. International Journal of Methods in Psychiatric Research Published by John Wiley & Sons Ltd

## 1 | INTRODUCTION

Specific phobias are with a 12-month prevalence of 6.4% the most prevalent anxiety disorder (Wittchen et al., 2011). Exposure-based cognitive behavioral therapy (CBT) is a first-line treatment (Bandelow et al., 2014) with medium to large effect sizes (Carpenter et al., 2018; Hofmann, Asnaani, Vonk, Sawyer, & Fang, 2012). Still, response rates indicate a clinically significant improvement only in about two-thirds of patients (Loerinc et al., 2015). Recent studies demonstrate relatively high rates of treatment dropout and relapse in anxiety disorders (Craske & Mystkowski, 2006; Fernandez, Salem, Swift, & Ramtahal, 2015; Scholten et al., 2013; Taylor, Abramowitz, & McKay, 2012). Thus, over one-third of patients may be left as nonresponders toward a first-line standard treatment with severe consequences for patients and increasing costs for societies. These figures underline the pressing need for intensified research efforts to better understand the mechanisms of exposure-based CBT and to identify markers predicting treatment (non-)response that can be applied to single patient predictions, thus allowing for *a priori* treatment decisions (Holmes, Craske, & Graybiel, 2014). Recently, machine learning approaches found their way into psychiatric research to enable the prediction of diagnosis, course, or treatment outcome on a single-case level. So far, most studies in psychiatric research used group comparison approaches that are perfect to identify, for example, a specific (neurobiological) mechanism. Still, these group-based studies are not well suited to make any prediction for the individual patient (Hahn, Nierenberg, & Whitfield-Gabrieli, 2017). Therefore, the use of machine learning can give insights into valuable predictors, which might then be investigated in future mechanistic studies in detail (Lueken & Hahn, 2016).

### 1.1 | Neuroimaging-based predictors of treatment outcome in anxiety disorders: Current state of evidence

Beyond the proof of efficacy for the average patient, current research on factors moderating the outcome of exposure-based CBT calls on strengthening the perspective on personalizing treatments (Richter, Pittig, Hollandt, & Lueken, 2017).

Several studies investigated the predictive power of clinical baseline features (Spinhoven et al., 2016) and genetic markers (Lester & Eley, 2013; Roberts et al., 2017). These studies are however limited by their group-based approach that does not allow guiding clinicians to select an adequate treatment for individual patients as a crucial prerequisite for translating personalized treatment approaches to clinical care.

Lueken et al. (2016) reviewed neurobiological markers related to treatment (pharmaco- or psychotherapy) response in anxiety disorders and identified the function of the anterior cingulate cortex (ACC) as the most promising marker, while evidence for the serotonin transporter genotype (5-HTTLPR) and cardiovascular flexibility was mixed. A recent cross-diagnostic meta-analysis identified that greater activation in a cluster located in the right cuneus extending into the right superior occipital gyrus and right middle occipital gyrus was predictive

of greater symptomatic improvement in anxiety disorders after psychological treatment (Marwood, Wise, Perkins, & Cleare, 2018).

A review of neuroimaging studies across anxiety disorders suggests that the amygdala, insula, hippocampus, and ACC constitute relevant predictors of treatment response in anxiety disorders. In addition, abnormalities in the hippocampus, amygdala, left middle temporal gyrus, fusiform gyrus, inferior occipital gyrus, left transversal temporal gyrus, inferior frontal gyrus, uncus, and areas associated with emotion regulation (dorsolateral prefrontal cortex and ACC) predict successful outcome of CBT (Santos, Carvalho, Van Ameringen, Nardi, & Freire, 2019). Of note, a number of neuroimaging studies on different forms of anxiety disorders was identified (Ball, Stein, Ramsawh, Campbell-Sills, & Paulus, 2014; Hahn et al., 2015; Månsson et al., 2015; Sundermann et al., 2017) that used machine learning methods to generate predictions on the single-case level and achieved 46–92% prediction accuracies. However, neuroimaging (mostly magnetic resonance tomography) was the data modality of choice and no study so far externally validated predictors using a second, independent dataset as adequate methodological standard as requested by current initiatives for predictive modeling studies (Collins, Reitsma, Altman, & Moons, 2015).

On a mechanistic level, corticolimbic circuitry activation in emotion regulation paradigms and hippocampus volume appeared to be predictive for CBT outcome in generalized anxiety disorder and panic disorder (Reinecke, Thilo, Filippini, Croft, & Harmer, 2014). Regarding social anxiety disorder, occipitotemporal brain activation during the presentation of angry faces versus neutral faces was positively associated with response to CBT (Doehrmann et al., 2013). Larger amygdala connectivity to a cluster encompassing the subgenual ACC, caudate and putamen, and lower amygdala connectivity with a cluster including the bilateral central sulcus and right temporal-occipital regions predicted enhanced response to CBT in patients with a social anxiety disorder (Whitfield-Gabrieli et al., 2016).

### 1.2 | The application of machine learning in psychiatry and psychotherapy

Despite initial group-based results in neuroimaging, the application of results to clinical practice is insufficient, as they do not translate into meaningful information for the individual patient (Walter et al., 2018; Woo, Chang, Lindquist, & Wager, 2017). Recently, the inference statistical approach, which is applied to find differences in efficacy between groups or mechanisms of several interventions, has been complemented by machine learning approaches. Machine learning methods enable the use of a set of multimodal predictors such as genetics, imaging, and clinical data, for (several) outcome variables (behavior and clinical characteristics) avoiding the need of multiple comparisons and making the detection of subtle variations, for example, in the brain, possible (Bzdok & Meyer-Lindenberg, 2018; Orrù, Pettersson-Yeo, Marquand, Sartori, & Mechelli, 2012; Woo et al., 2017). Multivariate pattern recognition, embedded within a machine learning framework, is a technology that has strongly influenced medical research (Darcy, Louie, & Roberts, 2016) and that bears potential

also for the field of mental health research and patient care (Orrù et al., 2012; Woo et al., 2017). By means of machine learning, an individual patient prediction of treatment (non-)response is made possible (Lueken & Hahn, 2016) and can inform about personalized treatment selection, the need of augmentation with other techniques or the treatment dose, to help in sparing ineffective treatments, associated side effects on patient compliance, disease chronification or aggravation, and direct and indirect costs. Hence, individual patient prediction is needed for a stratified treatment selection, which then needs to be tested in comparative randomized controlled trials.

However, despite their high prevalence, anxiety disorders are strongly underrepresented in predictive modeling (2.5% from all neuropsychiatric conditions), and predictive modeling accounts are still dominated by mere cross-sectional classification analyses (case/control distinction). Longitudinal data on theranostic markers (i.e., markers that predict treatment outcome; Woo et al., 2017) and cross-site validations are largely missing.

### 1.3 | Study aims and hypotheses

In the last 25 years, the field of neuroimaging research on anxiety disorders has significantly shifted from the mere characterization of pathophysiology to the description of how psychotherapy can change the brain (Messina, Sambin, Palmieri, & Viviani, 2013). Most recently, personalized medicine approaches are accumulating to detect individual pre-treatment markers of clinical response. However, the current state of evidence is limited by several gaps that are targeted by the present clinical study. Spider phobia can be used as a model disorder for pathological fear circuitry dysfunction showing high internal validity (rarely comorbidities; usually no psychopharmacological treatment) that allows for the examination of basic (neural) mechanisms. The present clinical study protocol aims to (i) identify biobehavioral markers of treatment nonresponse for a first-line treatment (e.g., behavioral exposure); (ii) overcome traditional univariate, correlational, or group-based approaches by applying state-of-the-art machine learning methodology that uses multivariate pattern recognition suitable for predictions on the individual patient; (iii) combine multiple units of analysis including (epi-)genetics, neural systems, and clinical readouts to directly compare the predictive value of these data domains, combinations thereof, and cost-efficient proxy-measures; and (iv) explicitly include an external (e.g., out-of-sample) cross-validation protocol based on our bicentric study design to evaluate the robustness and generalizability of predictors identified. Our primary hypothesis is that *a priori* prediction of treatment (non-)response in a second, independent sample based on multimodal markers is possible with sufficient prediction accuracy.

## 2 | METHODS

This clinical study is part of the Transregional Collaborative Research Center (CRC-TRR58) "Fear, Anxiety, Anxiety Disorders" funded by the German Research Foundation.

### 2.1 | Sample characteristics and recruitment pathway

We included patients with specific phobia of the animal subtype (spider phobia) assessed with the structured clinical interview for DSM-IV (SCID-IV) (Wittchen, Wunderlich, Gruschwitz, & Zaudig, 1997), who were aged between 18 and 65 years, right-handed, fluent in German language, had a Caucasian descent (back to maternal and paternal grandparents; limited to Caucasian descent due to genetic analyses/genotyping), and were willing to participate in a highly controlled behavioral exposure delivered via virtual reality (VR). Exclusion criteria comprised a lifetime diagnosis of other comorbid anxiety disorders (panic disorder, agoraphobia, social phobia, and generalized anxiety disorder), obsessive-compulsive disorder, posttraumatic stress disorder, severe major depression, borderline personality disorder, bipolar I disorder, psychotic disorders, substance dependence (except nicotine), or acute suicidality. Comorbid mild to moderate depression (unless currently treated) and other specific phobias of the animal subtype were allowed if spider phobia was the primary diagnosis. Patients with current (psycho-)pharmacological treatment, current or past psychotherapy, neurological diseases, pregnant women, and those fulfilling MRI-related exclusion criteria were excluded. Demographic and clinical characteristics of the sample at pre-treatment can be found in Table 1.

