Journal of Affective Disorders Reports 26 (2026) 101202

Contents lists available at ScienceDirect

Journal of Affective Disorders Reports

journal homepage: www.sciencedirect.com/journal/journal-of-affective-disorders-reports

Research Paper

Personalized treatment selection between cognitive behavioural therapy and short-term psychodynamic supportive psychotherapy for adult outpatient depression: a proof-of-concept application of the Personalized Advantage Index

Susanne Bremer-Hoeve [a,d,*], Suzanne C. van Bronswijk [b,c], Aartjan Beekman [d], Maartje F. Miggiels [a,d], Sipke P. Stuiver [e], Ellen Driessen [f,g], Maarten K. van Dijk [a]

[a] Dimence Institute of Mental Health, Deventer, the Netherlands
[b] Department of Psychiatry and Neuropsychology, School for Mental health and Neuroscience, Faculty of Health Medicine and Life Sciences, Maastricht University, Maastricht, the Netherlands
[c] Department of Psychiatry and Psychology, Maastricht University Medical Center, Maastricht, the Netherlands
[d] Department of Psychiatry, Amsterdam University Medical Center, Amsterdam, the Netherlands
[e] Context psychotherapy, Deventer, the Netherlands
[f] Department of Clinical Psychology, Behavioural Science Institute, Radboud University, Nijmegen, the Netherlands
[g] Depression Expertise Center, Pro Persona Mental Health Care, Nijmegen, the Netherlands

ARTICLE INFO

Keywords:
Depression
Machine learning
Cognitive behavioural therapy
Short-term psychodynamic supportive psychotherapy
Precision medicine

ABSTRACT

Objective: Minimal efficacy differences have been found between cognitive behavioural therapy (CBT) and short-term psychodynamic supportive psychotherapy (SPSP) but it is unclear which of these treatments work best for whom. This study aimed to explore the feasibility of applying a Personalized Advantage Index (PAI) as a proof-of-concept approach to optimize treatment outcomes by tailoring the selection of Cognitive Behavioral Therapy (CBT) or Short-term Psychodynamic Supportive Psychotherapy (SPSP) for individuals with depression.
Method: Data gathered from a randomized controlled trial comparing short-term psychodynamic supportive psychotherapy (SPSP, n = 148) and cognitive behavioural therapy (CBT, n = 142) for adults with depression were used. Elastic-Net regularization was applied to identify pre-treatment patient characteristics associated with treatment outcome, explicitly to reduce overfitting and manage multicollinearity given the exploratory nature and sample size of the study. The PAI approach was utilized to generate personalized treatment recommendations. Finally, we examined whether outcomes differed between individuals who did or did not receive their PAI-indicated treatment.
Results: A significant difference in post-treatment depression severity scores was found between those receiving their PAI-indicated treatment and those receiving their non-indicated PAI treatment ($d = 0.36$).
Conclusion: Although based on a modest sample and without external validation, the findings suggest that the Personalized Advantage Index (PAI) may hold promise as a feasibility-oriented, proof-of-concept approach for precision in psychotherapy allocation. However, the results should be interpreted cautiously, and replication in independent samples is required before clinical implementation can be considered.

1. Introduction

The predictive relationship between pre-treatment individual characteristics and treatment outcomes has been studied by many researchers in Major Depressive Disorder (MDD, Kravitz et al., 2004; Huibers et al., 2015; Vittengl et al., 2017; Cohen and DeRubeis, 2018; Cohen et al., 2020). While the average effects of various psychotherapy approaches for MDD is generally comparable (Cuijpers et al., 2021; Lorenzo-Luaces et al., 2021) individual responses to these modalities can vary substantially. Consequently, one type of psychotherapy might

* Corresponding author at: Dimence Institute of Mental Health, Deventer, the Netherlands.
E-mail addresses: s.bremer@dimencegroep.nl (S. Bremer-Hoeve), scvanbronswijk@gmail.com (S.C. van Bronswijk), a.beekman@amsterdamumc.nl (A. Beekman), m.miggiels@dimence.nl (M.F. Miggiels), info@ptcontext.nl (S.P. Stuiver), ellen.driessen@ru.nl (E. Driessen), m.vandijk@dimence.nl (M.K. van Dijk).

https://doi.org/10.1016/j.jadr.2026.101202
Received 4 June 2026; Received in revised form 21 September 2026; Accepted 19 October 2026
Available online 19 October 2026
2666-9153/© 2026 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

prove more advantageous for a specific individual compared to other available options (Cohen and DeRubeis, 2018; Kaiser et al., 2022; Simon and Perlis, 2010). So, understanding of baseline characteristics that influence the magnitude of treatment response, so called moderators, is essential. For example, in a direct comparison of two treatments, Driessen et al. (2016) found that short-term psychodynamic supportive psychotherapy (SPSP) was relatively more effective for patients with low comorbid anxiety and with a longer duration of the current depressive episode, whereas patients with a shorter depressive episode appeared to benefit more from cognitive behavioural therapy (CBT). Later, Cohen et al. (2020) identified additional moderators, including personality characteristics such as extraversion and anxiety sensitivity.

