



Personalized Psychotherapy for Outpatients with Major Depression and Anxiety Disorders: Transdiagnostic Versus Diagnosis-Specific Group Cognitive Behavioural Therapy

Anita Eskildsen¹ · Nina Reinholt^{2,3} · Suzanne van Bronswijk⁴ · René B. K. Brund¹ · Anne B. Christensen^{3,5} · Morten Hvenegaard² · Mikkel Arendt^{1,6} · Anja Alrø¹ · Stig Poulsen⁷ · Nicole K. Rosenberg² · Marcus J. H. Huibers⁸ · Sidse Arnfred^{3,5}

Published online: 26 May 2020
© The Author(s) 2020

Abstract

Background Only about half of all patients with anxiety disorders or major depression respond to cognitive behaviour therapy (CBT), even though this is an evidence-based treatment. Personalized treatment offers an approach to increase the number of patients who respond to therapy. The aim of this study was to examine predictors and moderators of (differential) treatment outcomes in transdiagnostic versus diagnosis-specific group CBT.

Methods A sample of 291 patients from three different mental health clinics in Denmark was randomized to either transdiagnostic or diagnosis-specific group CBT. The study outcome was the regression slope of the individual patient's repeated scores on the WHO-5 Well-being Index. Pre-treatment variables were identified as moderators or predictors through a two-step variable selection approach.

Results While the two-step approach failed to identify any moderators, four predictors were found: level of positive affect, duration of disorder, the *detachment* personality trait, and the coping strategy of cognitive reappraisal. A prognostic index was constructed, but did not seem to be robust across treatment sites.

Conclusions Our findings give insufficient evidence to support a recommendation of either transdiagnostic or diagnosis-specific CBT for a given patient or to predict the response to the applied group therapies.

Keywords Anxiety disorders · Major depression · Precision medicine · Unified protocol · Cognitive behaviour therapy

Nina Reinholt, Suzanne van Bronswijk, René B. K. Brund have equally contributed.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10608-020-10116-1>) contains supplementary material, which is available to authorized users.

✉ Anita Eskildsen
anita.e@rm.dk

¹ Department of Affective Disorders, Aarhus University Hospital, Palle Juul-Jensens Boulevard 175, 8200 Aarhus N, Denmark

² Mental Health Services in Capital Region of Denmark, Copenhagen, Denmark

³ Department of Clinical Medicine, Faculty of Health and Medical Science, University of Copenhagen, Copenhagen, Denmark

⁴ Department of Clinical Psychological Science, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands

⁵ Psychiatry West, Region Zealand Mental Health Service, Slagelse, Denmark

⁶ Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

⁷ Department of Psychology, University of Copenhagen, Copenhagen, Denmark

⁸ Department of Clinical Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Introduction

Major depressive disorder (MDD) and anxiety disorders are the most prevalent, often co-occurring, emotional disorders found in the western world. The disorders are frequently associated with functional impairment and carry high socio-economic costs (Whiteford et al. 2013; Wittchen et al. 2011). Although cognitive behaviour therapy (CBT) and other evidence-based treatments exist for these disorders and generally show good effects, they do not work for every individual (Dunlop et al. 2017; Hofmann et al. 2012; Loerinc et al. 2015; Springer et al. 2018).

National treatment guidelines recommend CBT as one of the treatments for anxiety and for depression (NICE 2009, 2011, 2013; Danish Health Authority 2007a, b). CBT manuals traditionally address a single disorder by targeting specific psychological mechanisms (i.e., cognitive distortions) that are believed to maintain a particular disorder. Within the CBT framework, however, the frequent occurrence of co-morbidity among emotional disorders has spurred the development of transdiagnostic CBT, i.e., treatments that apply a unified set of interventions to address several anxiety disorders, depression, and other emotional disorders (Barlow et al. 2016; McEvoy et al. 2009; Reinholt and Krogh 2014; Reinholt et al. 2017). The Unified Protocol for Transdiagnostic Treatment of Emotional Disorders (UP, Barlow et al. 2011a, b; Barlow et al. 2018) targets negative affectivity by stimulating emotion regulation strategies away from avoidance towards more adaptive strategies (Barlow et al. 2014; Sauer-Zavala et al. 2012). This builds on empirical evidence which suggests that negative affectivity is an important transdiagnostic process underlying all emotional disorders (Brown et al. 1998; Clark and Watson 1991; Krueger et al. 2005). In contrast, the UP integrates standard cognitive and behavioral techniques with mindfulness-based techniques, with the main focus on learning to accept emotional experiences and manage difficult situations, despite strong emotions.

In a recent randomized controlled trial for anxiety disorders, the UP and diagnosis-specific CBT protocols showed comparable symptom reductions (Barlow et al. 2017). Other studies furthermore suggest that transdiagnostic CBT (e.g., the UP) delivered in groups for patients with anxiety disorders and/or MDD can be effective in reducing anxiety and depressive symptoms (Bullis et al. 2015; de Ornelas Maia et al. 2015; Laposa et al. 2017; Norton and Hope 2005; Norton and Barrera 2012; Osmo et al. 2015; Reinholt et al. 2017; Zemestani et al. 2017). However, no randomized controlled trials comparing diagnosis-specific CBT with the UP for groups including patients with MDD have been published (Arnfred et al. 2017; Reinholt et al. 2017).

While the reported average effects of the UP versus diagnosis-specific CBT are comparable, it is possible that some patients would have greater benefit from a broader focus on emotion regulation as applied in the UP, while others would benefit more from the specific symptom-focused approach applied in diagnosis-specific CBT. Clinical practice would therefore benefit from knowledge to assist the matching of individuals with the most optimal treatment (personalized approaches to clinical decision making).

Studying individual pre-treatment characteristics that reliably predict (differential) treatment outcomes is a possible approach to the identification of patients who would benefit most from transdiagnostic and/or diagnosis-specific CBT. Such individual characteristics may be either prescriptive or prognostic. Prescriptive variables (i.e., moderators) can help identify for whom or under what conditions a treatment has a certain causal effect on outcome. Moderators are thus useful in stratifying a population into subgroups: those who would experience greater improvement from the transdiagnostic CBT and those for whom diagnosis-specific CBT would be best (Kraemer 2013). Treatment outcomes can be predicted by prognostic variables (i.e., predictors), irrespective of treatment type. Researchers have recently developed multivariable models or algorithms to predict (differential) outcomes by integrating multiple predictors and/or moderators (Cohen and DeRubeis 2018). Introduced by DeRubeis and colleagues, the Personalized Advantage Index (PAI, DeRubeis et al. 2014a, b) is a promising approach to the prediction of differential treatment effects, which has been replicated by others (Huibers et al. 2015; Keefe et al. 2018; Vittengl et al. 2017; Webb et al. 2019; Zilcha-Mano et al. 2016). This two-step approach first selects relevant predictors and moderators of treatment outcomes, then uses the variables to construct a PAI to recommend a specific treatment for each individual. The recommendation is based on quantitative estimates of the predicted advantage of the optimal treatment over the non-optimal treatment. Further examples of multivariable prediction models to guide treatment selection are provided by the “matching factor” (Barber and Muenz 1996), the “nearest-neighbours” approach (Lutz et al. 2006), and the M* approach (Kraemer 2013; Niles et al. 2017a, b; Niles et al. 2017a, b; Smagula et al. 2016; Wallace et al. 2013). A number of multivariable prediction models containing only prognostic information have been developed over the past years. Studies using the Prognostic Index (PI) are other promising attempts in which a quantified estimate of the individual’s prognosis can help determine the needed level of care (Lorenzo-Luaces et al. 2017; van Bronswijk et al. 2019a). Such studies would be particularly relevant with regard to group CBT, since the treatment of patients with anxiety and depression is often fraught with uncertainty regarding treatment outcomes. The evidence concerning predictors of CBT effects is mainly derived from individual

treatment studies (Marker et al. 2019); whether the same predictors are relevant for group CBT is not known. However, there is evidence to suggest that:

- (a) co-morbid depression in group CBT for anxiety is a poor predictor of treatment effects (Talkovsky et al. 2016)
- (b) interpersonal problems (measured before treatment) predict lower effect of group CBT for depression (Mcevoy et al. 2013)
- (c) motivation predicts the effect of group CBT for anxiety disorders (Marker et al. 2019)
- (d) baseline severity and level of depression moderate differential effect of group CBT and mindfulness-based stress reduction for anxiety disorders (Arch and Ayers 2013).

As the evidence cited above appears from single studies with relatively modest sample sizes, our knowledge is limited as to what can help us predict who will improve during any type of group CBT and who will improve more from UP in group compared with diagnosis-specific group CBT.

The aim of the present paper was to examine moderators of treatment effects of the UP and the diagnosis-specific group CBT, using the PAI approach in a multi-site randomized controlled non-inferiority trial. The sampled patients suffered relatively severe and chronic symptoms, and the majority had failed to respond to either psychotherapy or medication. As this is typical of patients in Danish psychiatric outpatient settings, the search for moderators and predictors is even more relevant.