Patients were recruited via local advertisements, flyers, posters, social media, university recruitment systems, specialized outpatient centers, and medical practices. We approximated the required sample size via the binomial distribution. The primary hypothesis is to achieve an accuracy rate of our prediction of responders/nonresponders using the trained classifiers, which is significantly higher than the probability of guessing. Based on the literature (e.g., Gloster et al., 2011; Loerinc et al., 2015), a response rate of approx. 50% can be expected from patients. Regarding the response prediction, an improvement of the accuracy rate of 10%, that is, from 50% guessing probability to 60% prediction accuracy would already represent a clinically relevant number. A sample size of  $N = 75$  with  $B(75; 0.5)$  reaches significance at 45 versus 30 correctly predicted patients ( $p = 0.032$ ), which corresponds to an accuracy rate of 60%. Taking a dropout rate of 25% into consideration, we included 100 patients per site. In total, 204 patients have been recruited at both sites, 104 patients (=104% of recruitment goal) in Münster (MS) and 100 patients (=100%) in Würzburg (WÜ). The study protocol has been reviewed by the Ethics Committees of the Medical Faculties at Würzburg University (proposal number 330/15) and Münster University (proposal number 216-212-b-S). After explaining the study protocol to all participants, written informed consent was obtained. The study has been registered at ClinicalTrials.gov (ID: NCT03208400).

### 2.2 | Study design

The clinical study is a prospective longitudinal investigation employing virtual reality exposure as a first-line treatment for specific phobia. The intervention consisted of a massed one-session virtual reality

**TABLE 1** Demographic and clinical characteristics of the sample at pretreatment, means (SD), except where noted

Variables	Sample Münster (n = 87)	Sample Würzburg (n = 87)
<b>Demographic characteristics</b>		
Female gender [n (%)]	72 (82.8%)	75 (86.2%)
Age (years)	27.16 (8.33)	29.39 (9.63)
Years of education	14.76 (2.78)	14.33 (3.34)
<b>Clinical characteristics</b>		
SPQ	22.74 (2.03)	23.19 (2.35)
BAT final distance	176.22 (73.52)	171.96 (62.26)
Age of onset -SP (years)	6.13 (4.86)	8.22 (4.05)
Comorbid major depression [n (%)]	3 (3.3%)	2 (2.3%)
Comorbid subordinate animal phobia [n (%)]	2 (2.3%)	1 (1.1%)
CGI [n(%)]		
Mildly ill	9 (10.3%)	15 (17.2%)
Moderately ill	41 (47.1%)	32 (36.8%)
Markedly ill	35 (40.2%)	37 (42.5%)
Severely ill	2 (2.3%)	3 (3.4%)
FEAS anxiety	101.16 (11.98)	101.59 (14.14)
FEAS disgust	109.49 (14.86)	110.26 (11.94)
Promis Cross D	4.28 (4.12)	4.45 (4.26)
Promis Specific Phobia	11.59 (8.53)	11.10 (9.37)
STAI-Trait	35.80 (8.53)	36.29 (9.00)
BDI-II total	3.59 (4.03)	3.52 (4.24)
<b>Additional psychological characteristics</b>		
ASI-3	14.87 (9.81)	15.56 (9.95)
UI-18 reduced ability to act	11.99 (4.57)	11.57 (5.14)
UI-18 burden	13.65 (4.97)	12.94 (4.84)
UI-18 vigilance	14.31 (5.19)	15.32 (5.92)
GSE	3.01 (0.36)	2.94 (0.42)

Note. Since data collection is still in progress, only complete datasets were used for analyses here (n=87). SPQ, Spider Phobia Questionnaire; BAT, behavioral avoidance test; CGI, Clinical Global Impression; FEAS, Fragebogen zu Ekel und Angst vor Spinnen (questionnaire regarding disgust and fear of spiders); PROMIS, patient-reported outcomes measurement information system (PROMIS Cross D, across anxiety disorder diagnosis); STAI-Trait, trait-version of the State-Trait Anxiety Inventory; BDI-II, Beck Depression Inventory-II; ASI-3, Anxiety Sensitivity Index-3; UI-18, Uncertainty Intolerance 18; GSE, General Self-Efficacy Scale.

exposure treatment (VRET). Because VRET is an effective treatment option compared with a waiting list (Garcia-Palacios, Hoffman, Carlin, Furness, & Botella, 2002) and comparable with evidence-based CBT (Oprîş et al., 2012), no control condition was included. Instead, treatment responders and nonresponders as indicated by our primary endpoint measures (see below) will be compared regarding pre-treatment baseline characteristics. Furthermore, machine learning will be used to generate predictions for the individual patient. In order to carry out a sophisticated external cross-validation approach, the identical study

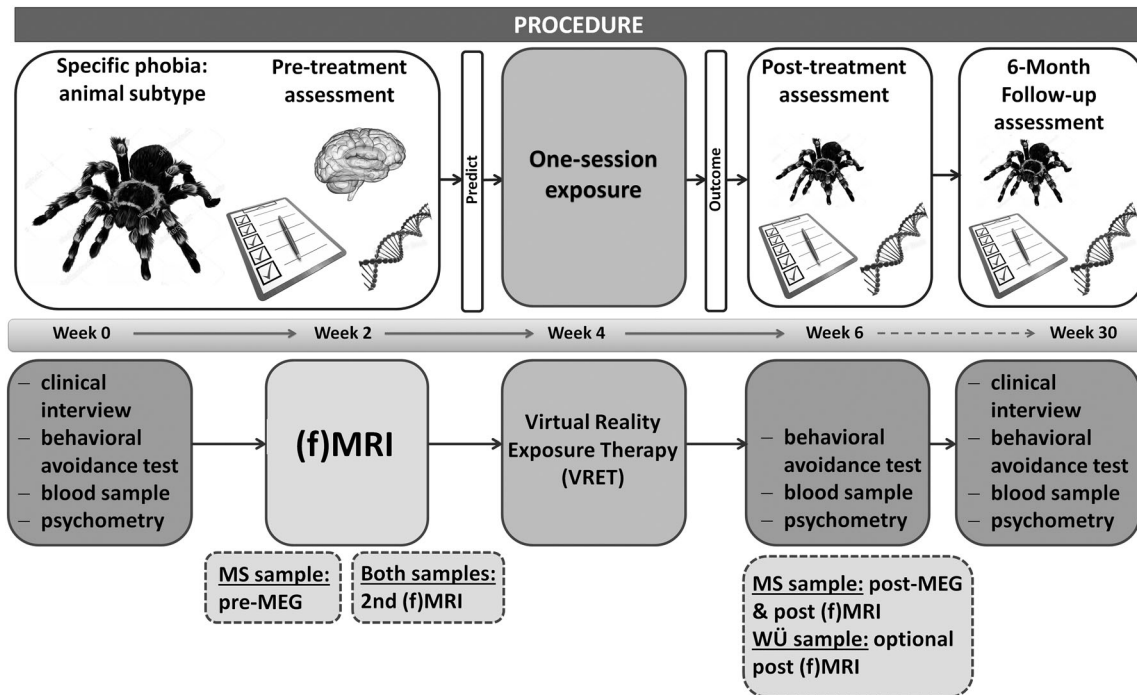
protocol is implemented in Würzburg and Münster. Patients from site A will thus serve to establish a classifier using machine learning algorithms. First, this classifier will be trained and internally validated via a leave-one-out procedure. Afterwards, the classifier will be cross-validated in the second, independent sample at site B for external validation, that is, generalizability. While at site A, the predictive pattern will be generated after completing the treatment, at site B treatment response will be predicted *a priori* in a double-blind manner (e.g., statisticians will not be informed which patients are responders or nonresponders). Contrasting the observed versus the predicted response rates will serve as a test of the hypothesis.

The study protocol (see Figure 1 for details) started with the baseline assessment collecting clinical/psychometric, behavioral, and (epi-)genetic data; neuroimaging data were gathered in a separate MRI visit. Next, the intervention (VRET) took place. After the VRET, clinical/psychometric, behavioral, and (epi-)genetic data were assessed again in a post-treatment (mean time between VRET and post-assessment over both sites was 5.32 days with a standard deviation of 6.07 days) and in a 6-month follow-up (FU) assessment. Total duration including follow-up is approx. 30 weeks, the time between the first four visits was scheduled to be 1 week each. A subsample of patients completed additional MRI visits (MS and WÜ) and/or magnetoencephalographic (MEG) assessments (MS) that are part of two other projects of the Transregional Collaborative Research Center CRC-TRR58 that share an overarching recruitment pipeline with the present study.

The MEG assessments were conducted before and after the VRET. Each assessment consisted of a fear conditioning and generalization paradigm (e.g., Onat & Büchel, 2015) with spider-related and spider-unrelated unconditioned aversive audiovisual stimuli (US). Gabor gratings with different tilt angles were either paired (CS+) or remained unpaired (CS-) with the US. In two separate blocks (spider-related vs spider-unrelated US), we obtained visually evoked magnetic fields, pupil responses, and behavioral measures in response to CS+, CS-, and so-called generalization stimuli, that is, stimuli ranging on a perceptual continuum between CS+ and CS-.

### 2.3 | VRET

In preparation for the VRET, patients were given a detailed psycho-educational manual (adapted from Herrmann et al., 2017) to read at home outlining the function of fear, its components and their interplay in general (vicious circle of fear), and more specific transferred to spider phobia and how it should be treated. Its content was discussed and explained before the VRET session started to clarify the rationale of behavioral exposure and its mechanism of action, which is suggested to induce new, inhibitory fear learning (Craske, Liao, Brown, & Vervliet, 2012). Patients also had the opportunity to ask questions. Before and after the VRET, patients completed a protocol assessing their expectations and apprehensions and which of them eventuated. The Igroup Presence Questionnaire (IPQ, Schubert, 2003) was used for measuring the sense of presence experienced in a virtual environment. The software used (VT+ research systems, VTplus



**FIGURE 1** Schematic representation of the study protocol. Pre-treatment assessment encompasses a baseline assessment to gather clinical and psychometric data, and a blood sample for genetic and epigenetic analyses is drawn. The behavioral avoidance test (BAT) serves as a quantification of avoidance behavior. A separate MRI session (structural and functional) completes the pre-treatment assessments, which will be used for prediction of treatment outcome. At the Münster site, baseline measurements are accompanied by an additional MEG measurement (prior to fMRI measurement); at both sites, there takes an additional fMRI measurement place, which is not part of the prediction track. Treatment itself consists of a one-session massed exposure therapy in virtual reality. Approx. 1 week after treatment, clinical, psychometric, behavioral, and epigenetic data are collected again, followed by an MEG- and fMRI measurement at the Münster site and optional fMRI measurement at Würzburg. Baseline measurements will be repeated at the follow-up assessment 6 months after the post-treatment assessment

GmbH, Würzburg) provides several scenarios from which five standard scenarios every patient should ideally complete were selected. The spiders used in the chosen VRET-scenarios differed. However, they were not exactly the same as the *Grammostola rosea* bird spider used in the in vivo behavioral avoidance test (BAT, please find detailed information below). While in the first scenario, the spider is quite small and black; the spiders in the other scenarios rather resemble a cross spider, although with an unrealistic body size. Due to the restriction that no other spider species are available in the used VR software, we therefore manipulated the size of the spider, the number of spiders, and the situational conditions. Before (anticipatory anxiety) and throughout the scenarios patients are constantly asked to give fear ratings on a scale from “0 = no fear at all” to “100 = extremely strong fear.” Within each scenario, we defined specific anchor points (e.g., standing right below the spider hanging from the doorframe etc.) that should be achieved by each patient. If the fear rating is <20 or if the rating is stagnating three times in succession, patients proceed with the next scenario. The number of fear ratings and the time interval between them were adapted on an individual level. The VR environment is generated using Steam Source engine (Valve Corp., Bellevue, Washington, USA) and displayed via a Z800 3D Visor (WÜ) or Oculus Rift DK2 (MS) head-mounted display (HMD; WÜ: eMagin, NY, USA; MS: Oculus VR, LLC). Maximum duration of the intervention is 2.5 hours.