This variability underscores the potential benefits of personalized treatment strategies, where therapy selection is informed by individual patient profiles. One promising approach to achieve this involves the use of multivariable prediction models that incorporate predictive individual characteristics to guide treatment selection. Selecting the optimal treatment based on these personalized predictions presents a promising approach to optimize depression outcomes (Cohen and DeRubeis, 2018).

An example of such a treatment selection approach is the Personalized Advantage Index (PAI). The PAI represents a preference score that indicates the extent to which one treatment is more advantageous than the other (DeRubeis et al., 2014). PAI scores are based on multivariable models that rely on various statistical and machine learning techniques.

Since DeRubeis et al. (2014) developed the PAI approach for cognitive therapy (CT) and antidepressant medication (ADM), it has been replicated and extended to various treatment options for depression, including continuation CT versus fluoxetine (Vittengl et al., 2017), CT versus interpersonal psychotherapy (IPT) (van Bronswijk et al., 2019) and CBT versus short-term psychodynamic supportive psychotherapy (SPSP) (Cohen et al., 2020). In the PAI study comparing SPSP and CBT, Cohen et al. (2020) identified several patient groups for whom SPSP might be more effective than CBT, and vice versa. For example, SPSP was found to produce better outcomes for introverted individuals and for those who reported higher levels of anxiety sensitivity at the start of treatment (Cohen et al., 2020). Although the specific predictors contributing to individualized treatment recommendations varied between studies, these findings consistently support the broader premise of individualized treatment allocation rather than the stability of any single predictor. At the same time, most PAI models have been developed using datasets of limited size, which constrains their generalizability. Consequently, existing PAI studies are best interpreted as exploratory or proof-of-concept investigations that demonstrate the feasibility and potential clinical relevance of personalized treatment selection, rather than as finalized predictive algorithms ready for clinical implementation. Randomized controlled trials are generally not designed or powered to detect effect-modifying factors, let alone to provide definitive tests of complex multivariable models such as the Personalized Advantage Index (PAI). Conducting trials large enough to achieve this would require larger sample sizes—for instance, simulations by Luedtke et al. (2019) suggest that at least 300 patients per treatment arm (≈600 total) are required to ensure adequate statistical power for detecting clinically meaningful treatment selection effects, which is highly unlikely in psychotherapy research. Given this limitation, it is crucial to leverage *existing* data from rigorously conducted head-to-head trials comparing different psychotherapies. The aim of this study is therefore twofold: (1) to explore the potential usefulness of the PAI approach for optimizing treatment allocation in depression, explicitly framed as another proof-of-concept analysis, and (2) to contribute to a growing database that can facilitate future meta-analyses of relevant data.

2. Method

2.1. Design

Data were used from a two-phase RCT comparing SPSP and CBT, that was called the D*Phase study (Miggiels et al., 2021, 2025) and that took place at Dimence, an institute of mental health in the Netherlands. The study design was approved by the medical research ethics committee of Medical Centre Leeuwarden (MREC: RTPO970, CCMO: NL 56, 047.099.16) and is described in more detail in the published study protocol (Miggiels et al., 2021). In the study protocol, the intention to conduct a PAI analysis was mentioned in the statistical analyses section. However, detailed procedures for PAI development were not prespecified in the protocol and were explicitly deferred to subsequent publications.

The D*Phase study consisted of two treatment phases. In the first phase patients were randomised 1:1 to receive 16 sessions of either SPSP or CBT over eight weeks. Patients who did not sufficiently respond to this first phase of treatment (defined as <50% improvement on the IDS-SR) were subsequently offered a second treatment phase, in which they were randomised 1:1:1 to (a) continuing the initial treatment with the same therapist, (b) continuing the initial treatment with another therapist, or (c) switching to the other treatment modality with another therapist, for up to another 16 twice-weekly sessions (Miggiels et al., 2021). The present PAI analyses are based exclusively on data from the first treatment phase.

2.2. Participants

All participants were referred to one of the outpatient clinics specialized in the treatment of anxiety and mood disorders. Participants were included if they met the criteria for single or recurrent unipolar major depression with a moderate or severe episode as primary diagnosis confirmed by the Mini-International Neuropsychiatric Interview (Sheehan et al., 1998).

Exclusion criteria were insufficient understanding of the Dutch language, presence of psychotic symptoms, substance dependency (with the exception of nicotine), high risk of suicide requiring immediate intervention, unstable dosage of ADM (same minimal effective target dose of medication for less than six weeks), having received psychotherapeutic treatment in the last year that was carried out in specialized mental healthcare, and the inability to commit to treatment requirements (i.e., attending therapy sessions).

After the initial diagnostic examination, individuals who provided written informed consent were randomly assigned to one of the treatment conditions. A total of 308 individuals were included, and 290 individuals were randomized (SPSP: 148, CBT: 142). For more details see Miggiels et al. (2025).