Should we fail to identify any moderators, we had planned to continue with an examination of general predictors to develop a PI for the estimation of personalized prediction of outcome, irrespective of the received treatment. We likewise intended to proceed to gauge the site specificity of the final model to test the generalizability of the results to other contexts. We are not aware of any other studies that have examined moderators of treatment outcome in transdiagnostic versus diagnosis-specific CBT. Furthermore, as the knowledge of predictors of group CBT treatment outcomes is very limited, we used a data-driven analysis strategy, rather than testing specific hypotheses. The study was thus exploratory in nature, aiming at using findings for the design of hypothesis-driven studies.

Methods

Design and Participants

Data came from a randomized controlled trial of the efficacy of transdiagnostic and diagnosis-specific group CBT

for MDD, social anxiety disorder, agoraphobia, and/or panic disorder (Arnfred et al. 2017). The trial is registered at Clinicaltrials.gov (NCT02954731); the primary results are planned for publication in a separate paper (Reinholt et al. under review).

Adult patients were recruited from three Danish psychiatric outpatient clinics, which form part of the publicly funded mental health services. The referral criterion for this secondary service was failure to respond to treatment (medication and/or psychotherapy) in primary care, i.e., with general practitioners or with private practice psychiatrists or psychologists. The patients may thus be classified as resistant to treatment. They were assessed using the full MINI Neuropsychiatric Interview, version 7.0.2 (except for module P; Sheehan et al. 1998). The trial inclusion criteria were (1) a principal DSM-5 diagnosis of MDD (single-episode or recurrent; approximately 50%), social anxiety disorder (approximately 25%), or agoraphobia/panic disorder (approximately 25%), (2) age 18–65 years, (3) no use of antidepressants or unchanged use of antidepressants for at least four weeks before intervention onset, with no anticipation of change in use (4) sufficient knowledge of the Danish language. Exclusion criteria were (1) moderate or high risk of suicide, (2) alcohol or substance use disorders, (3) a diagnosis of cluster A or B (DSM-5) personality disorder, (4) co-morbidity of pervasive developmental disorder, psychotic disorders, eating disorders, bipolar disorder, severe physical illness, or untreated attention deficit and hyperactivity disorder, (5) complex psychopharmacological treatment (receiving more than three types of medication for mental health disorders), (6) concurrent psychotherapy, and (7) patients declining to stop the use of anxiolytics within the first four weeks of intervention.

After written informed consent had been secured, a total of 291 patients were randomly assigned to transdiagnostic or to diagnosis-specific CBT.

Treatments

For both versions of CBT, treatment consisted of one individual session followed by 14 weekly two-hour group sessions. A group version of David Barlow's Unified Protocol for Transdiagnostic Treatment of Emotional Disorders was used as transdiagnostic CBT manual (UP, Barlow et al. 2011a, b). The manual was modified from a Danish translation of the published UP for individual therapy, with incorporation of recommendations on group delivery by the UP Institute, Center for Anxiety and Related Disorders (CARD), Boston University (Reinholt et al. 2017). Unpublished manuals based on standard group CBT manuals were used as diagnosis-specific CBT manuals (Arendt and Rosenberg 2012; Craske and Barlow 2008; Due Madsen 2008; Turk

et al. 2008), with adaptations for the Danish mental health service context.

Therapists

The 57 therapists (13 for UP and 44 for diagnosis-specific CBT) who participated in the study were licensed psychologists ($n = 38$), medical doctors in training as psychiatry specialists ($n = 7$), other professions trained in psychotherapy ($n = 7$), or psychiatrists ($n = 5$). At study onset, the therapists held an average of 8.84 years (95% CI 6.98–10.71) of clinical experience. All groups were led by at least one experienced psychologist or psychiatrist and one co-therapist who could be less acquainted with the manual. At least one of the two therapists had completed pre-trial training, followed by monthly supervision. Each therapist led between one and five groups (mean = 2.22). Regarding years of experience, no statistical significant difference was found between the UP therapists and the diagnosis-specific CBT therapists (Two-sample Wilcoxon rank-sum test ($z = 1.355$, $p = 0.1754$)).

Therapy sessions were audiotaped; 20% of the sessions were randomly selected (stratified for session number, diagnosis, and site) for assessment of therapeutic competence and protocol adherence. Competence levels were rated using the Revised Cognitive Therapy Scale (Blackburn et al. 2001, ranging from 1 to 6); while they were considered good to excellent (mean score between 4 and 6) in both conditions, mean competence scores were found to be statistically significant higher among UP therapists because of a lower competence level in diagnosis-specific CBT at one site. The potential implications of this difference are addressed in the primary RCT paper (Reinholt et al. under review). Using a scale developed for the specific manuals, adherence was measured as a percentage, with an 80% score considered "good adherence". This was achieved in both treatment arms.

Measures

Primary Outcome

As we included patients with a range of different diagnoses, we selected well-being as a user-relevant and transdiagnostic outcome measure. This was measured using the WHO Well-being index (WHO-5), which has been shown to have high clinimetric validity and to be a valid outcome measure in clinical trials. For the Danish population in general, a mean WHO-5 score of 70 has been found; 10 points is considered the threshold for a clinically relevant change. The scale may also be used as a screening tool for depression, in which case a cut-off score of ≤ 50 is used (Topp et al. 2015). The WHO-5 was completed at 16 time points: pre-treatment, before each group session, and immediately after the final treatment session (post-treatment). To aggregate the weekly

WHO-5 measures, we calculated a regression slope for each individual. This was a simple linear regression with one response variable (WHO-5) and one explanatory variable (weeks) for each patient (Triola 2010). To establish a robust estimate of the treatment effect for each patient, this model was chosen as an alternative to using end-point status or the difference between pre- and post-measures, which fails to take advantage of the multiple data points and leaves the outcome measure vulnerable to outliers. We did, however, repeat the analyses with change scores as the outcome (the difference between WHO-5 pre- and post-treatment) to allow for comparison with future studies, which may not have access to repeated measures throughout the treatment. The results of these analyses are listed in Online Appendix C.

Pre-treatment Variables

Forty-two pre-treatment variables were available as candidate predictors or moderators. To reduce multicollinearity between pre-treatment variables and thereby eliminate redundant information, we examined correlations between variables corrected for attenuation (cut-off < 0.7), as it has previously been done in similar studies (van Bronswijk et al. 2019a, b). Removing as few as possible, eight variables were removed in the moderator analyses and seven in the predictor analyses (including the pre-treatment measure of WHO-5). The clinical variables that were included in the analyses appear in Table 1. The following scales were used: The six-item Hamilton Anxiety Rating Scale (HAM-A6) (Hamilton 1959), Beck Depression Inventory-II (BDI) (Beck et al. 1996), Positive and Negative Affect Schedule (PANAS) (Watson et al. 1988), Reflective Functioning Questionnaire (RFQ) (Fonagy et al. 2016), Emotion Regulation Questionnaire (ERQ) (Gross and John 2003), Emotion Regulations Skills Questionnaire (ERSQ) (Berking et al. 2008), Perseverative Thinking Questionnaire (PTQ) (Ehring et al. 2011), Personality Inventory for DSM-5—Short Form (PID-5 SF) (Krueger et al. 2012), Level of Personality Functioning—Brief Form 2.0 (LPFS-BF) (Hutsebaut et al. 2016), Life Event Checklist for DSM-5 (LEC-5) (Gray et al. 2004) and Standardized Assessment of Personality—Abbreviated Scale (SAPAS) (Moran et al. 2003). More details on the scales were reported in Arnfred et al. 2017; they may also be found in Online Appendix A (supplemental material). Information on demographics and backgrounds is listed in Online Appendix B.

Statistical Analyses

Missing Data and Description of Variables

Missing variables were imputed, using a non-parametric random forest approach (R package "MissForest"; Stekhoven