## 2.4 | Measures of treatment adherence among study therapists

VRET was manualized, and study therapists were trained in cooperation and counterchecked the VRET conduction at the other site. Regular telephone conferences led by the principal investigators with longstanding clinical experience were conducted throughout the recruitment phase to discuss open questions and difficulties that appear during the conduction of the VRET.

## 2.5 | Description of VR scenarios

After an accommodation phase to the VR environment, the first scenario started with a rather small but moving spider sitting in a plastic box without a lid. Patients had to approach the box as close as possible and bend over it to watch the spider carefully. The second scenario consisted of a bigger spider hanging from the doorframe, and patients have to walk toward that door and finally stop in the doorframe beneath the spider and look up to it. In the third scenario, a big spider is crawling on the floor, and patients had to approach, to obstruct the way and to crouch down. The fourth scenario contained two spiders, one on the floor and one on the wall. Patients had to approach, to focus only on the spider on the floor and to crouch

**TABLE 2** Overview of assessments in chronological order arranged according to the type of measurement

Assessment	Baseline	MRI	VRET	Post-treatment	Follow-up
<b>Clinical</b>					
SCID	X				X
CGI	X			X	X
SPQ	X			X	X
Fragebogen Ekel und Angst vor Spinnen (FEAS) <sup>a</sup>	X			X	X
State-Trait Anxiety Inventory (STAI-Trait)	X			X	X
Beck Depression Inventory-II (BDI-II)	X			X	X
<b>Behavioral</b>					
BAT	X			X	X
<b>Neurobiological</b>					
Blood sampling	X			X	X
EDA	X			X	X
(f)MRI		X			
<b>Additional psychological characteristics</b>					
Igroup Presence Questionnaire (IPQ)			X		
Anxiety Sensitivity Index (ASI)	X			X	X
General Self-Efficacy Scale (GSE)	X			X	X
PROMIS Scales for DSM-5 (anxiety)	X			X	X
Intolerance of Uncertainty Scale(UI-18)	X			X	X
Beck Anxiety Inventory (BAI)	X				
List of threatening Experiences (LTE)	X				
Liebowitz Social Anxiety Scale (LSAS)	X				
Allgemeine Depressionsskala (ADS-K) <sup>b</sup>	X				
Agoraphobic Cognitions Questionnaire (ACQ)	X				
Penn State Worry Questionnaire (PSWQ)	X				
Social Phobia and Anxiety Inventory (SPAI)	X				
Positive and Negative Affect Schedule (PANAS-Trait)	X				
Childhood Trauma Questionnaire (CTQ)	X				
Life Calendar	X				
Kurzer Fragebogen zu Belastungen (KFB) <sup>c</sup>	X				
Brief COPE	X				
Fragebogen zur Angst vor Spinnen (FAS) <sup>d</sup>	X				
Behavioral inhibition system-behavioral activation system (BIS-BAS)				X	
Trier Inventory for Chronic Stress (TICS)				X	
Stressverarbeitungsfragebogen (SVF-78) <sup>e</sup>				X	
Cognitive Emotion Regulation Questionnaire (CERQ)				X	
Social Desirability Scale (SDS-CM)				X	
Temperamentskala (TEMPS-A) <sup>f</sup>				X	
Social Support Appraisals Scale (SS-A)				X	
Berliner Social Support Skalen (BSSS) <sup>g</sup>				X	

IPQ (Schubert, Friedmann, & Regenbrecht, 2001), ASI (Alpers & Pauli, 2001), BDI-II (Hautzinger, Keller, & Kühner, 2006), GSE (Schwarzer & Jerusalem, 1999), PROMIS (Wahl, Löwe, & Rose, 2011), STAI (Laux, 1981), UI-18 (Gerlach, Andor, & Patzelt, 2008), BAI (Margraf & Ehlers, 2007), LTE (Brugha & Cragg, 1990), LSAS (Stangier & Heidenreich, 2004), ADS (Hautzinger, Bailer, Hofmeister, & Keller, 2012), ACQ (Ehlers, Margraf, & Chambless, 2001), PSWQ (Stöber, 1998), SPAI (Fydrich, 2002), PANAS (Krohne, Egloff, Kohlmann, & Tausch, 1996), CTQ (Wingenfeld et al., 2010), Life calendar (Canli et al., 2006), KFB (Flor, 1991), COPE (Knoll, Rieckmann, & Schwarzer, 2005), FAS (Rinck et al., 2002), FEAS (Schaller, Gerdes, & Alpers, 2006), BIS-BAS (Strobel, Beauducel, & Debener, 2001), TICS (Schulz, Schlotz, & Becker, 2004), SVF-78 (Janke, 2002), CERQ (Loch, Hiller, & Witthöft, 2011), SDS-CM (Luck & Timaeus, 1969), TEMPS-A (Akiskal, Brieger, Mundt, Angst, & Marneros, 2002), SS-A (Lairreiter, 1996), BSSS (Schwarzer & Schulz, 2003).

<sup>a</sup>“Questionnaire on Disgust and Fear of Spiders”.

<sup>b</sup>German version of the Center for Epidemiological Studies Depression Scale (CES-D-scale, NIMH).

<sup>c</sup>Brief questionnaire about stresses and strains”.

<sup>d</sup>German version of Fear of Spiders Questionnaire (FSQ, Szymanski & O'Donohue, 1995).

<sup>e</sup>“Coping with Stress Inventory”.

<sup>f</sup>German version of the “Temperament Evaluation of Memphis, Pisa, Paris and San Diego Autoquestionnaire”.

<sup>g</sup>“Berlin Social Support Scales”.

down. In the fifth and last scenario, four big spiders were crawling on the floor. Patients had to approach two of them and to crouch down.

## 2.6 | Assessments: Before, after, and follow-up

An overview of all assessments can be found in Table 2.

## 2.7 | Primary outcome: Spider Phobia Questionnaire

A German translation of the Spider Phobia Questionnaire (Klorman, Weerts, Hastings, Melamed, & Lang, 1974) was used as a dimensional measure of psychopathology as this questionnaire is recommended to assess spider phobia (Hamm, 2006). The questionnaire consists of 31 items that have to be rated as “true” or “false,” maximum score per item is 1. The English version shows a satisfactory internal consistency of 0.91 (Cronbach's Alpha) and a test–retest correlation of 0.94 (Muris & Merckelbach, 1996). A sum score of at least 20 is chosen as inclusion criterion, as this is the cutoff score for clinically significant symptom severity (Öst, 1996). A reduction of at least 30% of the SPQ sum score from pre- to post-assessment will characterize clinically meaningful treatment response. SPQ scores are assessed pre-treatment, post-treatment, and after 6-month follow-up.

## 2.8 | Secondary outcome measures

An in vivo BAT (Figure 2) was used to assess generalization of treatment effects to a real spider. A bird spider (*Grammostola rosea*) was placed in a plastic box with a closed lid. The box was placed on a slide

3 m away from the patient who was then asked to slowly drag the box with the spider toward him–/herself as close as possible by using a crank. The final distance between patient and spider (in centimeter, quantification of avoidance behavior) served as the dependent variable. During the BAT, electrodermal activity (EDA) was recorded alongside with Ag/AgCl electrodes on the hypothenar of the left hand using isotonic electrode paste as contact medium and Brain Vision hard- and software for data acquisition (Brain Vision ExG Amplifier and Brain Vision Recorder; Brain Products, Munich, Germany). Patients were also asked to rate their fear on a scale from “0 = no fear at all” to “100 = extremely strong fear” for given anchor points (anticipation, at the doorstep, beginning of the BAT, after final distance, end of the BAT, and after spider left the room). In addition, observation of concomitant behavior was noted using a standardized scheme (i.e., if the patient is able to tolerate the stepwise approaching spider). All BAT outcomes are assessed pre-treatment, post-treatment, and after 6-month follow-up. A reduction of at least 50% of the final distance from pre- to post-assessment will characterize a secondary outcome of treatment response. Moreover, the Clinical Global Impressions Scale (CGI-S; Guy, 1976) used to rate symptom severity pre-treatment, post-treatment, and at follow-up as well as the SPQ score at follow-up will serve as additional secondary outcome measures.