2.3. Interventions

The first phase of the treatment was an eight-week period in which participants received 16 sessions of SPSP or CBT. All these sessions were given by therapists who had at least a master's level in medicine or psychology, including master-level psychologists, clinical psychologists, psychotherapists, residents in psychiatry, and psychiatrists,and completed a three-day course in both treatments. Treatment was delivered by a total of 40 therapists, who had a mean of 10.3 years (SD = 7.8) of general clinical experience and 8.7 years (SD = 7.0) of experience treating depressed patients specifically; on average, each therapist provided 7.2 (SD = 5.4) treatment trajectories within the study (Miggiels et al., 2025).

To facilitate and assess treatment integrity (competence and adherence), each treatment session was audio-recorded. These recordings were discussed during two-weekly group supervision provided by trained supervisors. The SPSP training and supervision sessions were

based on the Handbook of Short-Term Psychoanalytic Supportive Psychotherapy (Jonghe de, 2014). The CBT training was based on the Dutch protocol for CBT (Bockting and Huibers, 2011). Due to the COVID19 pandemic, therapy sessions were provided online through video calls from March 2020 to September 2022. A total of 56 individuals (22 CBT and 34 SPSP) received one or more online therapy sessions. There were no significant differences in the post-treatment depression severity scores between individuals with at least one online therapy session ($M = 31.3$, $SD = 14$) and those with only face-to-face therapy sessions ($M = 30.1$, $SD = 14.45$, $t(263) = .53$, $p = .60$). In addition, there were no significant differences in the number of sessions between individuals with at least one online therapy session ($M = 14.2$, $SD = 2.66$) and those with only face-to-face therapy sessions ($M = 13.4$, $SD = 2.98$, $t(263) = 1.86$, $p = .06$).

2.4. Measurements

2.4.1. Outcome

The primary outcome measure for the D*Phase study was the self-report version of the Inventory of Depressive Symptomatology (IDS-SR; Dutch translation Altrecht GGZ; Rush et al., 1986, 1996). It consists of 30 items scored on a 4-point scale. The IDS-SR is a treatment-sensitive measure for outpatients with depression and has adequate psychometric properties (Rush et al., 2000; Trivedi et al., 2004). In the current sample, internal consistency at baseline was good (Cronbach's $\alpha = 0.87$).

2.4.2. Pre-treatment variables

Pre-treatment variables, which provided information on specific individual characteristics, were all measured at baseline. Baseline variables were assessed in this RCT when previous research indicated them to be predictive of treatment outcome (Kessler et al., 2017; Miggiels et al., 2021). These baseline variables included social demographics, depression and clinical characteristics, comorbidity, and measures of stress and adversity.

2.5. Statistical analysis

2.5.1. Variable description and missing data

Data management and analyses were performed using RStudio (RStudio Team, 2023).

A correlation matrix was computed for all the variables to detect collinearity. No variables needed to excluded from the analyses due to a high correlation (cor. $>$ 0.70). Variance Inflation Factors (VIF) were additionally calculated for all baseline variables, all of which were below 3.6 (range 1.07–3.54), confirming the absence of problematic multicollinearity.

Eight individuals were excluded from the study due to significant changes in ADM during the treatment period, which may influence the depression outcome and might therefore bias the results. In addition, following Cohen et al. (2020), three other individuals were excluded from the study on the basis of having $\geq$50% missing baseline predictors. Finally, 32 individuals who attended fewer than four therapy sessions were excluded from the analyses. Previous research has shown that reliable and clinically meaningful change estimates are difficult to establish in very brief treatment episodes. For example, Delgadillo et al. (2014) and Robinson et al. (2020) observed that clients typically require several sessions before reliable or clinically significant improvement can be detected, with most improvement occurring after approximately four sessions. The final dataset thus included 265 participants (CBT: 129, SPSP: 136). Missing data imputation of the IDS-SR outcome and pre-treatment variables was performed with an iterative imputation method based on a random forest algorithm (R package 'MissForest', Stekhoven and Bühlmann, 2012). MissForest was chosen because it can handle mixed variable types and model complex relationships without requiring distributional assumptions (Stekhoven and Bühlmann, 2012). To assess its performance in the current dataset, observed values were artificially masked completely at random at a rate matching the actual missingness in the dataset, imputed, and subsequently compared with their original values. This evaluation provides an indication of imputation accuracy within the present dataset rather than a general validation of the method across datasets (Stekhoven and Bühlmann, 2012).

2.5.2. Variable transformation

Categorical variables were made binary, and to prevent computational issues, binary variables without adequate variance (smallest category $<$ 10%) were removed (Kuhn and Johnson, 2013). Outliers for continuous variables were winsorized and skewed variables were transformed using the Box–Cox method (Box and Cox, 1964; R package 'Caret'; Kuhn, 2008). An important advantage of the Box-Cox method is that it searches for the most suitable transformation (Kuhn and Johnson, 2013). Finally, the continuous variables were standardized, and categorical variables were mean-centered by recoding the categories (−0.5 and 0.5).