Table 1 Clinical characteristics of groups

	UP (N = 148) N (%)	Diagnosis-specific CBT (N = 143) N (%)
Disorder and co-morbidity		
Primary diagnosis		
MDD	73 (49.32)	73 (51.05)
Social anxiety disorder	40 (27.03)	40 (27.97)
Agoraphobia/panic disorder	35 (23.65)	30 (20.98)
Both MDD and anxiety disorder	60 (40.54)	52 (36.36)
More than one co-morbid disorder ^a	38 (25.68)	30 (21.13)
Previous episodes of primary disorder		
0 previous episodes	65 (44.22)	66 (46.48)
1 previous episode	31 (21.09)	33 (23.24)
2–3 previous episodes	25 (17.01)	26 (18.31)
> 3 previous episodes	26 (17.69)	17 (11.97)
Previous admission to psychiatric hospital	22 (15.49)	23 (16.55)
Previous psychological treatment	116 (82.27)	122 (89.05)
Concentration difficulties ^b	57 (40.14)	49 (36.57)
Current psychoactive drug (yes/no)	107 (72.30)	95 (66.90)
	Mean (CI)	Mean (CI)
Duration of psychoactive drug treatment	82.98 (58.37:107.58)	85.08 (51.32:118.84)
Duration of disorder (in months)	57.94 (41.71:74.17)	51.83 (37.92:65.75)
Age of onset	21.85 (20.08:23.63)	23.72 (21.58:25.85)
Well-being (WHO-5)	31.70 (28.64:34.76)	31.57 (28.15:34.98)
Depression severity (BDI-II)	27.12 (25.27: 28.97)	27.07 (25.01:29.12)
Anxiety severity (HAM-A6)	9.11 (8.39:9.83)	8.01 (7.29:8.74)
Personality trait		
Negative affect (PID-5)	1.51 (1.42:1.61)	1.41 (1.31:1.52)
Detachment (PID-5)	1.21 (1.11:1.31)	1.17 (1.06:1.27)
Antagonism (PID-5)	0.33 (0.27:0.39)	0.34 (0.28:0.41)
Disinhibition (PID-5)	0.98 (0.89:1.07)	0.87 (0.79:0.96)
Psychoticism (PID-5)	0.47 (0.41:0.53)	0.41 (0.34:0.47)
Level of Personality Functioning (LPFS-BF)	1.42 (1.33:1.50)	1.43 (1.34:1.53)
Cognitive Reappraisal (ERQ)	3.68 (3.47:3.88)	3.63 (3.37:3.88)
Expressive Suppression (ERQ)	3.94 (3.71:4.18)	4.03 (3.79:4.27)
Positive Affect Schedule (PANAS)	21.12 (20.09:22.14)	20.58 (19.49:21.67)
Emotion regulation (ERSQ)	1.75 (1.64:1.85)	1.78 (1.68:1.89)
Perseverative Thinking (PTQ)	40.53 (38.81:42.24)	40.77 (38.96:42.58)
Certainty about mental states (RFQ)	20.91 (18.91:22.91)	22.13 (19.85:24.41)
	N (%)	N (%)
SAPAS (> 2 yes-answers)	52 (35.86)	47 (33.57)
Treatment course		
Median number of sessions completed (25 and 75 percentile)	10 (3:13)	10 (5:13)
WHO-5 slope (Mean and CI)	0.59 (0.42:0.76)	0.64 (0.46:0.82)

UP unified protocol, BDI-II beck depression inventory, ERSQ emotion regulation skills questionnaire, ERQ emotion regulation questionnaire, HAM-A6 hamilton anxiety rating scale, LPFS-BF level of personality functioning scale-brief form, PANAS positive and negative affect schedule, PID-5 personality inventory for DSM-V-short form, PTQ perseverative thinking questionnaire, RFQ reflective functioning questionnaire, SAPAS standardised assessment of personality—abbreviated scale.

^aAnxiety disorders, affective disorders or eating disorders

^bSelf-reported moderate to severe difficulties with concentrating

and Bühlmann 2012) using all available data as input. This imputation approach has been shown to be accurate and comparable to multiple imputation and has been used in similar studies before (van Bronswijk et al. 2019a, b; Waljee et al. 2013). The imputation method was tested by applying the same method to the portion of the dataset with no missing values, using artificially produced missing data, subsequently comparing the imputed values with the actual data values by estimating the normalized root mean squared error (NRMSE) for continuous data and the proportion of falsely classified entries (PFC) for categorical data, as suggested by Stekhoven and Buhlmann (2012). Since the missing variables may not have been missing at random, we wanted to test whether imputing the outcome variable could lead to biased results (Sullivan et al. 2017). We therefore repeated the analyses on the dataset in which the outcome variable was not imputed as a sensitivity analysis (see Online Appendix C).

In the PAI analysis, we were interested in the differential treatment effects of the two treatments. To this end, we built a model based on individuals that had received a meaningful course of CBT. Previous research has indicated that four treatment contacts should be considered the lowest meaningful amount of psychotherapy (Delgadillo et al. 2014; Robinson et al. 2020). This four-sessions-or-more rule has been applied in a previous PAI study (Cohen et al. 2020) and was also used in this study. With regard to the PI analyses, we did not use the same limited sample, as we were more interested in general predictors irrespective of amount of treatment.

Pre-treatment Variable Transformation

Discrete variables were centred and continuous variables with a near-to-normal distribution were standardized. Continuous variables with a non-normal distribution were transformed using log transformation or a square root transformation based on visual inspection of histograms.

Pre-treatment Variable Selection

Different statistical methods have been applied to select pre-treatment variables for personalized prediction models (Cohen and DeRubeis 2018). In 2018, Keefe et al. introduced a two-step machine learning approach using a random forest method for model-based recursive partitioning followed by a stepwise AIC-penalized bootstrapped method. Since the development of prediction models in mental health is still in its infancy, there is little agreement on the most appropriate machine learning techniques. However, based on the following strengths, we chose to use the two-step method introduced by Keefe et al. (2018):

1. *Random forest* Previous research gives reason to believe that treatment response is a multifactorial process involving multiple variables, each having a small effect (Cohen and DeRubeis 2018). Because of this, it is important to apply a method that, while handling a large number of pre-treatment variables, is capable of preventing slightly weaker pre-treatment variables from being dominated by stronger pre-treatment variables. The random forest methodology for model-based recursive partitioning is well suited for this purpose.
2. *Stepwise AIC-penalized bootstrapped method* A second variable selection approach was added to ensure that the pre-treatment variables were predictive in multiple bootstrapped replications of the data. This approach also enabled us to determine the relationship between the pre-treatment variable and the outcome, based on the direction of the estimated regression coefficients.

In the first step, we applied a model-based recursive partitioning method using a random forest algorithm (R package "mobForest"; Garge et al. 2013, 2018). With this algorithm, multiple trees are created by repeatedly splitting bootstrapped samples into two subgroups. Splits are based on pre-treatment variables that lead to significantly different model behaviour on either side of the split. A regression model with the slope of WHO-5 as the dependent variable and treatment as the independent variable ($y = \text{treatment}$) acted as the pre-determined model for the current analyses. A "split variable" therefore indicates an association with treatment differences (i.e., a potential moderator). A random subset of 10 variables was available to inform each split producing 10,000 different tree-like structures by repeatedly splitting the sample on the variable with the strongest moderator impact until a minimum subgroup of 10 individuals was achieved. By using different random subsets of variables, moderators with smaller effects were less likely to be dominated by stronger moderators (Strobl et al. 2008). An α -level of 0.10 was set for splitting. This method provides a variable importance score for each pre-treatment variable, thus indicating the predictive value. Variables were carried forward to the next step if their permutation accuracy importance score exceeded the threshold, which was the absolute value of the variable importance score of the lowest ranking variable (Garge et al. 2013). As the study was a multi-centre study, the variable site was included in the analyses, together with the pre-treatment variables.

As a second step, we tested whether the variables identified in the first step would be selected as significant moderators in at least 60% of the 1000 bootstrapped samples using a multiple linear regression model with backward elimination (with an α -level of 0.05, R package "bootstepAIC"; Austin and Tu 2004; Rizopoulos and Rizopoulos 2009). The regression model was specified with the slope of the WHO-5 as the

dependent variable and the selected first-step pre-treatment variables (in the model-based recursive partitioning method) as the independent variables along with their interactions with treatment. The outcomes of this analysis were the number of times each variable was selected as a statistically significant moderator in the bootstrapped samples and whether the regression coefficient was positive or negative. The moderator was considered robust if selected in at least 60% of the samples. The 60% cut-off has previously been shown to construct a parsimonious model with good model fitting (Austin and Tu 2004), and similar studies has used it as a cut-off (Cohen et al. 2020; Keefe et al. 2018; van Bronswijk et al. 2019a, b; Zilcha-Mano et al. 2016).

Our results may have been substantially affected by the parameter selection, e.g., the number of randomly preselected predictor variables and the minimum required number of individuals at each node to form the tree-like structure (Strobl et al. 2009). We therefore planned to assess the potential influence of different parameter selections on the variables selected by the model-based recursive partitioning method (as reported in Online Appendix C).

Building and Evaluating the Personalized Advantage Index

If the two-step variable selection approach identified robust moderators, we planned to create a personalized advantage index (PAI). The PAI combines the identified pre-treatment variables in a multiple linear regression model and the slope of the WHO-5 as the dependent variable. The independent variables are the pre-treatment variables identified as predictors and the variables identified as moderators and their interactions with treatment. For instance, if two general predictors and two moderators were selected in the two-step approach described above, the equation for the PAI would be: $\text{Slope}_{\text{WHO-5}} = \beta_0 + \beta_1 * \text{predictor}_1 + \beta_2 * \text{predictor}_2 + \beta_3 * \text{moderator}_1 + \beta_4 * \text{moderator}_2 + \beta_5 * \text{moderator}_1 * \text{treatment} + \beta_6 * \text{moderator}_2 * \text{treatment}$. Individual outcome predictions for each treatment are then calculated based on this regression model using a fivefold cross-validation. In such a validation, the sample is split into five equal groups and individual outcomes are predicted using the regression model with weights based on the data of the other four groups to which the individual patient does not belong (Picard and Cook 1984). Two separate predictions are made for each individual, one for each treatment (UP and diagnosis-specific). The difference (positive or negative) in predicted outcomes (coefficients of the WHO-5 slope) constitutes the PAI, indicating whether UP or diagnosis-specific CBT should be recommended for the treatment of the given patient. A quantitative estimate of the advantage of the recommended treatment is thus provided.