## 2.9 | Other clinical and psychometric assessments

The SCID-IV-TR Axis I Disorders (American Psychiatric Association, 2000) was conducted to confirm the diagnosis of primary spider phobia and to check for comorbid diagnoses pre-treatment and at follow-



**FIGURE 2** In vivo behavioral avoidance test (BAT). A bird spider placed in a plastic box with a closed lid is used to assess generalization of treatment effects to a real spider. The box is placed on a slide 3 m away from the patient who then slowly drags the box with the spider toward himself as close as possible using a crank. The final distance between patient and spider (i.e., quantification of avoidance behavior) serves as the dependent variable. Patients are asked to rate their fear, observation of behavior is noted using a standardized scheme, and electrodermal activity (EDA) is recorded alongside. All outcomes are assessed pre-treatment, post-treatment, and after 6-month follow-up

up. Several questionnaires were assessed via the online survey application LimeSurvey (LimeSurvey GmbH, Hamburg, Germany) at all three assessment dates. A follow-up paper-and-pencil interview to assess avoidance behavior, encounters with spiders (actively precipitated or accidentally), experienced fear compared with pre-treatment, and an evaluation of the VRET intervention was also handed out (see supplement).

## 2.10 | (Epi-)Genetic assessment

2.7 ml EDTA-Monovette was drawn for differential blood count. Two 9 ml EDTA-Monovettes were drawn and immediately stored at  $-80^{\circ}\text{C}$  for DNA extraction. DNA will be isolated from human whole blood using the FlexiGene DNA Kit (QIAGEN, Hilden, Germany). All samples will be genotyped for several genetic polymorphisms in respective candidate genes like NPSR1 (Dannlowksi et al., 2011), CRHR1 (Weber et al., 2016), and MAOA (Reif et al., 2014) and relevant epigenetic signatures (e.g., Dannlowksi et al., 2014; Schartner et al., 2017; Ziegler et al., 2015). Results will be controlled for possible confounding factors in epigenetic analyses like smoking status (Philibert et al., 2010). Blood was taken pre-treatment for genetic and pre-treatment, post-treatment, and follow-up for epigenetic analyses.

## 2.11 | Neuroimaging battery

### 2.11.1 | (f)MRI assessments

For each patient, a structural T1 dataset to assess brain morphometry and a resting state measurement (8 min) to assess the functional organization of the brain at rest (eyes closed but awake) without any specific task-related activation was conducted with lights switched off. The first task was the Sustained and Phasic Fear Paradigm (SPF; Münsterkötter et al., 2015) to investigate the neural networks involved in the processing of a phasic fear response toward an actual threat and in the processing of anticipatory sustained anxiety toward an imminent and unpredictable hazard. It consists of 15 active and 14 inactive runs (block design). During inactive blocks, participants fixate on a dot presented in the middle of the screen for 15 s. Active blocks consist of 10 pictures each presented for 1.7 s and followed by a fixation dot (300 ms). Three fear conditions are presented in pseudorandomized order: in the sustained fear condition, where participants are told that pictures of spiders could appear, pictures of empty rooms are presented, and in one-third of the runs, a picture of a spider appears in the last quarter of the run. In the phasic fear condition, participants are instructed that they will see pictures of spiders, whereas in the no fear/safety condition, only pictures of empty rooms are presented. Each fear condition is presented five times followed by a no stimulus block, respectively. After each active run, participants evaluate their subjective appraisal regarding the pictures on a 4-point scale from *very unpleasant* to *very pleasant*. Total task duration is 9:45 min.

The second task was the Hariri face-matching paradigm that is widely used to investigate amygdala responsiveness to fearful and

angry faces (e.g., Dannlowksi et al., 2011; Hariri, Tessitore, Mattay, Fera, & Weinberger, 2002). The paradigm consists of four blocks of a face-matching task alternating with five blocks of a sensorimotor control task. During the face-matching task, patients view three triangularly arranged faces (all three expressing either anger or fear) from the Ekman and Friesen (Ekman & Friesen, 1976) stimulus set. They are instructed to select one out of two faces (bottom) that is identical to a target face (top). Each face-matching block consists of six images, balanced for gender and emotion (angry or fearful). During the sensorimotor control blocks, patients view three triangularly arranged geometric shapes (circles and ellipses) and select one of two shapes (bottom) that is identical to a target shape (top). Each sensorimotor control block consists of six different shape trios. All blocks are preceded by an instruction ("match faces" or "match shapes" in German) that lasts 2 s. In the face-processing blocks, each of the six face trios is presented for 4 s with a variable interstimulus interval of 1.5–5.5 s (mean 3.5 s), for a total block length of 47 s. In the sensorimotor control blocks, each of the six shape trios is presented for 4 s with a fixed interstimulus interval of 1.5 s, for a total block length of 35 s. Total task duration is 6:30 min. Performance (accuracy and reaction time) is recorded.

### 2.11.2 | (f)MRI data acquisition and quality control pathway

At both sites, 3-Tesla MRI scanners were used (WÜ: Siemens Skyra, MS: Siemens Prisma). Structural T1 dataset was collected using magnetization-prepared rapid gradient echo (MPRAGE): matrix =  $256 \times 256$ , slices = 176, FOV = 256, voxel size =  $1 \times 1 \times 1$  mm, TE = 2.26 ms (WÜ), TE = 2.28 ms (MS), TR = 1.9 s (WÜ), TR = 2.13 s (MS), flip angle =  $9^{\circ}$  (WÜ), flip angle =  $8^{\circ}$  (MS). Functional images were collected with a T2\* weighted echo planar imaging (EPI) sequence sensitive to blood oxygenation level-dependent (BOLD) contrast in ascending order (matrix =  $64 \times 64$ , slices = 33, FOV = 210, voxel size =  $3.3 \times 3.3 \times 3.8$  mm, slice thickness = 3.8 mm, 10% slice gap, TE = 30 ms (WÜ), TE = 29 ms (MS), TR = 2.0 s, flip angle =  $90^{\circ}$ ). Slices covered the whole brain and were positioned transaxially parallel to the anterior–posterior commissural line with a tilted angle of  $20^{\circ}$ . Stimuli were presented via MR-compatible LCD goggles (WÜ) or via a back-projection monitor (MS) and headphones using Presentation 14 - (Neurobehavioral Systems; www.neurobs.de).

Functional images are spatially and temporally aligned, normalized into standard stereotactic space and smoothed with a 8-mm full width at half maximum (FWHM) Gaussian kernel. At the first level, realignment parameters are included in the model as regressors of no interest to account for motion artifacts. The BOLD response for each condition of the SPF paradigm (phasic fear, sustained fear, no fear, and baseline) and of the Hariri face-matching paradigm (faces and shapes) is modeled by the canonical hemodynamic response function using the general linear model (GLM) to analyze brain activation differences related to the onset of the different stimuli. Parameter estimates ( $\beta$ ),  $t$ -, and  $F$ -statistic images are calculated. Significance thresholds for predefined region-of-interest (ROI) analyses and



exploratory whole-brain analyses are set to  $p < 0.05$  FWE corrected. MRI data quality control encompasses visual inspection of the structural T1 image concerning morphometry and artifacts, visual inspection of functional activation patterns for the first level contrasts, and close scrutiny of movement and rotation parameters. A global value (min–max range) for movement (x-/y-/z-axis) or rotation (pitch/roll/yaw) greater than 3.3 mm or resp.  $3.3^\circ$  leads to preclusion. Peak values (max. value from one scan to another) greater than 3.3 mm or resp.  $3.3^\circ$  are also discarded. Global and peak values are checked for each patient and paradigm, respectively.

## 2.12 | Machine learning analysis approach

Machine learning includes hypothesis-free methods to detect classification patterns of brain activity distinguishing two or more groups of subjects, particularly in high-dimensional data spaces. Support vector machine (SVM; Vapnik, 1995) and Gaussian process classifier (GPC) have been shown to be well suited for hemodynamic imaging data to predict treatment outcome following CBT or electroconvulsive therapy (Hahn et al., 2015; Redlich et al., 2016). This set of tools has been used to discriminate subforms of affective disorders, including unipolar and bipolar depression in the state of depression with up to 90% accuracy (Grotegerd et al., 2013; Grotegerd, Stuhmann, et al., 2014; Redlich et al., 2014). Analysis for the present study will be based on software tools for machine learning that have been developed by our workgroup (Grotegerd, Redlich, et al., 2014; PHOTON, 2018). Multiple information sources of different, heterogeneous data—such as hemodynamic imaging, genetic, and behavioral data—will be integrated by appropriate pattern classification methods supplemented by deep learning technology and neural networks (Figure 3). Namely, we will integrate clinical data (compare for Table 2), anatomical (T1-weighted images) and functional imaging data (fear-processing relevant contrasts) and fear-relevant genotypes [e.g., NPSR1, monoaminergic systems (MAOA, SCL6A4/5-HTT), corticotropin releasing hormone (CRHR1), brain-derived neurotrophic factor (BDNF, Val66Met), and oxytocin (OXTR)]. With regard to overfitting and shortcomings of small samples in contrast to the number of available

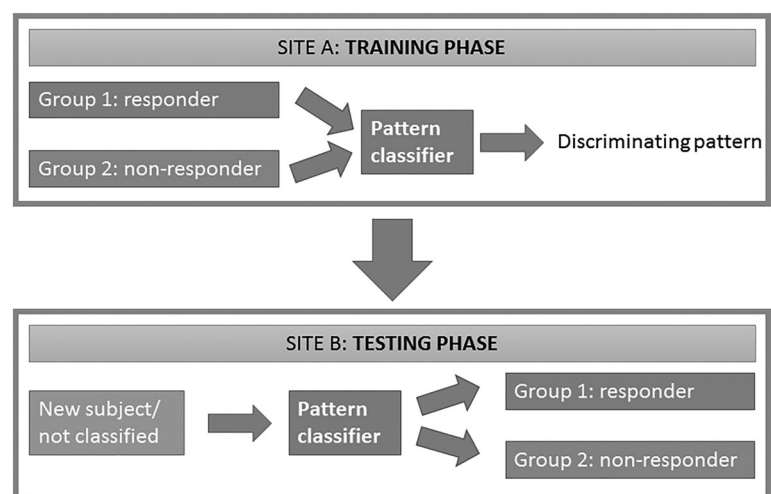
features, such data modality combination techniques provide a more reliable generalization. With kernel-based pattern algorithms (e.g., SVM and GPC), multiple kernel learning enables the combination of different data modalities, each represented by a single kernel (Sonnenburg, Rätsch, Schäfer, & Schölkopf, 2006). The actual classification algorithm learns based on a combination of the kernels using a specific function (e.g., linear combination or multiplication of kernels). Hyperparameters of these machine learning algorithms will be optimized in a nested cross-validation approach. This approach is conservative and reality-oriented and allows us to investigate the robustness regarding potential site effects. Permutation tests will be used to test accuracy significance.