2.5.3. Model development

Elastic-Net Regularization (ENR, R package 'glmnet', Friedman et al., 2010) was used to identify variables related to treatment outcome. ENR is a highly interpretable technique that combines the advantages of Lasso and Ridge Regression by combining penalizations. The advantages of ENR are that it copes well with a high number of variables and correlations between them (Friedman et al., 2010; Kuhn and Johnson, 2013). To identify moderators, prognostic models for each treatment group were created. The results of these prognostic models contain variables with their coefficient values, indicating an association between the variables and the dependent variable. Variables were chosen as potential moderators by either appearing exclusively in one model or by appearing in both models but with distinct coefficient values (Cohen et al., 2019). ENR as a technique was hereby selected specifically to reduce the chance of overfitting and manage multicollinearity, an important safeguard given the relatively modest sample size and exploratory nature of the present analysis.

2.5.4. Generating the PAI

Once variables were selected using ENR, personalized treatment recommendations were generated using the PAI approach (DeRubeis et al., 2014). The prognostic models were used to predict the outcome for each individual in CBT and SPSP using tenfold cross-validation (CV). It should be emphasized that this cross-validation procedure was conducted internally within the same dataset and does not constitute external validation. Consequently, the developed PAI model is intended to evaluate conceptual feasibility rather than predictive performance in independent samples. The tenfold CV divides the sample randomly into ten approximately equal groups. The model was fitted for nine groups, and post-treatment IDS-SR score predictions were generated for each treatment in the holdout group. This process is repeated 10 times. The differences between these two predicted IDS-SR scores in CBT and SPSP is the individuals PAI score, indicating the optimal treatment. The predicted difference score in post-treatment IDS-SR was the dependent variable in the developed PAI model with ENR, using the selected baseline variables in the CBT and SPSP model as potential predictors.

2.5.5. Evaluating the PAI model

The overall performance of the model was evaluated using R^2 that indicates the amount of variability in IDS-SR explained by the pre-treatment variables with a value between 0 and 1 (the closer to 1 the better; Kuhn and Johnson, 2013).

To determine the utility of the PAI for treatment recommendation, the average post-treatment IDS-SR score of individuals receiving their PAI-indicated treatment were compared with that of individuals receiving their PAI-non-indicated treatment using *t*-tests and effect sizes (Cohen's d). Additionally, this comparison was also applied separately for two PAI treatment indications (CBT and SPSP) as well. For

individuals whose PAI indicated CBT, the post-treatment IDS-SR scores were compared between individuals who received CBT (indicated) and those who received SPSP (non-indicated). For individuals whose PAI indicated SPSP, the post-treatment IDS-SR scores were compared between individuals who received SPSP (indicated) and those who received CBT (non-indicated; DeRubeis et al., 2014).

3. Results

3.1. Variable description and missing data

Table 1 presents an overview of the baseline variables. Of all baseline variables, 0.9% of all values were missing and 9.8% values of the outcome variable were missing and thus imputed. The performance of the imputation method was accurate when applied to a dataset with simulated missing values (*NRMSE* = 0.08 and *PFC* = 0.17).

3.2. Model development

A total of eleven out of twenty variables were selected in the CBT model using ENR, with an R^2 of 0.47, and the SPSP model consisted of four out of twenty variables with an R^2 of 0.37 (see Supplement 1). A total of twelve out of twenty variables were identified when comparing the two ENR models. The following variables were selected: age, gender, employment, household income, education level, pre-treatment depression severity, psychological distress, functional impairment, positive mental health, personality organization, defense mechanisms, and traumatic life events. The PAI model was developed using ENR with the twelve variables selected from the ENR models with the predicted difference score in post-treatment IDS-SR as the dependent variable.

3.3. Evaluating PAI

There was a significant difference in the post-treatment depression severity scores between individuals receiving PAI-indicated treatment (M = 27.7, SD = 15.04, N = 129) and those receiving non-indicated PAI treatment (M = 32.8, SD = 13.23, N = 136, t(263) = 2.9, p = 0.003, Cohen's d = .36). At the same time, these two groups did not significantly differ with regard to other relevant characteristics as measured at baseline, such as baseline level of depression severity, depression duration or comorbidity. Compared to those indicated for CBT, patients indicated for SPSP tended to be younger, more often male, more likely to live below the poverty line, and to show greater functional impairment (SDS), more impaired personality organization (IPO), lower self-esteem (RSES), and more immature defense mechanisms (DSQ-42).

As shown in Fig. 1, individuals receiving SPSP with a PAI-indication SPSP reported lower post-treatment depression severity scores than individuals receiving CBT with a PAI-indication SPSP (t(208) = 2.4, p = 0.02, Cohen's d = .33), and those receiving CBT with a PAI-indication CBT reported lower post-treatment depression severity scores than individuals receiving non-indicated treatment SPSP (t(53) = 1.8, p = 0.085, Cohen's d = .48), although this last result did not reach conventional levels of statistical significance.

4. Discussion

This study examined individual treatment recommendations for either SPSP or CBT, using the PAI approach. The aims of this study were (1) to explore the potential usefulness of the PAI approach for optimizing treatment allocation in depression, explicitly framed as a another proof-of-concept analysis, and (2) to contribute to a growing database that can facilitate future meta-analyses of relevant data.