To evaluate the PAI, we planned to compare the actual slopes for patients who had received the PAI-indicated

treatment with the slopes for patients who had received their non-indicated treatment. This would be done by t-testing (DeRubeis et al. 2014a, b).

Building and Evaluating the Prognostic Index (PI)

If no robust moderators were identified in the two-step variable selection approach, we planned to build a PI to predict overall treatment effects for each individual. To identify robust predictors, we planned to use a variable selection procedure similar to the one described above, but with three adjustments: (1) in the first step we would adjust the pre-determined model to identify general predictors of outcome rather than moderators of differential outcome ($y = \text{intercept} + \text{variables}$); (2) we would not include an interaction with treatment in the regression model of the second step; and (3) we would use the whole sample ($n = 291$) rather than only those patients who completed ≥ 4 sessions, since differential treatment effects were irrelevant. After the two-step variable selection, we planned to use the identified predictors to build a multiple linear regression model with the slope of the WHO-5 as the dependent variable and the identified predictors as the independent variables ($\text{Slope}_{\text{WHO-5}} = \beta + \beta_1 * \text{predictor}_1 + \beta_2 * \text{predictor}_2 \dots + \beta_n * \text{predictor}_n$). Using a fivefold cross-validation, individual outcome predictions could then be calculated on the basis of this regression model (Picard and Cook 1984).

To evaluate the PI, we planned to compare the predicted slope for each individual (based on the weights of the other four folds, to which the individual did not belong) with the actual "observed" slope for that individual. The difference between the two WHO-5 slopes could then be calculated, and the association between these scores examined using Pearson's correlation analysis.

Testing the Generalizability of Variable Selection and Model Fitting

Since we used the full sample for the two-step variable selection approach and model fitting, the PAI/PI may have been inflated due to double-dipping (Fiedler 2011; Vul et al. 2009). Therefore, as a sensitivity analysis, we planned to revisit all the steps using data combining Sites 1 and 3 ($n = 168$) and comparing the results with those of Site 2 ($n = 123$; see Online Appendix C).

Results

Variable Description and Missing Data

The PI analyses included 291 randomized patients (UP: $n = 148$; diagnosis-specific CBT: $n = 143$). The means and

frequencies of the clinical characteristics of the patients are listed in Table 1. Online Appendix B lists demographic and background information. The median number of completed sessions in both groups was 10. Four or more group sessions were completed by 228 patients (UP: n = 110; diagnosis-specific CBT: n = 118), who were included in the PAI analyses.

There were no significant differences between the two treatment groups regarding demographic or clinical characteristics except that the UP group had higher levels of anxiety symptoms on HAM-A.

A total of 14.5% of the observations were missing. We tested the imputation method by artificially producing the missing data on the data of the 81 patients with no missing pre-treatment data or missing outcome data before Session 5. With a 0.33 error term for the continuous variables (NRMSE), and a 0.41 error term for the discrete variables (PFC), the imputation method appeared satisfactory.

Variable Selection for Personalized Advantage Index

Age, the Positive Affect Schedule, level of the *detachment* personality trait, duration of disorder, and BDI-II were identified by running the model-based recursive partitioning method, with age having the largest variable importance score. None of these variables, however, were selected as moderators in at least 60% of the bootstrap samples using the backwards elimination technique, indicating that no robust moderators were available for building a PAI.

Variable Selection for PI

Thirteen predictors of outcome were identified by running the model-based recursive partitioning method in the search of predictors. They are depicted in Table 2.

Only level of positive affect (measured by the PANAS), duration of disorder, level of *detachment* personality trait, and cognitive reappraisal were selected in at least 60% of the bootstrap samples (Table 2, in bold). More than 99% of the bootstrap samples showed negative coefficients, indicating that longer duration of disorder, higher levels of positive affect, detachment, and cognitive reappraisal were associated with less improvement during treatment.

Estimating PI Scores Using Fivefold Cross Validation

The four selected predictors were combined into the following multiple linear regression model: $Slope_{WHO-5} = \beta + \beta_1 * \log(\text{positive affect}) + \beta_2 * \log(\text{length of disorder}) + \beta_3 * (\text{detachment}) + \beta_4 * (\text{cognitive reappraisal})$. Comparing the actual "observed" slope with the predicted slope (based on weights in four of five folds), we found a mean difference of 0.01, with a 95% CI between - 0.10 and 0.13. The correlation (Pearson's *r*) was 0.25 ($p < 0.001$).

Generalizability of Variable Selection and Model Fitting

When we separately conducted the two variable selection steps at each treatment site (Site 2 versus Sites 1 and 3), we

Table 2 Predictors selected with model-based recursive partitioning technique (N = 291)

Variable importance rank (High to low)	Variable	Percentage (%) of bootstrapped samples that selected variable	Significant β	
			Positive	Negative
1	Depression severity (BDI-II)	37.8	9.52	90.48
2	Detachment (PID-5)	78.1	0.13	99.87
3	Level of personality functioning (LPFS-BF)	28.0	56.43	43.57
4	Positive affect (PANAS)	93.0	0.00	100.00
5	Duration of disorder	87.1	0.00	100.00
6	Perseverative thinking (PTQ)	43.9	99.77	0.23
7	Cognitive reappraisal (ERQ)	62.0	0.48	99.52
8	Expressive Suppression (ERQ)	36.8	0.82	99.18
9	Emotion regulation (ERSQ)	23.1	55.41	44.59
10	Psychoticism (PID-5)	40.1	95.26	4.74
11	Certainty about mental states (RFQ)	49.6	99.60	0.40
12	Anxiety severity (HAM-A6)	26.6	8.27	91.73
13	Disinhibition (PID-5)	32.0	3.44	96.56

Bold values indicate the variables selected as predictors in at least 60% of the bootstrap samples

BDI-II beck depression inventory, *HAM-A6* hamilton anxiety rating scale, *ERSQ* emotion regulation skills questionnaire, *ERQ* emotion regulation questionnaire, *LPFS-BF* level of personality functioning scale-brief form, *PANAS* positive and negative affect schedule, *PID-5* personality inventory for DSM-V-short form, *PTQ* perseverative thinking questionnaire, *RFQ* reflective functioning questionnaire

were unable to identify any predictors, which were selected both in the primary PI and in the PIs on the two separate samples (see Online Appendix C).

Sensitivity Analyses

No considerable change in results was observed when the analyses were repeated without imputing the outcome variable except that only two variables were selected in both steps: positive affect and duration of disorder. Changing the parameters in the model-based recursive partitioning method also failed to modify the results (details in Online Appendix C).

Discussion

We were unable to identify any robust moderators of differential treatment outcome in this study. This suggests that the pre-treatment patient variables trialled here are of no use in predicting whether patients with social anxiety disorder, agoraphobia/panic disorder, and/or MDD will improve more or less from the UP compared with diagnosis-specific group CBT. We identified four predictors of outcome: level of positive affect, duration of disorder, the *detachment* personality trait, and the *cognitive reappraisal* coping strategy. The predictors were negatively associated with outcome, indicating that higher pre-treatment levels of positive affect, i.e., feeling enthusiastic, active, and alert, predicted less improvement during treatment. This may be surprising but could result from the fact that positive affect and well-being are highly correlated and that a high level of well-being at baseline allows less room for improvement and thus a lower slope. This finding corresponds with those of a study of social phobia in which high positive affect before treatment was found to predict less improvement in quality of life during treatment (Sewart et al. 2019). Cognitive reappraisal as an emotion regulation strategy is thought to intervene early in the emotion-generative process and typically leads to more positive emotions and less negative emotions (Gross and John 2003). Our finding that this coping strategy predicts less improvement may therefore be surprising, whereas it may be less surprising that a longer duration of disorder and the detachment personality trait (the tendency to avoid socioemotional experience) predicted less improvement. According to the sensitivity analysis concerning site specificity, however, the PI model was not reliable across treatment sites. This indicates that further study is needed to improve the PI. Variables such as motivation and expectancy as well as interpersonal problems could be relevant for inclusion in future studies. Measures of stability/strain in everyday life should likewise be considered, since these aspects of the patient's life may interfere with home work engagement

that has been shown to mediate the effect of CBT in some studies (Cammin-Nowak et al. 2013; Westra et al. 2007).

Our results provide no evidence to support a preference for either UP or diagnosis-specific group CBT for a given patient. Accordingly, treatment selection may be based on patient preferences or logistics as well as on the variables examined. However, other variables, not addressed in the present study (e.g., interpersonal problems or prior experience with CBT), should be investigated in future studies before any conclusions can be drawn as to the existence of moderators. In some treatment settings, a logistic advantage may be obtained by using transdiagnostic group CBT rather than diagnosis-specific group CBT since patients would not have to wait until a sufficient number of patients with the same diagnosis were ready to enter their designated group. Moreover, shifting from diagnosis-specific to transdiagnostic CBT may potentially reduce the costs of training therapists, as they would need training only in one manual.