## 3 | DISCUSSION

### 3.1 | Incremental value and outlook

The overall aim of the CRC-TRR58 is to concentrate expertise in the fields of molecular biology, (epi)genetics, neurophysiology, behavioral biology, neuroimaging, psychology, and psychiatry, in order to provide a better understanding of the mechanistic basis of fear, anxiety, and anxiety disorders and ultimately promote their improved treatment and prevention. Embedded in the CRC-TRR58, the presented study protocol addresses challenges in current clinical research and represents the first step into a new direction paving the way to more personalized treatments.

Focusing on the clinical translation and intervention for improved anxiety control, the benefits through the identification of biobehavioral markers of treatment nonresponders are twofold: the advancement of personalized treatments and the potential for optimization of treatment. Previous research characterizing the mechanisms of action underlying exposure and patient features associated with treatment outcome is dominated by approaches focusing on the group level. This mechanistic approach aimed at optimizing models of disease and treatments targeting at disorder-specific brain circuits (Lueken & Hahn, 2016). One major shortcoming of this group-level approach is the lack of individual or patient-specific prediction. The already



**FIGURE 3** Schematic representation of the machine learning procedure. The pattern classifier is trained at site A to generate a discriminating pattern maximally distinguishing responders from nonresponders based on data from the two groups. The classifier is then tested at site B by classifying a new subject as responder or nonresponder

mentioned rather unsatisfying response rates and effect sizes (Gloster et al., 2011; Huhn et al., 2014; Loerinc et al., 2015; Taylor et al., 2012) warrant researchers to more strongly focus on novel methods that are able to generate single patient predictions and thus may guide more personalized treatment approaches. There is evidence suggesting that data-derived subgroups within specific patient groups are better suited to predict treatment outcome than DSM or ICD diagnoses that include heterogeneous endophenotypes (Hahn et al., 2017). By further establishing these predictive approaches into practice, objectively measurable endophenotypes could be used for early disease detection, individualized treatment selection, and dosage adjustment to reduce the burden of disease. Further, the identification of biobehavioral markers can give hints at what might be the missing part in treatment nonresponse. For instance, knowing what differentiates treatment nonresponders from responders might give a starting point for treatment modification or even new development. Knowing that reactivity of the ACC or connectivity frontolimbic connectivity in response to emotional stimuli might be predictive for treatment response in anxiety disorders (Lueken et al., 2016), one might think about pre-exposition interventions addressing the ACC function, for example, interventions focusing on emotion regulation or attentional bias modification training. On the other hand, noninvasive brain stimulation methods could be used to modify these predictive areas, to augment the efficacy of exposure sessions (Herrmann et al., 2017). Beside an individualized combination of first-line treatments and the development of new interventions, addressing specific biobehavioral markers of treatment nonresponse can be launched.

The personalized treatment selection/modulation can be realized as machine learning offers the possibility of building objective algorithmic frameworks for single patient treatment response prediction across a diversity of psychiatric conditions (Bzdok & Meyer-Lindenberg, 2018; Woo et al., 2017). Because traditional disease categories are increasingly questioned to represent underlying neurobiological classes (Hyman, 2007), machine learning seems to be an appropriate tool enabling the detection of complex patterns in the brain, behavior, and genes. There are as well several important methodological benefits of multivariate predictive models (Woo et al., 2017). First of all, the direction of inference is reversed, that is, brain features, (epi-)genetic, and clinical characteristics serve as a set of predictors and treatment response as an outcome. Second, the problem of multiple testing can be overcome by the integration of existing data into one model. And lastly, the prognostic value is assessed by evaluating the performance of the model in an independent sample, thereby yielding valid estimates of effect size and ergo clinical significance (Woo et al., 2017). In addition, the usage of a bicentric cross-validation is close to reality; still, robustness to potential site-effects needs to be tested. The applied combination of multiple units of analysis, for example, (epi-)genetic, neural systems, and clinical readouts, can be adequately processed via machine learning methods. By means of several units of analysis, the optimal (cost-efficient) predictors become identifiable. The applied units of analysis in the present study could be extended in further studies using, for example, electronic momentary assessment.

In the present study, a one-session exposure-based CBT was used, and exposure is realized in virtual reality. Several studies showed efficacy of one-session exposure therapy in specific phobia (Andersson et al., 2009, 2013; Öst, 1996; Vika, Skaret, Raadal, Öst, & Kvale, 2009), superiority of five versus one session could not be proven (Ost, Brandberg, & Alm, 1997; Öst, Hellström, & Kåver, 1992; Vika et al., 2009), and on a meta-analytical base, multi-sessions were only marginally superior compared with one-session therapies (Wolitzky-Taylor, Horowitz, Powers, & Telch, 2008). Still, interpreting the results of these studies, one has to bear in mind the small number of studies. Evidence for the efficacy of VRET is comparable with exposure-based CBT in vivo (Emmelkamp, Bruynzeel, Drost, & van der Mast, 2001; Gilroy, Kirkby, Daniels, Menzies, & Montgomery, 2000, 2003), and the acceptance and commitment of patients is even higher as for exposure-based CBT in vivo (García-Palacios, Botella, Hoffman, & Fabregat, 2007). As technical aspects are improving, VRET seems to be a good alternative to exposure-based CBT in vivo (Botella, Fernández-Álvarez, Guillén, García-Palacios, & Baños, 2017), when therapist, as well as patient, keep in mind the following traps: (1) cognitive avoidance in the form of “focus on unrealness” during exposure, (2) possibly limited action radius for the patient, (3) the need to actively change procedure and circumstances, and (4) the translation into daily life of the patient. Additionally, the transdiagnostic aspects of treatment response should be targeted in future clinical research. Research focus should be on commonalities and differences in treatment response signatures across anxiety disorders or disorders from the internalizing spectrum and the question of an overarching signature of treatment non-response.

In conclusion, the present study offers the possibility to investigate theranostic markers for a model disorder of fear circuitry dysfunctions. The design was laid out to maximize the internal validity of such a novel proof of concept; if our approach turns out to be successful, future studies are needed to test these markers in more heterogeneous samples and settings in order to evaluate whether predicting treatment nonresponse in more ecologically valid settings (or, ultimately, the next patient out there) is possible with sufficient accuracy. Current challenges in psychiatric research are addressed as data from several modalities (psychometric, behavioral, neural systems, and molecular-genetic data) are integrated into a prediction model using machine learning. Cross-validation will be executed in an independent second-site sample. The detection of pre-treatment theranostic markers of clinical response could help in supporting clinical decision-making on individually tailored treatment approaches or, respectively, to spare ineffective treatment and its related financial costs. With this study, we hope to further bridge the gap between basic and clinical research and—as a long-term goal—to bring stratified therapy approaches into reach.

## ACKNOWLEDGEMENTS

This work was funded by the Deutsche Forschungsgemeinschaft (DFG)—Projektnummer 44541416-TRR 58 (CRC-TRR58, Projects C09

and Z02 to UD and UL, Project C08 to MJ and TS, Project C07 to TS and MJH) and the Interdisciplinary Center for Clinical Research (IZKF) of the medical faculty of Münster (Grant Dan3/012/17 to UD).

We would like to thank Tina Jocham, Jana Scharnagl, and Inge Gröbner (Dept. of Psychiatry, University Hospital of Würzburg); Dominik Grotegerd, Manuel Kraft, Merle Gebauer, Elena Wilkens, Jonathan Repple, Nina Muck, Stella Fingas, Janina Werner, Anna Kraus, and Kordula Vorspohl (Dept. of Psychiatry, University of Münster); Harald Kugel, Jochen Bauer, and Birgit Vahrenkamp (Dept. of Clinical Radiology, University of Münster), Lea Borgmann, Jaqueline Brieke, Aylin Fuchs, Carolin Heinemann, Annika Hense, Valeria Kleinitz, Kaja Looock, Johannes Lücke, and Kathrin Rüb (Institute of Medical Psychology and Systems Neuroscience); Tilman Coers, Julia Wandschura, Marielle Clerc, Hannah Casper, Sarah Hein, Karin Wilken, Andreas Wollbrink, Ute Trompeter, and Hildegard Deitermann (Institute for Biomagnetism and Biosignalanalysis) for their help and support.

## CONFLICT OF INTEREST

The following authors report no conflicts of interest concerning the content of this paper: H. Schwarzmeier, E. J. Lehr, J. Böhnlein, F. R. Seeger, K. Roesmann, B. Gathmann, M. J. Herrmann, N. Siminski, M. Junghöfer, T. Straube, D. Grotegerd, and U. Dannlowski.