Using machine learning techniques, twelve pre-treatment variables were used to calculate PAI scores for each individual that determined the indicated treatment. These are all variables that are regularly available in routine clinical care and have clinical face-validity. There was a significant difference in post-treatment depression severity scores between individuals receiving their PAI-indicated treatment and those receiving their non-indicated PAI treatment with a small-to-moderate effect size (Cohen's *d* = .36). These matching effects occurred within

Table 1
Description and comparison of pre-treatment variables in short-term psychodynamic supportive psychotherapy v. cognitive behavioural therapy.

	Short-term psychodynamic supportive psychotherapy (n = 136)	Cognitive behavioural therapy (n = 129)
Social demographics		
Age, mean (S.D.)	34.5 (13.2)	37 (13.2)
Female gender, n (%)	81 (60%)	79 (61%)
Marital status, married, n (%)	35 (26%)	38 (29%)
Education level		
Low, n (%)	19 (14%)	14 (11%)
Intermediate, n (%)	87 (64%)	80 (62%)
High, n (%)	30 (22%)	35 (27%)
Household income		
Below poverty line, n (%)	50 (37%)	32 (25%)
Don't know or won't tell, n (%)	30 (22%)	34 (26%)
Above poverty line, n (%)	56 (41%)	63 (49%)
Employment, yes, n (%)	82 (60%)	73 (57%)
Depression & clinical characteristics		
Pharmacotherapy at start treatment, n (%)	57 (42%)	53 (41%)
Episode Duration (> 2 years), n (%)	53 (39%)	56 (43%)
Depression severity, IDS-SR, mean (S.D.)	40.1 (9.9)	40.6 (10.1)
Defence mechanisms, DSQ, mean (S.D.)	184 (27.6)	182 (26.7)
Comorbidity		
Comorbid classification according to the MINI, yes, n (%)	58 (43%)	51 (40%)
Psychological distress, BSI, mean (S.D.)	2.1 (.7)	2 (0.7)
Agoraphobia, AGO, mean (S.D.)	10 (8.9)	10.3 (8.7)
Self-esteem, RSES, mean (S.D.)	11.1 (4.2)	11.8 (4.4)
Rumination, RRS, mean (S.D.)	56.1 (10.2)	56.1 (10.1)
Personality organization, IPO, mean (S.D.)	164.5 (34.5)	159.2 (34.4)
Stress and adversity		
Life events (lifetime), LEC-5, mean (S.D.)	2.8 (1.3)	2.9 (1.3)
Traumatic life-events, List of Threatening Experiences, mean (S.D.)	18.9 (9.4)	20.1 (9.9)
Other		
Positive mental health, MHC-SF, mean (S.D.)	22.9 (10)	21.5 (10)
Functional impairment, SDS, mean (S.D.)	21 (4.8)	21.2 (4.2)

AGO, Agoraphobia Scale; BSI, Brief Symptom Inventory; DSQ, Defense Style Questionnaire; IDS-SR, Inventory of Depressive Symptomatology – Self-Report; IPO, Inventory of Personality Organisation; LEC-5, Checklist Life Events for DSM-5; MHC-SF, Mental Health Continuum-Short Form; RRS, Rumination Response Scale; RSES, Rosenberg Self-Esteem Scale; SDS, Sheehan Disability Scale.
* $p < 0.05$.