Several aspects may be considered when attempting to explain the lack of identified moderators in this study. Despite differences in the treatment targets in the UP and diagnosis-specific CBT and the heterogeneity of the UP group (only this group had patients with different diagnoses), a likely explanation is that the two therapy formats were too similar in nature and used similar or overlapping interventions. For instance, although the therapies were delivered with different rationale and goals, exposure was an important intervention in both arms of the study. This may have rendered the search for moderators less relevant here than in previous PAI studies, where psychotherapy was compared with medications (e.g., DeRubeis et al. 2014a, b; Vittengl et al. 2017) or CBT was compared with either interpersonal therapy (Huibers et al. 2015; van Bronswijk et al. 2019a, b) or psychodynamic therapy (Cohen et al. 2020).

Another important question concerns the quality of treatment. Both treatment formats led to limited improvement as a WHO-5 slope of some 0.6 (95% CI 0.43–0.77) corresponds to an 8.4-point change (6.02–0.78) over the 14 weeks. The mean change obtained in our study thus failed to reach the clinically significant level, which was defined as a 10-point improvement on the WHO-5 scale (Arnfred et al. 2017). However, as the current study was set in secondary services, the included patients had previously failed to respond to medication and/or psychotherapy offered by primary services (typically delivered by a general practitioner or a private practice psychiatrist or psychologist). The patients may therefore be classified as treatment resistant, with a very low level of functioning in daily life (e.g., less than 10% worked full-time at time of study enrolment; 22% were classified as students). In addition, the use of the well-being index as an outcome measure may have diminished the obtained treatment effect, since positive variables tend to be less responsive as treatment targets (Sewart et al. 2019).

However, changes in well-being and symptoms were similar in magnitude, which suggests that this explanation is less plausible here (Reinholt et al., under review). Taking into account the severity and chronicity of the sample as well as the relatively short duration of the treatment, the positive nature of the outcome measure, and the fact that the rating of treatment quality and adherence was high, we find it unlikely that the modest treatment effect reflects poor treatment quality. It cannot be ruled out, however, that moderators may have been identified if we had used a less chronic sample, longer duration of treatment, or another outcome measure, thereby increasing the potential of change during treatment. The homogeneity of the study population is another possible explanation for the failure to detect differences in outcome, since the participants were all non-responders to first-line treatment. Likewise, a high proportion of the group may have been what DeRubeis and colleagues have termed as *intractable*, i.e., patients who would experience no improvement no matter the quality of therapy provided (DeRubeis et al. 2014a, b). However, the highly varied response renders this explanation less likely. The slope range of -3.56 to 6.24 corresponds to a change in WHO-5 range from -49.8 to 87.36 points.

It is also possible that the samples from the three treatment sites differed since it was not possible to identify the same predictors in the three sites when conducting the analyses separately. However, treatment site was not selected as an important variable in the first variable selection step, at which the samples were pooled.

It may be speculated that we overestimated the capacity of multivariable prediction models at this stage of model development. Perhaps the difficulty of building prediction algorithms is greater than expected. Other factors may have played a role, such as the numerous interactions between variables, which may have rendered the models too complex, combined with the possibility that pre-treatment variables were less important than processes during or outside treatment (i.e., non-specific treatment factors such as alliance and group cohesion or unexpected events in the family or at work, etc., Kazantzis et al. 2018; Zilcha-Mano 2017).

Proponents of the complex network approach to psychopathology argue that symptoms are not a reflection of an underlying disease or dimension (e.g., neuroticism), but rather that they constitute the disease itself in the form of a complex network of elements interacting in ways that tend to maintain themselves (Hofmann et al. 2016). In this perspective, response to treatment is caused by critical transitions in network states; hence, pre-treatment factors do not necessarily offer important information on the likelihood of change or the resilience of a given network, which may explain the scarcity of predictors of treatment effect found in the current study. Furthermore, it is possible that the similarities between the two treatment conditions may have provided a similar push

to the complex networks, which may explain the failure to find any moderator of treatment effect. In a complex network perspective, the current understanding of moderation as the effect of baseline characteristics may be too narrow, inasmuch as a reassessment of the same patient may be needed to capture the dynamics of a complex ideographic network (Hofmann and Hayes 2019).

Variable selection methods provide a further point of discussion. Such approaches may be biased towards multi-category variables, which are more likely to be selected in the model (Kim and Loh 2001; Strobl et al. 2007). To reduce bias in the current study, predictors were standardized before running the model-based recursive partitioning model. We cannot rule out, however, a bias in the variable selection and the estimation of variable importance scores. The effect may be that less optimal predictors are carried forward to the next step. Additionally, the sample size of our study may have been too small. Power calculation is no simple task in analyses such as these, but a newly published simulation study by Luedtke and colleagues conclude that a sample size of 300 per treatment arm is required for sufficient statistical power to use multivariable prediction models with four predictors for comparison of two or more treatments (Luedtke et al. 2019). On the other hand, the current study has a larger sample size than most of the previous PAI studies.

Some strengths of the study should also be noted. To the best of our knowledge, this is the first study to examine moderators of differential treatment outcome in UP and diagnosis-specific CBT. In contrast to previous studies comparing UP and diagnosis-specific CBT, our sample of 291 patients included a subgroup with a primary diagnosis of MDD, thus allowing us to examine predictors and moderators in both anxiety disorders and MDD. We used a two-step machine learning approach to identify relevant predictors and moderators of treatment outcome rather than basing PAI treatment selection on simple linear regression models, as in earlier approaches. Our approach thus incorporated internal validation techniques such as bootstrapping to maximize the stability and generalizability of these models. The random forest approach is a comparatively stable model when n is relatively low and the number of predictors high (Bureau et al. 2005; Heidema et al. 2006). Without compromising the stability of the model, this approach also allows for the inclusion of many predictors and prevents weaker predictors from being dominated by stronger ones. In addition, the approach is capable of estimating linear and non-linear associations (Strobl et al. 2008).

Conclusion

The current study, which has compared UP with diagnosis-specific group CBT for anxiety and depression, did not enable us to identify moderators of treatment effect.

Although four predictors were identified, they were insufficiently robust to be selected across treatment sites. We cannot rule out that the lack of identified moderators and robust predictors may be a consequence of the factors that characterized our study: the chronicity of the sample, the chosen outcome measure, or the statistical methods used. We conclude that there is insufficient evidence to support a preference for either UP or diagnosis-specific group CBT for a given patient.

Acknowledgements The study was supported by TRYGFonden (ID 114241), the Capital Region Mental Health Services, Region Zealand Mental Health Services, the Jascha Foundation, the Psychiatric Research Foundation of Central Denmark Region, the Ivan Nielsen Foundation for Individuals with Particular Mental Disorders, and the Foundation for Research on Mental Disorders. The funding sources were not involved in the study design, collection, analysis, data interpretation, writing of the paper, or the decision to submit the article for publication. We are grateful to the UP Institute, Center for Anxiety and Related Disorders (CARD), Boston University for their invaluable advice concerning the manual and group delivery and to all the participating therapists and patients.

Compliance with Ethical Standards

Conflict of Interest Anita Eskildsen, Nina Reinholdt, Suzanne van Bronswijk, René B.K. Brund, Anne B. Christensen, Morten Hvenegaard, Mikkel Arendt, Anja Alrø, Stig Poulsen, Nicole K. Rosenberg, Marcus J.H. Huibers, and Sidse Arnfred declare that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the Danish National Committee on Health Research Ethics and with the 1964 Helsinki Declaration and its later amendments.

Informed Consent Informed consent was obtained from all individuals who participated in the study.