## ORCID

Hanna Schwarzmeier  <https://orcid.org/0000-0002-7341-4744>

Joscha Böhnlein  <https://orcid.org/0000-0002-9870-5599>

## REFERENCES

- Akiskal, H. S., Brieger, P., Mundt, C., Angst, J., & Marneros, A. (2002). Temperament und affektive Störungen Die TEMPS-A-Skala als Konvergenz europäischer und US-amerikanischer Konzepte. *Der Nervenarzt*, *73*(3), 262–271. <https://doi.org/10.1007/s00115-001-1230-y>
- Alpers, G. W., & Pauli, P. (2001). *Angstsensitivitätsindex*. Würzburg: Julius-Maximilians-Universität.
- American Psychiatric Association (2000). *DSM-IV-TR: Diagnostic and statistical manual of mental disorders, text revision*. Washington, DC: American Psychiatric Association. Washington, D.C.: American Psychiatric Association
- Andersson, G., Waara, J., Jonsson, U., Malmaeus, F., Carlbring, P., & Ost, L.-G. (2009). Internet-based self-help versus one-session exposure in the treatment of spider phobia: a randomized controlled trial. *Cognitive Behaviour Therapy*, *38*(2), 114–120. <https://doi.org/10.1080/16506070902931326>
- Andersson, G., Waara, J., Jonsson, U., Malmaeus, F., Carlbring, P., & Ost, L.-G. (2013). Internet-based exposure treatment versus one-session exposure treatment of snake phobia: A randomized controlled trial. *Cognitive Behaviour Therapy*, *42*(4), 284–291. <https://doi.org/10.1080/16506073.2013.844202>
- Ball, T. M., Stein, M. B., Ramsawh, H. J., Campbell-Sills, L., & Paulus, M. P. (2014). Single-subject anxiety treatment outcome prediction using functional neuroimaging. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, *39*(5), 1254–1261. <https://doi.org/10.1038/npp.2013.328>
- Bandelow, B., Wiltink, J., Alpers, G. W., Benecke, C., Deckert, J., Eckhardt-Henn, A., ... Beutel, M. E. (2014). Deutsche S3-Leitlinie Behandlung von Angststörungen.
- Botella, C., Fernández-Álvarez, J., Guillén, V., García-Palacios, A., & Baños, R. (2017). Recent progress in virtual reality exposure therapy for phobias: A systematic review. *Current Psychiatry Reports*, *19*(7), 13–42. <https://doi.org/10.1007/s11920-017-0788-4>
- Brugha, T. S., & Cragg, D. (1990). The list of threatening experiences: The reliability and validity of a brief life events questionnaire. *Acta Psychiatrica Scandinavica*, *82*(1), 77–81. <https://doi.org/10.1111/j.1600-0447.1990.tb01360.x>
- Bzdok, D., & Meyer-Lindenberg, A. (2018). Machine learning for precision psychiatry: Opportunities and challenges. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(3), 223–230. <https://doi.org/10.1016/j.bpsc.2017.11.007>
- Canli, T., Qiu, M., Omura, K., Congdon, E., Haas, B. W., Amin, Z., ... Lesch, K. P. (2006). Neural correlates of epigenesis. *Proceedings of the National Academy of Sciences*, *103*(43), 16033–16038. <https://doi.org/10.1073/pnas.0601674103>
- Carpenter, J. K., Andrews, L. A., Witcraft, S. M., Powers, M. B., Smits, J. A. J., & Hofmann, S. G. (2018). Cognitive behavioral therapy for anxiety and related disorders: A meta-analysis of randomized placebo-controlled trials. *Depression and Anxiety*, *35*(6), 502–514. <https://doi.org/10.1002/da.22728>
- Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. M. (2015). Transparent reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD statement. *Annals of Internal Medicine*, *162*(1), 55–63. <https://doi.org/10.7326/M14-0697>
- Craske, M. G., Liao, B., Brown, L., & Vervliet, B. (2012). Role of inhibition in exposure therapy. *Journal of Experimental Psychopathology*, *3*(3), jep.026511), 322–345. <https://doi.org/10.5127/jep.026511>
- Craske, M. G., & Mystkowski, J. L. (2006). Exposure therapy and extinction: Clinical studies. In M. G. Craske, D. Hermans, & D. Vansteenwegen (Eds.), *Fear and learning: From basic processes to clinical implications* (pp. 217–233). Washington, DC: Am. Psychiatr. Assoc.
- Dannlowski, U., Kugel, H., Franke, F., Stuhrmann, A., Hohoff, C., Zwanzger, P., ... Domschke, K. (2011). Neuropeptide-S (NPS) receptor genotype modulates basolateral amygdala responsiveness to aversive stimuli. *Neuropsychopharmacology*, *36*(9), 1879–1885. <https://doi.org/10.1038/npp.2011.73>
- Dannlowski, U., Kugel, H., Redlich, R., Halik, A., Schneider, I., Opel, N., ... Hohoff, C. (2014). Serotonin transporter gene methylation is associated with hippocampal gray matter volume. *Human Brain Mapping*, *35*(11), 5356–5367. <https://doi.org/10.1002/hbm.22555>
- Darcy, A. M., Louie, A. K., & Roberts, L. W. (2016). Machine learning and the profession of medicine. *JAMA*, *315*(6), 551–552. <https://doi.org/10.1001/jama.2015.18421>
- Doehrmann, O., Ghosh, S. S., Polli, F. E., Reynolds, G. O., Horn, F., Keshavan, A., ... Gabrieli, J. D. (2013). Predicting treatment response in social anxiety disorder from functional magnetic resonance imaging. *JAMA Psychiatry*, *70*(1), 87–97. <https://doi.org/10.1001/2013.jamapsychiatry.5>
- Ehlers, A., Margraf, J., & Chambless, D. (2001). *AKV - Fragebogen zu körperbezogenen Ängsten, Kognitionen und Vermeidung*. Göttingen: Hogrefe Verlag.
- Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto: Consulting Psychologists Press.
- Emmelkamp, P. M. G., Bruynzeel, M., Drost, L., & van der Mast, C. A. (2001). Virtual reality treatment in acrophobia: A comparison with exposure in vivo. *Cyberpsychology & behavior: The impact of the Internet, multimedia, and virtual reality on behavior and society*, *4*(3), 335–339. <https://doi.org/10.1089/109493101300210222>

- Fernandez, E., Salem, D., Swift, J. K., & Ramtahal, N. (2015). Meta-analysis of dropout from cognitive behavioral therapy: Magnitude, timing, and moderators. *Journal of Consulting and Clinical Psychology, 83*(6), 1108–1122. <https://doi.org/10.1037/ccp0000044>
- Flor, H. (1991). Kurzer Fragebogen zur Erfassung von Belastungen (KFB). In Huber (Ed.), *Psychobiologie des Schmerzes*. Huber: Bern.
- Fydrich, T. (2002). *SPAI-Soziale Phobie und Angst Inventar*. Göttingen: Hogrefe.
- Garcia-Palacios, A., Botella, C., Hoffman, H., & Fabregat, S. (2007). Comparing acceptance and refusal rates of virtual reality exposure vs. in vivo exposure by patients with specific phobias. *Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society, 10*(5), 722–724. <https://doi.org/10.1089/cpb.2007.9962>
- Garcia-Palacios, A., Hoffman, H., Carlin, A., Furness, T. ., & Botella, C. (2002). Virtual reality in the treatment of spider phobia: A controlled study. *Behaviour Research and Therapy, 40*(9), 983–993. [https://doi.org/10.1016/S0005-7967\(01\)00068-7](https://doi.org/10.1016/S0005-7967(01)00068-7)
- Gerlach, A. L., Andor, T., & Patzelt, J. (2008). Die Bedeutung von Unsicherheitsintoleranz für die Generalisierte Angststörung Modellüberlegungen und Entwicklung einer deutschen Version der Unsicherheitsintoleranz-Skala. *Zeitschrift für Klinische Psychologie und Psychotherapie, 37*(3), 190–199. <https://doi.org/10.1026/1616-3443.37.3.190>
- Gilroy, L. J., Kirkby, K. C., Daniels, B. A., Menzies, R. G., & Montgomery, I. M. (2000). Controlled comparison of computer-aided vicarious exposure versus live exposure in the treatment of spider phobia. *Behavior Therapy, 31*(4), 733–744. [https://doi.org/10.1016/S0005-7894\(00\)80041-6](https://doi.org/10.1016/S0005-7894(00)80041-6)
- Gilroy, L. J., Kirkby, K. C., Daniels, B. A., Menzies, R. G., & Montgomery, I. M. (2003). Long-term follow-up of computer-aided vicarious exposure versus live graded exposure in the treatment of spider phobia. *Behavior Therapy, 34*(1), 65–76. [https://doi.org/10.1016/S0005-7894\(03\)80022-9](https://doi.org/10.1016/S0005-7894(03)80022-9)
- Gloster, A. T., Wittchen, H.-U., Einsle, F., Lang, T., Helbig-Lang, S., Fydrich, T., ... Arolt, V. (2011). Psychological treatment for panic disorder with agoraphobia: A randomized controlled trial to examine the role of therapist-guided exposure in situ in CBT. *Journal of Consulting and Clinical Psychology, 79*(3), 406–420. <https://doi.org/10.1037/a0023584>
- Grotegerd, D., Redlich, R., Almeida, J. R. C., Riemenschneider, M., Kugel, H., Arolt, V., & Dannlowski, U. (2014). MANIA—A Pattern Classification Toolbox for Neuroimaging Data. *Neuroinformatics, 12*(3), 471–486. <https://doi.org/10.1007/s12021-014-9223-8>
- Grotegerd, D., Stuhmann, A., Kugel, H., Schmidt, S., Redlich, R., Zwanzger, P., ... Dannlowski, U. (2014). Amygdala excitability to subliminally presented emotional faces distinguishes unipolar and bipolar depression—an fMRI and pattern classification study. *Human Brain Mapping, 35*(7), 2995–3007. <https://doi.org/10.1002/hbm.22380>
- Grotegerd, D., Suslow, T., Bauer, J., Ohrmann, P., Arolt, V., Stuhmann, A., ... Dannlowski, U. (2013). Discriminating unipolar and bipolar depression by means of fMRI and pattern classification: A pilot study. *European Archives of Psychiatry and Clinical Neuroscience, 263*(2), 119–131. <https://doi.org/10.1007/s00406-012-0329-4>
- Guy, W. (1976). CGI. Clinical Global Impressions. *ECDEU Assessment Manual for Psychopharmacology*. Retrieved from <http://ci.nii.ac.jp/naid/10018972210/en/>
- Hahn, T., Kircher, T., Straube, B., Wittchen, H.-U., Konrad, C., Ströhle, A., ... Lueken, U. (2015). Predicting treatment response to cognitive behavioral therapy in panic disorder with agoraphobia by integrating local neural information. *JAMA Psychiatry, 72*(1), 68–74. <https://doi.org/10.1001/jamapsychiatry.2014.1741>
- Hahn, T., Nierenberg, A. A., & Whitfield-Gabrieli, S. (2017). Predictive analytics in mental health: Applications, guidelines, challenges and perspectives. *Molecular Psychiatry, 22*(1), 37–43. <https://doi.org/10.1038/mp.2016.201>
- Hamm, A. (2006). Spezifische Phobien. Hogrefe.
- Hariri, A. R., Tessitore, A., Mattay, V. S., Fera, F., & Weinberger, D. R. (2002). The amygdala response to emotional stimuli: A comparison of faces and scenes. *NeuroImage, 17*, 317–323. <https://doi.org/10.1006/nimg.2002.1179>
- Hautzinger, M., Bailer, M., Hofmeister, D., & Keller, F. (2012). *Allgemeine Depressions Skala. Manual. 2. überarbeitete und neu normierte Auflage*. Göttingen: Hogrefe Verlag GmbH & Co KG.
- Hautzinger, M., Keller, F., & Kühner, C. (2006). *Das Beck Depressionsinventar II. Deutsche Bearbeitung und Handbuch zum BDI II*. Frankfurt a. M: Harcourt Test Services.
- Herrmann, M. J., Katzorka, A., Busch, Y., Gromer, D., Polak, T., Pauli, P., & Deckert, J. (2017). Medial prefrontal cortex stimulation accelerates therapy response of exposure therapy in acrophobia. *Brain Stimulation, 10*(2), 291–297. <https://doi.org/10.1016/j.brs.2016.11.007>
- Hofmann, S. G., Asnaani, A., Vonk, I. J. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research, 36*(5), 427–440. <https://doi.org/10.1007/s10608-012-9476-1>
- Holmes, E. A., Craske, M. G., & Graybiel, A. M. (2014). Psychological treatments: A call for mental-health science. *Nature, 511*(7509), 287–289. <https://doi.org/10.1038/511287a>
- Huhn, M., Tardy, M., Spineli, L. M., Kissling, W., Förstl, H., Pitschel-Walz, G., ... Leucht, S. (2014). Efficacy of pharmacotherapy and psychotherapy for adult psychiatric disorders: A systematic overview of meta-analyses. *JAMA Psychiatry, 71*(6), 706–715. <https://doi.org/10.1001/jamapsychiatry.2014.112>
- Hyman, S. E. (2007). Can neuroscience be integrated into the DSM-V? *Nature Reviews. Neuroscience, 8*(9), 725–732. <https://doi.org/10.1038/nrn2218>
- Janke, W. (2002). Streßverarbeitungsfragebogen:(SVF); mit SVF 120 und SVF 78. Hogrefe Verlag für Psychologie.
- Klorman, R., Weerts, T. C., Hastings, J. E., Melamed, B. G., & Lang, P. J. (1974). Psychometric description of some specific-fear questionnaires. *Behavior Therapy, 5*(3), 401–409. [https://doi.org/10.1016/S0005-7894\(74\)80008-0](https://doi.org/10.1016/S0005-7894(74)80008-0)
- Knoll, N., Rieckmann, N., & Schwarzer, R. (2005). Coping as a mediator between personality and stress outcomes: A longitudinal study with cataract surgery patients. *European Journal of Personality, 19*(3), 229–247. <https://doi.org/10.1002/per.546>
- Krohne, H. W., Egloff, B., Kohlmann, C.-W., & Tausch, A. (1996). Untersuchungen mit einer deutschen Version der “Positive and Negative Affect Schedule” (PANAS). *Diagnostica, 42*, 139–156. <https://doi.org/10.1037/t49650-000>
- Laireiter, A. (1996). *Skalen Sozialer Unterstützung [Social Support Scales Manual]*. Mödling: Schuhfried.
- Laux, L. (1981). Das State-Trait-Angstinventar (STAI): Theoretische Grundlagen und Handanweisung. Retrieved from <https://opus4.kobv.de/opus4-bamberg/frontdoor/index/index/docId/31823>
- Lester, K. J., & Eley, T. C. (2013). Therapygenetics: Using genetic markers to predict response to psychological treatment for mood and anxiety disorders. *Biology of Mood & Anxiety Disorders, 3*(1), 4. <https://doi.org/10.1186/2045-5380-3-4>
- Loch, N., Hiller, W., & Witthöft, M. (2011). Der Cognitive Emotion Regulation Questionnaire (CERQ). *Zeitschrift für Klinische Psychologie und Psychotherapie, 40*(2), 94–106. <https://doi.org/10.1026/1616-3443/a000079>
- Loerinc, A. G., Meuret, A. E., Twohig, M. P., Rosenfield, D., Bluett, E. J., & Craske, M. G. (2015). Response rates for CBT for anxiety disorders: Need for standardized criteria. *Clinical Psychology Review, 42*, 72–82. <https://doi.org/10.1016/j.cpr.2015.08.004>
- Luck, H., & Timaeus, E. (1969). Scales for the measurement of manifest anxiety (MAS) and social desirability (SDS-E and SDS-CM). *Diagnostica,*