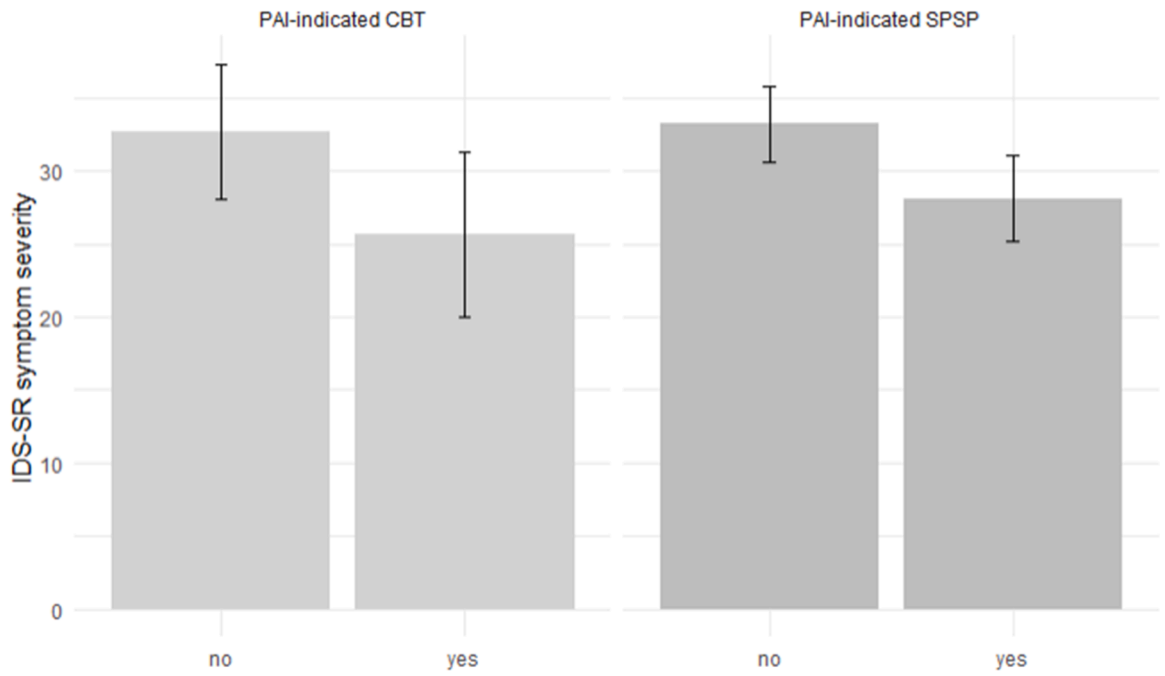


Fig. 1. Comparison of the observed post-treatment IDS-SR symptom severity scores for individuals randomized to their indicated treatment versus individuals randomized to their non-indicated treatment.
Note. CBT = cognitive behavioural therapy, PAI = Personalized Advantage Index, SPSP = short-term psychodynamic supportive psychotherapy.

both treatment conditions as well (SPSP: Cohen's $d = .33$, CBT: Cohen's $d = .48$), although the result pertaining to CBT did not reach conventional levels of statistical significance. The observed profile of patients indicated for SPSP, characterized by greater functional impairment, more impaired personality organization, and more immature defense mechanisms, is broadly consistent with psychodynamic theory and previous research emphasizing the relevance of personality functioning and interpersonal difficulties in psychodynamic treatment (Cohen et al., 2020; Driessen et al., 2016; De Jonghe, 2014). In contrast, differences in age, gender, and income are less clearly supported by the existing literature and may reflect local treatment allocation practices rather than established treatment-matching principles.

The small-to-moderate effect sizes in post-treatment depression severity scores between individuals receiving PAI-indicated treatment and those receiving non-indicated PAI treatment were similar to effect sizes found in previous PAI studies (Cohen et al., 2020; DeRubeis et al., 2014; Van Bronswijk et al., 2021a, 2021b). Although the specific predictors and the set of baseline variables used to derive the PAI models differed from those used and identified by Cohen et al. (2020) in their study that also focused on CBT and SPSP, the presence of a comparable matching effect supports the conceptual robustness of the PAI, also for these two types of treatment. These presence of such an additional, small-to-moderate matching effect could hold significant relevance and impact in optimizing therapeutic interventions in clinical practice. It also shows the potential of moving the psychotherapy field towards more tailored clinical decision making.

Although the present study found a small-to-moderate effect size, the findings should be interpreted with caution because the model was created and tested on the same dataset, an issue described as double-dipping (Vul et al., 2009; Fiedler, 2011). Ideally, the train/test method should be used to prevent overfitting of the model (James et al., 2003). Overfitting implies that the algorithm fits the developed dataset exceptionally well but demonstrates poor performance in predicting outcomes for new and unseen data (Bleeker et al., 2003; Kuhn, 2008; Lorenzo-Luaces et al., 2021). Using the train/test method was not possible in this study because of the insufficient sample size. Furthermore, this relatively small sample size for applying machine learning techniques was a second limitation, which also increases the risk of overfitting. Prior simulation study by Luedtke et al. (2019) indicated that multivariable prediction models require larger sample sizes. However, to reduce the risk of overfitting due to double dipping and a small sample sizes, 10-fold cross-validation was applied in the model building process. The third limitation was the lack of external validation. This is an essential next step for clinical applicability of the PAI, as it assesses the extent to which the findings can be extrapolated to different circumstances. The generalisability of our findings is further constrained by the composition of the sample. For example, the present analytic sample was predominantly female. Predictors should therefore be interpreted with this context in mind, and it remains an open question whether the PAI predictors or matching effect would generalise to more socioeconomically diverse samples. A fourth limitation concerns the training of therapists delivering both treatments: all received only a three-day course covering both modalities, which is substantially shorter than the years of specialised training typically required for proficiency. This may have affected the treatments asymmetrically, as SPSP was relatively new for most therapists compared with CBT. Although treatment integrity was monitored throughout via audio-recorded sessions and biweekly group supervision, this may not have fully compensated for the brief training and could have attenuated the observed PAI matching effect if patients whose PAI scores indicated SPSP were treated by comparatively less experienced therapists in that modality.

Because of these limitations results must thus be interpreted cautiously, and they necessitate a proof-of-concept interpretation rather than confirmation of predictive validity. Despite these limitations, the current findings suggest that PAI-guided treatment allocation warrants further investigation. Future research should prioritize external validation, pooled analyses across trials, and prospective evaluation in naturalistic clinical settings.

Ethical approval

The study was registered with the Netherlands Trial Register on August 26, 2016 (NL5753). The study has been approved by the medical

research ethical committee of Medical Centre Leeuwarden (number: METC: RTPO970, registration number NL 56047.099.16).

Informed consent from participants

The patients / participants provided their written informed consent to participate in this study.