Research Involving Animal Rights No animal studies were carried out for this study.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Arch, J. J., & Ayers, C. R. (2013). Which treatment worked better for whom? Moderators of group cognitive behavioral therapy versus adapted mindfulness based stress reduction for anxiety disorders. *Behaviour Research and Therapy*, 51(8), 434–442. <https://doi.org/10.1016/J.Brat.2013.04.004>.
- Arendt, M., & Rosenberg, N. K. (2012). *Kognitiv Terapi [Cognitive Therapy]* (1st ed.). Copenhagen: Hans Reitzel.
- Arnfred, S., Aharoni, R., Hvenegaard, M., Poulsen, S., Bach, B., Arendt, M., et al. (2017). Transdiagnostic group Cbt vs standard group Cbt for depression, social anxiety disorder and agoraphobia/panic disorder: Study protocol for a pragmatic, multicenter non-inferiority randomized controlled trial. *Bmc Psychiatry*, 17(1), 37. <https://doi.org/10.1186/S12888-016-1175-0>.
- Austin, P. C., & Tu, J. V. (2004). Bootstrap methods for developing predictive models. *The American Statistician*, 58(2), 131–137. <https://doi.org/10.1198/0003130043277>.
- Barber, J. P., & Muenz, L. R. (1996). The role of avoidance and obsessiveness in matching patients to cognitive and interpersonal psychotherapy: Empirical findings from the treatment for depression collaborative research program. *Journal of Consulting and Clinical Psychology*, 64(5), 951–958. <https://doi.org/10.1037/0022-006x.64.5.951>.
- Barlow, D. H., Allen, L. B., & Choate, M. L. (2016). Toward a unified treatment for emotional disorders. In D. H. Barlow (Ed.), *The neurotic paradox: Progress in understanding and treating anxiety and related disorders* (1st ed., pp. 141–165). New York: Routledge/Taylor & Francis Group.
- Barlow, D. H., Ellard, K. K., Fairholme, C. P., Farchione, T. J., Boisseau, C. L., & Ehrenreich-May, J. (2011a). *Unified protocol for transdiagnostic treatment of emotional disorders: Work book*. New York: Oxford University Press.
- Barlow, D. H., Farchione, T. J., Bullis, J. R., Gallagher, M. W., Murray-Latin, H., Sauer-Zavala, S., et al. (2017). The unified protocol for transdiagnostic treatment of emotional disorders compared with diagnosis-specific protocols for anxiety disorders: A randomized clinical trial. *Jama Psychiatry*, 74(9), 875–884. <https://doi.org/10.1001/Jamapsychiatry.2017.2164>.
- Barlow, D. H., Farchione, T. J., Fairholme, C. P., Ellard, K. K., Boisseau, C. L., Allen, L. B., et al. (2011b). *Unified protocol for transdiagnostic treatment of emotional disorders: Therapist guide*. New York: Oxford University Press.
- Barlow, D. H., Farchione, T. J., Sauer-Zavala, S., Latin, H. M., Ellard, K. K., Bullis, J. R., et al. (2018). *Unified protocol for transdiagnostic treatment of emotional disorders: therapist guide* (2nd ed.). New York: Oxford University Press.
- Barlow, D. H., Sauer-Zavala, S., Carl, J. R., Bullis, J. R., & Ellard, K. K. (2014). The nature, diagnosis, and treatment of neuroticism: Back to the future. *Clinical Psychological Science*, 2(3), 344–365. <https://doi.org/10.1016/J.Cpr.2009.11.004>.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Bdi*. San Antonio, TX: Psychological Corporation.
- Berking, M., Wupperman, P., Reichardt, A., Pejic, T., Dippel, A., & Znoj, H. (2008). Emotion-regulation skills as a treatment target in psychotherapy. *Behaviour Research and Therapy*, 46(11), 1230–1237. <https://doi.org/10.1016/J.Brat.2008.08.005>.
- Blackburn, I., James, I. A., Milne, D. L., Baker, C., Standart, S., Garland, A., et al. (2001). The Revised Cognitive Therapy Scale (Cts-R): Psychometric properties. *Behavioural and Cognitive Psychotherapy*, 29(4), 431–446. <https://doi.org/10.1017/S1352465801004040>.
- Brown, T. A., Chorpita, B. F., & Barlow, D. H. (1998). Structural relationships among dimensions of the Dsm-Iv anxiety and mood disorders and dimensions of negative affect, positive affect, and autonomic arousal. *Journal of Abnormal Psychology*, 107(2), 179–192. <https://doi.org/10.1037//0021-843x.107.2.179>.
- Bullis, J. R., Sauer-Zavala, S., Bentley, K. H., Thompson-Hollands, J., Carl, J. R., & Barlow, D. H. (2015). The unified protocol

- for transdiagnostic treatment of emotional disorders: Preliminary exploration of effectiveness for group delivery. *Behavior Modification*, 39(2), 295–321. <https://doi.org/10.1177/0145445514553094>.
- Bureau, A., Dupuis, J., Falls, K., Lunetta, K. L., Hayward, B., Keith, T. P., et al. (2005). Identifying Snps predictive of phenotype using random forests. *Genetic Epidemiology*, 28(2), 171–182. <https://doi.org/10.1002/Gepi.20041>.
- Cammin-Nowak, S., Helbig-Lang, S., Lang, T., Gloster, A. T., Fehm, L., Gerlach, A. L., et al. (2013). Specificity of homework compliance effects on treatment outcome in Cbt: Evidence from a controlled trial on panic disorder and agoraphobia. *Journal of Clinical Psychology*, 69(6), 616–629. <https://doi.org/10.1002/Jclp.21975>.
- Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology*, 100(3), 316–336. <https://doi.org/10.1037//0021-843x.100.3.316>.
- Cohen, Z. D., & Derubeis, R. J. (2018). Treatment selection in depression. *Annual Review of Clinical Psychology*, 14(1), 209–236. <https://doi.org/10.1146/Annurev-Clinpsy-050817-084746>.
- Cohen, Z. D., Kim, T. T., Van, H. L., Dekker, J. J. M., & Driessen, E. (2020). A demonstration of a multi-method variable selection approach for treatment selection: Recommending cognitive-behavioral versus psychodynamic therapy for mild to moderate adult depression. *Psychotherapy Research*, 30(2), 137–150. <https://doi.org/10.1080/10503307.2018.1563312>.
- Craske, M. G., & Barlow, D. H. (2008). Panic sobia. In D. H. Barlow (Ed.), (4th ed., pp. 1–64). New York: Guilford Press
- Danish Health Authority. (2007a). *Referenceprogram for Angstlidelser Hos Voksne (Guideline)*. Copenhagen: Danish Health Authority.
- Danish Health Authority. (2007b). *Referenceprogram for unipolar depression Hos Voksne (guideline)*. Copenhagen: Danish Health Authority.
- De Ornelas Maia, A. C. C., Nardi, A. E., & Cardoso, A. (2015). The utilization of unified protocols in behavioral cognitive therapy in transdiagnostic group subjects: A clinical trial. *Journal of Affective Disorders*, 172(1), 179–183. <https://doi.org/10.1016/J.Jad.2014.09.023>.
- Delgadillo, J., Mcmillan, D., Lucock, M., Leach, C., Ali, S., & Gilbody, S. (2014). Early changes, attrition, and dose-response in low intensity psychological interventions. *British Journal of Clinical Psychology*, 53(1), 114–130. <https://doi.org/10.1111/Bjcp.12031>.
- Derubeis, R. J., Cohen, Z. D., Forand, N. R., Fournier, J. C., Gelfand, L. A., & Lorenzo-Luaces, L. (2014a). The personalized advantage index: Translating research on prediction into individualized treatment recommendations. A Demonstration. *PLoS ONE*, 9(1), E83875. <https://doi.org/10.1371/Journal.Pone.0083875>.
- Derubeis, R. J., Gelfand, L. A., German, R. E., Fournier, J. C., & Forand, N. R. (2014b). Understanding processes of change: How some patients reveal more than others-and some groups of therapists less-about what matters in psychotherapy. *Psychotherapy Research*, 24(3), 419–428. <https://doi.org/10.1080/10503307.2013.838654>.
- Due Madsen, J. (2008). *Kognitiv Adfaerdsterapi Ved depression*. Virum: Dansk Psykologisk Forlag.
- Dunlop, B. W., Kelley, M. E., Aponte-Rivera, V., Mletzko-Crowe, T., Kinkead, B., Ritchie, J. C., et al. (2017). Effects of patient preferences on outcomes in the predictors of remission in depression to individual and combined treatments (Predict) study. *The American Journal of Psychiatry*, 174(6), 546–556. <https://doi.org/10.1176/Appi.Ajp.2016.16050517>.
- Ehring, T., Zetsche, U., Weidacker, K., Wahl, K., Schonfeld, S., & Ehlers, A. (2011). The perseverative thinking questionnaire (Ptq): Validation of a content-independent measure of repetitive negative thinking. *The Journal of Behavior Therapy and Experimental Psychiatry*, 42(2), 225–232. <https://doi.org/10.1016/J.Jbtep.2010.12.003>.
- Fiedler, K. (2011). Voodoo correlations are everywhere—not only in neuroscience. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, 6(2), 163–171. <https://doi.org/10.1177/1745691611400237>.
- Fonagy, P., Luyten, P., Moulton-Perkins, A., Lee, Y., Warren, F., Howard, S., et al. (2016). Development and validation of a self-report measure of mentalizing: The reflective functioning questionnaire. *PLoS ONE*, 11(7), 1–28. <https://doi.org/10.1371/Journal.Pone.0158678>.
- Garge, N., Bobashev, G., & Eggleston, B. (2013). Random forest methodology for model-based recursive partitioning: The mobforest package for R. *BMC Bioinformatics*, 14(1), 125–125. <https://doi.org/10.1186/1471-2105-14-125>.
- Garge, N., Eggleston, B., Bobashev, G., Carper, B., Jones, K., Hothorn, T., et al. (2018). *Model Based Random Forest Analysis*. Retrieved from <https://Cran.R-Project.Org/Web/Packages/Mobforest/Index.Html>
- Gray, M. J., Litz, B. T., Hsu, J. L., & Lombardo, T. W. (2004). Psychometric properties of the life events checklist. *Assessment*, 11(4), 330–341. <https://doi.org/10.1177/1073191104269954>.
- Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, 85(2), 348–362. <https://doi.org/10.1037/0022-3514.85.2.348>.
- Hamilton, M. (1959). The assessment of anxiety states by rating. *British Journal of Medical Psychology*, 32(1), 50–55. <https://doi.org/10.1111/J.2044-8341.1959.Tb00467.X>.
- Heidema, A. G., Boer, J. M., Nagelkerke, N., Mariman, E. C., Van Der, A. D. L., & Feskens, E. (2006). The challenge for genetic epidemiologists: How to analyze large numbers of snps in relation to complex diseases. *Bmc Genetics*, 7(1), 23–23. <https://doi.org/10.1186/1471-2156-7-23>.
- 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>.
- Hofmann, S. G., Curtiss, J., & McNally, R. J. (2016). A complex network perspective on clinical science. *Perspectives on Psychological Science*, 11(5), 597–605. <https://doi.org/10.1177/1745691616639283>.
- Hofmann, S. G., & Hayes, S. C. (2019). The future of intervention science: Process-based therapy. *Clinical Psychological Science*, 7(1), 37–50. <https://doi.org/10.1177/2167702618772296>.
- Huibers, M. J. H., Cohen, Z. D., Lemmens, L. H. J. M., Arntz, A., Peeters, F. P. M. L., Cuijpers, P., et al. (2015). Predicting optimal outcomes in cognitive therapy or interpersonal psychotherapy for depressed individuals using the personalized advantage index approach. *PLoS ONE*, 10(11), E0140771. <https://doi.org/10.1371/Journal.Pone.0140771>.
- Hutsebaut, J., Feenstra, D. J., & Kamphuis, J. H. (2016). Development and preliminary psychometric evaluation of a brief self-report questionnaire for the assessment of the Dsm-5 level of personality functioning scale: The Lpfs Brief Form (Lpfs-Bf). *Personality Disorders: Theory, Research, and Treatment*. <https://doi.org/10.1037/Per0000159>.
- Kazantzis, N., Luong, H. K., Usatoff, A. S., Impala, T., Yew, R. Y., & Hofmann, S. G. (2018). The processes of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research*, 42(4), 349–357. <https://doi.org/10.1007/S10608-018-9920-Y>.
- Keefe, J. R., Wiltsey Stirman, S., Cohen, Z. D., Derubeis, R. J., Smith, B. N., & Resick, P. A. (2018). In rape trauma Ptsd, patient characteristics indicate which trauma-focused treatment they are most