- 15(3), 134–141. Retrieved from. <http://psycnet.apa.org/record/1970-10448-001>
- Lueken, U., & Hahn, T. (2016). Functional neuroimaging of psychotherapeutic processes in anxiety and depression. *Current Opinion in Psychiatry*, 29(1), 25–31. <https://doi.org/10.1097/YCO.0000000000000218>
- Lueken, U., Zierhut, K. C., Hahn, T., Straube, B., Kircher, T., Reif, A., ... Domschke, K. (2016). Neurobiological markers predicting treatment response in anxiety disorders: A systematic review and implications for clinical application. *Neuroscience & Biobehavioral Reviews*, 66, 143–162. <https://doi.org/10.1016/j.neubiorev.2016.04.005>
- Månsson, K. N. T., Frick, A., Boraxbekk, C.-J., Marquand, A. F., Williams, S. C. R., Carlbring, P., ... Furmark, T. (2015). Predicting long-term outcome of Internet-delivered cognitive behavior therapy for social anxiety disorder using fMRI and support vector machine learning. *Translational Psychiatry*, 5(3), e530. <https://doi.org/10.1038/tp.2015.22>
- Margraf, J., & Ehlers, A. (2007). Das Beck-Angstinventar (BAI).
- Marwood, L., Wise, T., Perkins, A. M., & Cleare, A. J. (2018). Meta-analyses of the neural mechanisms and predictors of response to psychotherapy in depression and anxiety. *Neuroscience & Biobehavioral Reviews*, 95, 61–72. <https://doi.org/10.1016/j.neubiorev.2018.09.022>
- Messina, I., Sambin, M., Palmieri, A., & Viviani, R. (2013). Neural correlates of psychotherapy in anxiety and depression: a meta-analysis. *PLoS ONE*, 8(9), e74657. <https://doi.org/10.1371/journal.pone.0074657>
- Münsterkötter, A. L., Notzon, S., Redlich, R., Grotegerd, D., Dohm, K., Arolt, V., ... Dannlowski, U. (2015). Spider or no spider? Neural correlates of sustained and phasic fear in spider phobia. *Depression and Anxiety*, 32(9), 656–663. <https://doi.org/10.1002/da.22382>
- Muris, P., & Merckelbach, H. (1996). A comparison of two spider fear questionnaires. *Journal of Behavior Therapy and Experimental Psychiatry*, 27(3), 241–244. [https://doi.org/10.1016/S0005-7916\(96\)00022-5](https://doi.org/10.1016/S0005-7916(96)00022-5)
- Onat, S., & Büchel, C. (2015). The neuronal basis of fear generalization in humans. *Nature Neuroscience*, 18(12), 1811–1818. <https://doi.org/10.1038/nn.4166>
- Opris, D., Pinteá, S., García-Palacios, A., Botella, C., Szamosközi, Ş., & David, D. (2012). Virtual reality exposure therapy in anxiety disorders: A quantitative meta-analysis. *Depression and Anxiety*, 29(2), 85–93. <https://doi.org/10.1002/da.20910>
- Orrù, G., Pettersson-Yeo, W., Marquand, A. F., Sartori, G., & Mechelli, A. (2012). Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review. *Neuroscience & Biobehavioral Reviews*, 36(4), 1140–1152. <https://doi.org/10.1016/j.neubiorev.2012.01.004>
- Öst, L.-G. (1996). One-session group treatment of spider phobia. *Behaviour Research and Therapy*, 34(9), 707–715. [https://doi.org/10.1016/0005-7967\(96\)00022-8](https://doi.org/10.1016/0005-7967(96)00022-8)
- Ost, L. G., Brandberg, M., & Alm, T. (1997). One versus five sessions of exposure in the treatment of flying phobia. *Behaviour Research and Therapy*, 35(11), 987–996. [https://doi.org/10.1016/S0005-7967\(97\)00077-6](https://doi.org/10.1016/S0005-7967(97)00077-6)
- Öst, L.-G., Hellström, K., & Käver, A. (1992). One versus five sessions of exposure in the treatment of injection phobia. *Behavior Therapy*, 23(2), 263–281. [https://doi.org/10.1016/S0005-7894\(05\)80385-5](https://doi.org/10.1016/S0005-7894(05)80385-5)
- Philibert, R. A., Beach, S. R. H., Gunter, T. D., Brody, G. H., Madan, A., & Gerrard, M. (2010). The effect of smoking on MAOA promoter methylation in DNA prepared from lymphoblasts and whole blood. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics: The Official Publication of the International Society of Psychiatric Genetics*, 153B(2), 619–628. <https://doi.org/10.1002/ajmg.b.31031>
- PHOTON. (2018). PHOTON. Retrieved from <https://www.photon-ai.com/>
- Redlich, R., Almeida, J. R. C., Grotegerd, D., Opel, N., Kugel, H., Heindel, W., ... Dannlowski, U. (2014). Brain morphometric biomarkers distinguishing unipolar and bipolar depression: A voxel-based morphometry-pattern classification approach. *JAMA Psychiatry*, 71(11), 1222–1230. <https://doi.org/10.1001/jamapsychiatry.2014.1100>
- Redlich, R., Opel, N., Grotegerd, D., Dohm, K., Zaremba, D., Bürger, C., ... Dannlowski, U. (2016). Prediction of individual response to electroconvulsive therapy via machine learning on structural magnetic resonance imaging data. *JAMA Psychiatry*, 73(6), 557–564. <https://doi.org/10.1001/jamapsychiatry.2016.0316>
- Reif, A., Richter, J., Straube, B., Höfler, M., Lueken, U., Gloster, A. T., ... Deckert, J. (2014). MAOA and mechanisms of panic disorder revisited: from bench to molecular psychotherapy. *Molecular Psychiatry*, 19(1), 122–128. <https://doi.org/10.1038/mp.2012.172>
- Reinecke, A., Thilo, K., Filippini, N., Croft, A., & Harmer, C. J. (2014). Predicting rapid response to cognitive-behavioural treatment for panic disorder: The role of hippocampus, insula, and dorsolateral prefrontal cortex. *Behaviour Research and Therapy*, 62, 120–128. <https://doi.org/10.1016/j.brat.2014.07.017>
- Richter, J., Pittig, A., Hollandt, M., & Lueken, U. (2017). Bridging the gaps between basic science and cognitive-behavioral treatments for anxiety disorders in routine care. *Zeitschrift für Psychologie*, 225(3), 252–267. <https://doi.org/10.1027/2151-2604/a000309>
- Rinck, M., Bundschuh, S., Engler, S., Müller, A., Wissmann, J., Ellwart, T., & Becker, E. (2002). Reliability and validity of German versions of three instruments measuring fear of spiders. *Diagnostica; Informationsorgan Über Psychologische Tests Und Untersuchungsmethoden*, 48, 141–149. Retrieved from. <https://www.narcis.nl/publication/RecordID/oai:repository.ubn.ru.nl:2066%2F62495>
- Roberts, S., Wong, C. C. Y., Breen, G., Coleman, J. R. I., De Jong, S., Jöhren, P., ... Eley, T. C. (2017). Genome-wide expression and response to exposure-based psychological therapy for anxiety disorders. *Translational Psychiatry*, 7(8), e1219. <https://doi.org/10.1038/tp.2017.177>
- Santos, V. A., Carvalho, D. D., Van Ameringen, M., Nardi, A. E., & Freire, R. C. (2019). Neuroimaging findings as predictors of treatment outcome of psychotherapy in anxiety disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 91, 60–71. <https://doi.org/10.1016/J.PNPBP.2018.04.001>
- Schaller, E., Gerdes, A., & Alpers, G. W. (2006). Angst ungleich Ekel: Der Fragebogen zu Ekel und Angst vor Spinnen. Wissenschaftliche Beiträge zum 24. Symposium der Fachgruppe Klinische Psychologie und Psychotherapie (Vol. 105). Lengerich.
- Schartner, C., Ziegler, C., Schiele, M. A., Kollert, L., Weber, H., Zwanzger, P., ... Domschke, K. (2017). CRHR1 promoter hypomethylation: An epigenetic readout of panic disorder? *European Neuropsychopharmacology*, 27(4), 360–371. <https://doi.org/10.1016/j.euroneuro.2017.01.005>
- Scholten, W. D., Batelaan, N. M., van Balkom, A. J., Penninx, B. W. J. H., Smit, J. H., & van Oppen, P. (2013). Recurrence of anxiety disorders and its predictors. *Journal of Affective Disorders*, 147(1–3), 180–185. <https://doi.org/10.1016/j.jad.2012.10.031>
- Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The Experience of Presence: Factor Analytic Insights. *Presence Teleoperators and Virtual Environments*, 10(3), 266–281. <https://doi.org/10.1162/105474601300343603>
- Schubert, T. W. (2003). The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realism. *Zeitschrift Für Medienpsychologie*, 15(2), 69–71. <https://doi.org/10.1026//1617-6383.15.2.69>
- Schulz, P., Schlotz, W., & Becker, P. (2004). *Trierer Inventar zum chronischen Stress: TICS*. Göttingen: Hogrefe.
- Schwarzer, R., & Jerusalem, M. (1999). *Skalen zur Erfassung von Lehrer-und Schülermerkmalen*. Freie Universität Berlin: Berlin.
- Schwarzer, R., & Schulz, U. (2003). Soziale Unterstützung bei der Krankheitsbewältigung: Die Berliner Social Support Skalen (BSSS). *Diagnostica*, 49(2), 73–82.