Data statement

Individual participant data underlying the results in this article and analytical codes, will be shared after de- identification. Data will be shared only with researchers who provide a methodologically sound proposal and whose proposed use of the data has been approved by an independent review committee. Data can be used only to achieve the aims in the approved proposal. Proposals should be directed to m.van-dijk@dimencegroep.nl. Proposals may be submitted up to 36 months following article publication. A data access agreement is required and all access must comply with regulatory restrictions imposed on the original study.

Funding

This study was funded by Dimence Institute of Mental Health and the VCVGZ association.

CRediT authorship contribution statement

Susanne Bremer-Hoeve: Writing – original draft, Validation, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Suzanne C. van Bronswijk:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Aartjan Beekman:** Writing – review & editing, Supervision. **Maartje F. Miggiels:** Writing – review & editing, Project administration, Funding acquisition, Data curation, Conceptualization. **Sipke P. Stuiver:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Ellen Driessen:** Writing – review & editing, Data curation. **Maarten K. van Dijk:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Many thanks to the contribution of participants and staff of Dimence Institute of Mental Health who were involved in the study.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jadr.2026.101202.

References

Bleeker, S.E., Moll, H.A., Steyerberg, E.W., Donders, A.R., Derksen-Lubsen, G., Grobbee, . E., Moons, K.G, 2003. External validation is necessary in prediction research: a clinical example. J. Clin. Epidemiol. 56 (9), 826–832. https://doi.org/10.1016/s0895-4356(03)00207-5.

Bockting, C.L.H., Huibers, M.J.H., 2011. Protocollaire behandeling van patiënten met een depressieve stoornis. – Cognitieve gedragstherapie. In: Keijsers, G.P.J., van Minnen, A., Hoogduin, C.A.L. (Eds.), Protocollaire Behandelingen Voor Volwassenen Met Psychische Klachten, Protocollaire Behandelingen Voor Volwassenen Met Psychische Klachten, 1, pp. 251–288.

Box, G.E.P., Cox, D.R., 1964. An Analysis of transformations. J. R. Stat. Soc. Ser. B Methodol. 26 (2), 211–243. http://www.jstor.org/stable/2984418.

Cohen, Z.D., DeRubeis, R.J., 2018. Treatment selection in depression. Annu Rev. Clin. Psychol. 14, 209–236. https://doi.org/10.1146/annurev-clinpsy-050817-084746.

Cohen, Z.D., Kim, T.T., Van, H.L., Dekker, J.J., 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. Psychother. Res. 30 (2), 137–150.

Cuijpers, P., Quero, S., Noma, H., Ciharova, M., Miguel, C., Karyotaki, E., Cipriani, A., Cristea, I.A., Furukawa, T.A., 2021. Psychotherapies for depression: a network meta-analysis covering efficacy, acceptability and long-term outcomes of all main treatment types. World Psychiatry 20 (2), 283–293. https://doi.org/10.1002/wps.20860.

Delgadillo, J., McMillan, D., Lucock, M., Leach, C., Ali, S., Gilbody, S., 2014. Early changes, attrition, and dose-response in low intensity psychological interventions. Br. J. Clin. Psychol. 53 (1), 114–130. https://doi.org/10.1111/bjc.12031.

DeRubeis, R.J., Cohen, Z.D., Forand, N.R., Fournier, J.C., Gelfand, L.A., Lorenzo-Luaces, L., 2014. 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.

Driessen, E., Smits, N., Dekker, J.J., Peen, J., Don, F.J., Kool, S., Westra, D., Hendriksen, M., Cuijpers, P., Van, H.L., 2016. Differential efficacy of cognitive behavioral therapy and psychodynamic therapy for major depression: a study of prescriptive factors. Psychol. Med. 46 (4), 731–744.

Fiedler, K., 2011. Voodoo correlations are everywhere-not only in neuroscience. Perspect. Psychol. Sci. 6 (2), 163–171. https://doi.org/10.1177/1745691611400237.

Friedman, J., Hastie, T., Tibshirani, R., 2010. Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. 33 (1), 1. https://doi.org/10.18637/jss.v033.i01.

Huibers, M.J., Cohen, Z.D., Lemmens, L.H., Arntz, A., Peeters, F.P., Cuijpers, P., DeRubeis, R.J., 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.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2003. An Introduction to Statistical Learning with Applications in R. Springer.

Jonghe de, F., 2014. Handboek Kortdurende Psychoanalytische Steungevende Psychotherapie. Amsterdam: Benecke Beheer B.V.

Kaiser, T., Volkmann, C., Volkmann, A., Karyotaki, E., Cuijpers, P., Brakemeier, E.-L., 2022. Heterogeneity of treatment effects in trials on psychotherapy of depression. Clin. Psychol. 29 (3), 294–303. https://doi.org/10.1037/cps0000079.

Kessler, van Loo, H., Wardenaar, K., Bossarte, R., Brenner, L., Ebert, D., Sampson, N., 2017. Using patient self-reports to study heterogeneity of treatment effects in major depressive disorder. Epidemiol. Psychiatr. Sci. 26 (1), 22–36.