- likely to complete. *Depression and Anxiety*, 35(4), 330–338. <https://doi.org/10.1002/Da.22731>.
- Kim, H., & Loh, W. (2001). Classification trees with unbiased multiway splits. *Journal of the American Statistical Association*, 96(454), 589–604. <https://doi.org/10.1198/016214501753168271>.
- Kraemer, H. C. (2013). Discovering, comparing, and combining moderators of treatment on outcome after randomized clinical trials: A parametric approach. *Statistics in Medicine*, 32(11), 1964–1973. <https://doi.org/10.1002/Sim.5734>.
- Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for Dsm-5. *Psychological Medicine*, 42(9), 1879–1890. <https://doi.org/10.1017/S0033291711002674>.
- Krueger, R. F., Watson, D., & Barlow, D. H. (2005). Introduction to the special section: Toward a dimensionally based taxonomy of psychopathology. *Journal of Abnormal Psychology*, 114(4), 491–493. <https://doi.org/10.1037/0021-843x.114.4.491>.
- Laposa, J. M., Mancuso, E., Abraham, G., & Loli-Dano, L. (2017). Unified protocol transdiagnostic treatment in group format: A preliminary investigation with anxious individuals. *Behavior Modification*, 41(2), 253–268. <https://doi.org/10.1177/0145445516667664>.
- 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(1), 72–82. <https://doi.org/10.1016/J.Cpr.2015.08.004>.
- Lorenzo-Luaces, L., Derubeis, R. J., Van Straten, A., & Tiemens, B. (2017). A prognostic index (Pi) as a moderator of outcomes in the treatment of depression: A proof of concept combining multiple variables to inform risk-stratified stepped care models. *Journal of Affective Disorders*, 213(1), 78–85. <https://doi.org/10.1016/J.Jad.2017.02.010>.
- Luedtke, A., Sadikova, E., & Kessler, R. C. (2019). Sample size requirements for multivariate models to predict between-patient differences in best treatments of major depressive disorder. *Clinical Psychological Science*, 7(3), 445–461. <https://doi.org/10.1177/2167702618815466>.
- Lutz, W., Saunders, S. M., Leon, S. C., Martinovich, Z., Kosfelder, J., Schulte, D., et al. (2006). Empirically and clinically useful decision making in psychotherapy: Differential predictions with treatment response models. *Psychological Assessment*, 18(2), 133–141. <https://doi.org/10.1037/1040-3590.18.2.133>.
- Marker, I., Salvaris, C. A., Thompson, E. M., Tolliday, T., & Norton, P. J. (2019). Client motivation and engagement in transdiagnostic group cognitive behavioral therapy for anxiety disorders: Predictors and outcomes. *Cognitive Therapy and Research*, 43(5), 819. <https://doi.org/10.1007/S10608-019-10014-1>.
- Mcevoy, P. M., Burgess, M. M., & Nathan, P. (2013). The Relationship between interpersonal problems, negative cognitions, and outcomes from cognitive behavioral group therapy for depression. *Journal of Affective Disorders*, 150(2), 266–275. <https://doi.org/10.1016/J.Jad.2013.04.005>.
- Mcevoy, P. M., Nathan, P., & Norton, P. J. (2009). Efficacy Of Transdiagnostic treatments: A review of published outcome studies and future research directions. *Journal of Cognitive Psychotherapy*, 23(1), 20–33. <https://doi.org/10.1891/0889-8391.23.1.20>.
- Moran, P., Leese, M., Lee, T., Walters, P., Thornicroft, G., & Mann, A. (2003). Standardised assessment of personality—abbreviated scale (Sapas): Preliminary validation of a brief screen for personality disorder. *The British Journal of Psychiatry*, 183(3), 228–232. <https://doi.org/10.1192/Bjp.183.3.228>.
- Nice. (2009). *Depression: Treatment and management of depression in adults (No. 90)*. London, UK: National Institute for Health and Clinical Excellence.
- Nice. (2011). *Common Mental health problems: Identification and pathways to care (No. Cg123)*. London, UK: National Institute For Health And Care Excellence.
- Nice. (2013). *Social anxiety disorder: Recognition, assessment and treatment (No. Cg159)*. London, UK: National Institute for Health and Care Excellence.
- Niles, A. N., Loerinc, A. G., Krull, J. L., Roy-Byrne, P., Sullivan, G., Sherbourne, C. D., et al. (2017a). Advancing personalized medicine: Application of a novel statistical method to identify treatment moderators in the coordinated anxiety learning and management study. *Behavior Therapy*, 48(4), 490–500. <https://doi.org/10.1016/J.Beth.2017.02.001>.
- Niles, A. N., Wolitzky-Taylor, K. B., Arch, J. J., & Craske, M. G. (2017b). Applying a novel statistical method to advance the personalized treatment of anxiety disorders: A composite moderator of comparative drop-out from Cbt and act. *Behaviour Research and Therapy*, 91(1), 13–23. <https://doi.org/10.1016/J.Brat.2017.01.001>.
- Norton, P. J., & Barrera, T. L. (2012). Transdiagnostic Versus diagnosis-specific Cbt for anxiety disorders: A preliminary randomized controlled noninferiority trial. *Depression and Anxiety*, 29(10), 874–882. <https://doi.org/10.1002/Da.21974>.
- Norton, P. J., & Hope, D. A. (2005). Preliminary evaluation of a broad-spectrum cognitive-behavioral group therapy for anxiety. *Journal of Behavior Therapy and Experimental Psychiatry*, 36(2), 79–97. <https://doi.org/10.1016/J.Jbtep.2004.07.002>.
- Osma, J., Castellano, C., Crespo, E., & García-Palacios, A. (2015). The unified protocol for transdiagnostic treatment of emotional disorders in group format in a spanish public mental health setting. *Behavioral Psychology*, 23(3), 447–466.
- Picard, R. R., & Cook, R. D. (1984). Cross-validation of regression models. *Journal of the American Statistical Association*, 79(387), 575–583. <https://doi.org/10.1080/01621459.1984.10478083>.
- Reinholt, N., Aharoni, R., Winding, C., Rosenberg, N., Rosenbaum, B., & Arnfred, S. (2017). Transdiagnostic group Cbt for anxiety disorders: The unified protocol in mental health services. *Cognitive Behaviour Therapy*, 46(1), 29–43. <https://doi.org/10.1080/16506073.2016.1227360>.
- Reinholt, N., Hvenegaard, M., Christensen, A., Eskildsen, A., Poulsen, S., Hjørthøj, C., et al. Transdiagnostic versus diagnosis-specific cognitive-behavioral therapy for anxiety and depressive disorders: A randomized clinical noninferiority trial (Tract-Rct). *Under Review*.
- Reinholt, N., & Krogh, J. (2014). Efficacy of transdiagnostic cognitive behaviour therapy for anxiety disorders: A systematic review and meta-analysis of published outcome studies. *Cognitive Behaviour Therapy*, 43(3), 171–184. <https://doi.org/10.1080/16506073.2014.897367>.
- Rizopoulos, D., & Rizopoulos, M. (2009). Package ‘Bootsteapaic’. Retrieved from https://Cran.R-Project.Org/Web/Packages/Boots_tepaic/Index.Html
- Robinson, L., Delgado, J., & Kellett, S. (2020). The dose-response effect in routinely delivered psychological therapies: A systematic review. *Psychotherapy Research*, 30(1), 79–96. <https://doi.org/10.1080/10503307.2019.1566676>.
- Sauer-Zavala, S., Boswell, J. F., Gallagher, M. W., Bentley, K. H., Ametaj, A., & Barlow, D. H. (2012). The role of negative affectivity and negative reactivity to emotions in predicting outcomes in the unified protocol for the transdiagnostic treatment of emotional disorders. *Behaviour Research and Therapy*, 50(9), 551–557. <https://doi.org/10.1016/J.Brat.2012.05.005>.
- Sewart, A. R., Niles, A. N., Burklund, L. J., Saxbe, D. E., Lieberman, M. D., & Craske, M. G. (2019). Examining positive and negative affect as outcomes and moderators of cognitive-behavioral therapy and acceptance and commitment therapy for social