- Sonnenburg, S., Rätsch, G., Schäfer, C., & Schölkopf, B. (2006). Large scale multiple kernel learning. *The Journal of Machine Learning Research*, 7(12), 1531–1565.
- Spinhoven, P., Batelaan, N., Rhebergen, D., van Balkom, A., Schoevers, R., & Penninx, B. W. (2016). Prediction of 6-yr symptom course trajectories of anxiety disorders by diagnostic, clinical and psychological variables. *Journal of Anxiety Disorders*, 44, 92–101. <https://doi.org/10.1016/j.janxdis.2016.10.011>
- Stangier, U., & Heidenreich, T. (2004). In Collegium Internationale Psychiatriae Scalarum (Ed.), *Die Liebowitz Soziale Angst-Skala (LSAS)*. Weinheim: Beltz.
- Stöber, J. (1998). Reliability and validity of two widely-used worry questionnaires: Self-report and self-peer convergence. *Personality and Individual Differences*, 24(6), 887–890. [https://doi.org/10.1016/S0191-8869\(97\)00232-8](https://doi.org/10.1016/S0191-8869(97)00232-8)
- Strobel, A., Beauducel, A., & Debener, S. (2001). Eine deutschsprachige Version des BIS/BAS-Fragebogens von Carver und White. *Zeitschrift Für Differentielle Und Diagnostische Psychologie*, 22(3), 216–227. <https://doi.org/10.1024//0170-1789.22.3.216>
- Sundermann, B., Bode, J., Lueken, U., Westphal, D., Gerlach, A. L., Straube, B., ... Pfeleiderer, B. (2017). Support vector machine analysis of functional magnetic resonance imaging of interoception does not reliably predict individual outcomes of cognitive behavioral therapy in panic disorder with agoraphobia. *Frontiers in Psychiatry*, 8, 99. <https://doi.org/10.3389/fpsy.2017.00099>
- Taylor, S., Abramowitz, J. S., & McKay, D. (2012). Non-adherence and non-response in the treatment of anxiety disorders. *Journal of Anxiety Disorders*, 26(5), 583–589. <https://doi.org/10.1016/j.janxdis.2012.02.010>
- Vapnik, V. (1995). *The nature of statistical learning theory*. New York: Springer Verlag.
- Vika, M., Skaret, E., Raadal, M., Ost, L.-G., & Kvale, G. (2009). One- vs. five-session treatment of intra-oral injection phobia: A randomized clinical study. *European Journal of Oral Sciences*, 117(3), 279–285. <https://doi.org/10.1111/j.1600-0722.2009.00628.x>
- Wahl, I., Löwe, B., & Rose, M. (2011). Das Patient-Reported Outcomes Measurement Information System (PROMIS): Übersetzung der Item-Banken für Depressivität und Angst ins Deutsche. *Klinische Diagnostik Und Evaluation*, 4, 236–261.
- Walter, M., Alizadeh, S., Jamalabadi, H., Lueken, U., Dannlowski, U., Walter, H., ... Dwyer, D. B. (2018). Translational machine learning for psychiatric neuroimaging. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 91, 113–121. <https://doi.org/10.1016/j.pnpbp.2018.09.014>
- Weber, H., Richter, J., Straube, B., Lueken, U., Domschke, K., Schartner, C., ... Reif, A. (2016). Allelic variation in CRHR1 predisposes to panic disorder: Evidence for biased fear processing. *Molecular Psychiatry*, 21(6), 813–822. <https://doi.org/10.1038/mp.2015.125>
- Whitfield-Gabrieli, S., Ghosh, S. S., Nieto-Castanon, A., Saygin, Z., Doehrmann, O., Chai, X. J., ... Gabrieli, J. D. E. (2016). Brain connectomics predict response to treatment in social anxiety disorder. *Molecular Psychiatry*, 21(5), 680–685. <https://doi.org/10.1038/mp.2015.109>
- Wingenfeld, K., Spitzer, C., Mensebach, C., Grabe, H., Hill, A., Gast, U., ... Driessen, M. (2010). Die deutsche Version des Childhood Trauma Questionnaire (CTQ): Erste Befunde zu den psychometrischen Kennwerten. *PPmP - Psychotherapie · Psychosomatik · Medizinische Psychologie*, 60(11), 442–450. <https://doi.org/10.1055/s-0030-1247564>
- Wittchen, H.-U., Jacobi, F., Rehm, J., Gustavsson, A., Svensson, M., Jönsson, B., ... Steinhausen, H. C. (2011). The size and burden of mental disorders and other disorders of the brain in Europe 2010. *European Neuropsychopharmacology*, 21(9), 655–679. <https://doi.org/10.1016/j.euroneuro.2011.07.018>
- Wittchen, H.-U., Wunderlich, U., Gruschwitz, S., & Zaudig, M. (1997). SKID I. Strukturiertes Klinisches Interview für DSM-IV. Achse I: Psychische Störungen. Interviewheft und Beurteilungsheft. Eine deutschsprachige, erweiterte Bearb. d. amerikanischen Originalversion des SKID I.
- Wolitzky-Taylor, K. B., Horowitz, J. D., Powers, M. B., & Telch, M. J. (2008). Psychological approaches in the treatment of specific phobias: A meta-analysis. *Clinical Psychology Review*, 28(6), 1021–1037. <https://doi.org/10.1016/j.cpr.2008.02.007>
- Woo, C.-W., Chang, L. J., Lindquist, M. A., & Wager, T. D. (2017). Building better biomarkers: Brain models in translational neuroimaging. *Nature Neuroscience*, 20(3), 365–377. <https://doi.org/10.1038/nn.4478>
- Ziegler, C., Dannlowski, U., Bräuer, D., Stevens, S., Laeger, I., Wittmann, H., ... Domschke, K. (2015). Oxytocin receptor gene methylation: Converging multilevel evidence for a role in social anxiety. *Neuropsychopharmacology*, 40(6), 1528–1538. <https://doi.org/10.1038/npp.2015.2>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Schwarzmeier H, Leehr EJ, Böhnlein J, et al. Theranostic markers for personalized therapy of spider phobia: Methods of a bicentric external cross-validation machine learning approach. *Int J Methods Psychiatr Res*. 2020;29:e1812. <https://doi.org/10.1002/mpr.1812>