Kravitz, R.L., Duan, N., Braslow, J., 2004. Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. Milbank. Q. 82 (4), 661–687. https://doi.org/10.1111/j.0887-378X.2004.00327.x.

Kuhn, M., 2008. Building predictive models in R using the caret package. J. Stat. Softw. 28 (5), 1–26. https://doi.org/10.18637/jss.v028.i05.

Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling, 26. Springer.

Lorenzo-Luaces, L., Peipert, A., Romero, R.D.J., Rutter, L.A., Rodriguez-Quintana, N., 2021. Personalized medicine and cognitive behavioural therapies for depression: small effects, big problems, and bigger data. Int. J. Cogn. Ther. 14 (1), 59–85.

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. Clin. Psychol. Sci. 7 (3), 445–461. https://doi.org/10.1177/2167702618815466.

Miggiels, M., Ten Klooster, P., Beekman, A., Bremer, S., Dekker, J., Janssen, C., van Dijk, M.K., 2025. The D*phase-study: comparing short-term psychodynamic psychotherapy and cognitive behavioural therapy for major depressive disorder in a randomised controlled non-inferiority trial. J. Affect. Disord. 371, 344–351. https://doi.org/10.1016/j.jad.2024.10.122.

Miggiels, M.F., Ten Klooster, P.M., Bremer-Hoeve, S., Dekker, J.J.M., Huibers, M.J.H., Reefhuis, E., Van, H.L., van Dijk, M.K., 2021. The D*phase-study: study protocol for a pragmatic two-phased, randomised controlled (non-inferiority) trial that addresses treatment non-response and compares cognitive behavioural therapy and short-term psychodynamic supportive psychotherapy for major depression. BMC. Psychiatry 21 (1), 233. https://doi.org/10.1186/s12888-021-03156-8.

Robinson, L., Delgadillo, J., Kellett, S., 2020. The dose-response effect in routinely delivered psychological therapies: a systematic review. Psychother. Res. 30 (1), 79–96. https://doi.org/10.1080/10503307.2019.15666.

RStudio Team, 2023. RStudio: integrated Development for R. RStudio, PBC, Boston, MA URL. http://www.rstudio.com/.

Rush, A.J., Carmody, T., Reimitz, P.E., 2000. The Inventory of Depressive Symptomatology (IDS): clinician (IDS-C) and self-report (IDS-SR) ratings of depressive symptoms. Int. J. Methods Psychiatr. Res. 9 (2), 45–59. https://doi.org/10.1002/mpr.79.

Rush, A.J., Giles, D.E., Schlesser, M.A., Fulton, C.L., Weissenburger, J., Burns, C., 1986. The inventory for depressive symptomatology (IDS): preliminary findings. Psychiatry Res. 18 (1), 65–87. https://doi.org/10.1016/0165-1781(86)90060-0.

Rush, A.J., Gullion, C.M., Basco, M.R., Jarrett, R.B., Trivedi, M.H., 1996. The inventory of depressive symptomatology (IDS): psychometric properties. Psychol. Med. 26 (3), 477–486. https://doi.org/10.1017/s0033291700035558.

Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-International

Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J. Clin. Psychiatry 59 (Suppl 20), 22–57.

Simon, G.E., Perlis, R.H., 2010. Personalized medicine for depression: can we match patients with treatments? Am. J. Psychiatry 167 (12), 1445–1455. https://doi.org/10.1176/appi.ajp.2010.09111680.

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.

Trivedi, M.H., Rush, A.J., Ibrahim, H.M., Carmody, T.J., Biggs, M.M., Suppes, T., Crismon, M.L., Shores-Wilson, K., Toprac, M.G., Dennehy, E.B., Witte, B., Kashner, T. M., 2004. The Inventory of depressive symptomatology, clinician rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of depressive symptomatology, clinician rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. Psychol. Med. 34 (1), 73–82. https://doi.org/10.1017/s0033291703001107.

van Bronswijk, S.C., Bruijniks, S.J.E., Lorenzo-Luaces, L., Derubeis, R.J., Lemmens, L.H.J. M., Peeters, F.P.M.L., Huibers, M.J.H., 2021a. Cross-trial prediction in psychotherapy: external validation of the Personalized Advantage Index using machine learning in two Dutch randomized trials comparing CBT versus IPT for depression. Psychother. Res. 31 (1), 78–91. https://doi.org/10.1080/10503307.2020.1823029.

van Bronswijk, S.C., DeRubeis, R.J., Lemmens, L.H.J.M., Peeters, F.P.M.L., Keefe, J.R., Cohen, Z.D., Huibers, M.J.H., 2021b. Precision medicine for long-term depression outcomes using the Personalized Advantage Index approach: cognitive therapy or interpersonal psychotherapy? Psychol. Med. 51 (2), 279–289.

Vittengl, J.R., Anna Clark, L., 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 Res. 253, 174–181.

Vul, E., Harris, C., Winkielman, P., Pashler, H., 2009. Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. Perspect. Psychol. Sci. 4 (3), 274–290. https://doi.org/10.1111/j.1745-6924.2009.01125.x.