- anxiety disorder. *Behavior Therapy*, 50(6), 1112–1124. <https://doi.org/10.1016/J.Beth.2019.07.001>.
- Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., et al. (1998). The mini-international neuropsychiatric interview (MINI): The development and validation of a structured diagnostic psychiatric interview for Dsm-IV and Icd-10. *The Journal of Clinical Psychiatry*, 59(Suppl 20), 22.
- Smagula, S. F., Wallace, M. L., Anderson, S. J., Karp, J. F., Lenze, E. J., Mulsant, B. H., et al. (2016). Combining moderators to identify clinical profiles of patients who will, and will not, benefit from aripiprazole augmentation for treatment resistant late-life major depressive disorder. *Journal of Psychiatric Research*, 81, 112–118. <https://doi.org/10.1016/J.Jpsychires.2016.07.005>.
- Springer, K. S., Levy, H. C., & Tolin, D. F. (2018). Remission in Cbt for adult anxiety disorders: A meta-analysis. *Clinical Psychology Review*, 61, 1–8. <https://doi.org/10.1016/J.Cpr.2018.03.002>.
- Stekhoven, D. J., & Bühlmann, P. (2012). Missforest-non-parametric missing value imputation for mixed-type data. *Bioinformatics*, 28(1), 112–118. <https://doi.org/10.1093/Bioinformatics/Btr597>.
- Strobl, C., Boulesteix, A., & Augustin, T. (2007). Unbiased split selection for classification trees based on the gini index. *Computational Statistics and Data Analysis*, 52(1), 483–501. <https://doi.org/10.1016/J.Csda.2006.12.030>.
- Strobl, C., Boulesteix, A., Kneib, T., Augustin, T., & Zeileis, A. (2008). Conditional variable importance for random forests. *BMC Bioinformatics*, 9(1), 307–307. <https://doi.org/10.1186/1471-2105-9-307>.
- Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging and random forests. *Psychological Methods*, 14(4), 323–348. <https://doi.org/10.1037/A0016973>.
- Sullivan, T. R., Lee, K. J., Ryan, P., & Salter, A. B. (2017). Multiple imputation for handling missing outcome data when estimating the relative risk. *Bmc Medical Research Methodology*, 17(1), 134. <https://doi.org/10.1186/S12874-017-0414-5>.
- Talkovsky, A. M., Green, K. L., Osegueda, A., & Norton, P. J. (2016). Secondary depression in transdiagnostic group cognitive behavioral therapy among individuals diagnosed with anxiety disorders. *Journal of Anxiety Disorders*, 46, 56–64. <https://doi.org/10.1016/J.Janxdis.2016.09.008>.
- Topp, C. W., Ostergaard, S. D., Sondergaard, S., & Bech, P. (2015). The who-5 well-being index: A systematic review of the literature. *Psychotherapy and Psychosomatics*, 84(3), 167–176. <https://doi.org/10.1159/000376585>.
- Triola, M. F. (2010). *Elementary statistics using excel* (4th ed.). Boston: Addison Wesley.
- Turk, C. L., Heimberg, R. G., & Magee, L. (2008). Social anxiety disorder. In D. H. Barlow (Ed.), (4th ed., pp. 123–163). New York: Guilford Press.
- Van Bronswijk, S., Derubeis, C., Lemmens, R. J., Peeters, L. H. J. M., Keefe, F. P. M. L., Cohen, J. R., et al. (2019a). Precision medicine for long-term depression outcomes using the personalized advantage index approach: Cognitive therapy or interpersonal psychotherapy? *Psychological Medicine*, 1, 2. <https://doi.org/10.1017/S0033291719003192>.
- Van Bronswijk, S. C., Lemmens, L. H. J. M., Keefe, J. R., Huibers, M. J. H., Derubeis, R. J., & Peeters, F. P. M. L. (2019b). A prognostic index for long-term outcome after successful acute phase cognitive therapy and interpersonal psychotherapy for major depressive disorder. *Depression and Anxiety*, 36(3), 252–261. <https://doi.org/10.1002/Da.22868>.
- Vittengl, J. R., Clark, L. A., Thase, M. E., & Jarrett, R. B. (2017). Initial steps to inform selection of continuation cognitive therapy or fluoxetine for higher risk responders to cognitive therapy for recurrent major depressive disorder. *Psychiatry Research*, 253(1), 174–181. <https://doi.org/10.1016/J.Psychres.2017.03.032>.
- Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly high correlations in Fmri studies of emotion, personality, and social cognition. *Perspectives on Psychological Science*, 4(3), 274–290. <https://doi.org/10.1111/J.1745-6924.2009.01125.X>.
- Waljee, A. K., Mukherjee, A., Singal, A. G., Zhang, Y., Warren, J., Balis, U., et al. (2013). Comparison of imputation methods for missing laboratory data in medicine. *British Medical Journal Open*, 3(8), E002847. <https://doi.org/10.1136/Bmjopen-2013-002847>.
- Wallace, M. L., Frank, E., & Kraemer, H. C. (2013). A novel approach for developing and interpreting treatment moderator profiles in randomized clinical trials. *Jama Psychiatry*, 70(11), 1241–1247. <https://doi.org/10.1001/Jamapsychiatry.2013.1960>.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The panas scales. *Journal of Personality and Social Psychology*, 54(6), 1063–1070. <https://doi.org/10.1037//0022-3514.54.6.1063>.
- Webb, C. A., Trivedi, M. H., Cohen, Z. D., Dillon, D. G., Fournier, J. C., Goer, F., et al. (2019). Personalized prediction of antidepressant V. Placebo Response: Evidence from the embarc study. *Psychological Medicine*, 49(7), 1118–1127. <https://doi.org/10.1017/S0033291718001708>.
- Westra, H. A., Dozois, D. J. A., & Marcus, M. (2007). Expectancy, homework compliance, and initial change in cognitive-behavioral therapy for anxiety. *Journal of Consulting and Clinical Psychology*, 75(3), 363–373. <https://doi.org/10.1037/0022-006x.75.3.363>.
- Whiteford, H. A., Degenhardt, L., Rehm, J., Baxter, A. J., Ferrari, A. J., Erskine, H. E., et al. (2013). Global burden of disease attributable to mental and substance use disorders: Findings from the global burden of disease study 2010. *The Lancet*, 382(9904), 1575–1586. [https://doi.org/10.1016/S0140-6736\(13\)61611-6](https://doi.org/10.1016/S0140-6736(13)61611-6).
- Wittchen, H. U., Jacobi, F., Rehm, J., Gustavsson, A., Svensson, M., Jönsson, B., et al. (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>.
- Zemestani, M., Imani, M., & Ottaviani, C. (2017). A preliminary investigation on the effectiveness of unified and transdiagnostic cognitive behavior therapy for patients with comorbid depression and anxiety. *International Journal of Cognitive Therapy*, 10(2), 175–185. <https://doi.org/10.1521/Ijct.2017.10.2.175>.
- Zilcha-Mano, S. (2017). Is the alliance really therapeutic? Revisiting this question in light of recent methodological advances. *American Psychologist*, 72(4), 311–325. <https://doi.org/10.1037/A0040435>.
- Zilcha-Mano, S., Keefe, J. R., Chui, H., Rubin, A., Barrett, M. S., & Barber, J. P. (2016). Reducing dropout in treatment for depression: Translating dropout predictors into individualized treatment recommendations. *The Journal of Clinical Psychiatry*, 77(12), E1584–E1590. <https://doi.org/10.4088/Jcp.15m10081>.