Article

Using the Personalized Advantage Index for Individual Treatment Allocation to Blended Treatment or Treatment as Usual for Depression in Secondary Care

Nadine Friedl [1,*], Tobias Krieger [1], Karine Chevreul [2], Jean Baptiste Hazo [3], Jérôme Holtzmann [4], Mark Hoogendoorn [5], Annet Kleiboer [6], Kim Mathiasen [7,8], Antoine Urech [9], Heleen Riper [10,11,12] and Thomas Berger [1]

[1] Department of Clinical Psychology, University of Bern, 3012 Bern, Switzerland
[2] URC Eco Ile-de-France (AP-HP), Hotel Dieu, 1, Place du Parvis Notre Dame, 75004 Paris, France
[3] Eceve, Unit 1123, Inserm, University of Paris, Health Economics Research Unit, Assistance Publique-Hôpitaux de Paris, 75004 Paris, France
[4] University Hospital Grenoble Alpes, Mood disorders and emotional pathologies Unit; Pôle de Psychiatrie, Neurologie et Rééducation Neurologique, 38043 Grenoble, France
[5] Department of Computer Science, VU University Amsterdam Faculty of Sciences, De Boelelaan 1081m 1081 HV Amsterdam, The Netherlands
[6] Section Clinical Psychology, Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam and EMGO+ Institute for Health Care and Research, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands
[7] Department of Psychology, Faculty of Health Sciences, University of Southern Denmark, 5230 Odense M, Denmark
[8] Center of Telepsychiatry, University of Southern Denmark, 5000 Odense, Denmark
[9] INSELSPITAL, University Hospital Bern, University Clinic for Neurology, University Acute-Neurorehabilitation Center; 3010 Bern, Switzerland
[10] Department of Psychiatry and the Amsterdam Public Health Research Institute, GGZ inGeest/Amsterdam UMC, Vrije Universiteit, de Boelelaan 1117, 1081 HV Amsterdam, The Netherlands
[11] Department of Research and Innovation, GGZ inGeest Specialized Mental Health Care, Oldenaller 1, 1081 HJ Amsterdam, The Netherlands
[12] Department of Clinical, Neuro-and Developmental Psychology and the Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands
* Correspondence: nadine.friedl@psy.unibe.ch

Abstract: A variety of effective psychotherapies for depression are available, but patients who suffer from depression vary in their treatment response. Combining face-to-face therapies with internet-based elements in the sense of blended treatment is a new approach to treatment for depression. The goal of this study was to answer the following research questions: (1) What are the most important predictors determining optimal treatment allocation to treatment as usual or blended treatment? and (2) Would model-determined treatment allocation using this predictive information and the personalized advantage index (PAI)-approach result in better treatment outcomes? Bayesian model averaging (BMA) was applied to the data of a randomized controlled trial (RCT) comparing the efficacy of treatment as usual and blended treatment in depressive outpatients. Pre-treatment symptomatology and treatment expectancy predicted outcomes irrespective of treatment condition, whereas different prescriptive predictors were found. A PAI of 2.33 PHQ-9 points was found, meaning that patients who would have received the treatment that is optimal for them would

have had a post-treatment PHQ-9 score that is two points lower than if they had received the treatment that is suboptimal for them. For 29% of the sample, the PAI was five or greater, which means that a substantial difference between the two treatments was predicted. The use of the PAI approach for clinical practice must be further confirmed in prospective research; the current study supports the identification of specific interventions favorable for specific patients.

1. Introduction

Globally, 300 million people of all ages suffer from depression [1]. Depression is one of the most common problems seen in clinical practice, and it is associated with high societal costs, as well as great suffering [2]. Given this burden, the need for improved access to efficacious and cost-effective treatments is essential [3]. In the last decades, research has focused on examining different treatment options for depression. Especially, cognitive behavior therapy (CBT) and interpersonal therapy (IPT) can be seen as first-line treatments [4,5]. Moreover, current studies are aimed at scaling up treatments for depression. One way to do so is through internet-based therapies [5]. Whereas the most dominant format in which treatment is delivered is through face-to-face contact, internet-based therapies have received much attention in recent years [6]. The efficacy and cost-effectiveness of the latter have been supported by a growing number of research [7–9]. Even though only a few studies have directly compared internet-based with face-to-face CBT for depression, results suggest it to have similar overall effects [6].

A newer approach to depression treatment is to combine web-based technologies with face-to-face therapy, called blended treatment. Blended treatment includes any combination of face-to-face therapy and internet-based interventions, e.g., web-based components are used as an adjunctive intervention or are integrated during face-to-face therapy [10]. Although research that investigates the efficacy of blended treatment formats is still scarce, preliminary results suggest their feasibility and their efficacy in reducing symptoms of depression [11–15]. For example, a randomized controlled trial by Berger and colleagues [16] showed the superiority of blended treatment, consisting of an internet-based intervention as an adjunct to face-to-face psychotherapy, in comparison to regular face-to-face psychotherapy in a pragmatic randomized controlled study in patients with a unipolar affective disorder in routine care. Another recent study showed the noninferiority of blended treatment to conventional CBT for patients with depression and found blended treatment to be cost-effective [17]. Moreover, the blended treatment has also been evaluated in an inpatient setting where patients suffering from depression that received an online self-help program in addition to inpatient psychotherapy improved significantly more than patients who received online information about depression in addition to inpatient psychotherapy [18]. Furthermore, a recent systematic review showed that, compared to face-to-face therapy, a blended treatment may help maintaining initially achieved changes within psychotherapy in the long-term [19].

Potential benefits of blending treatments may be a greater reduction in depressive symptoms and increased cost-effectiveness [20], as well as an improvement of patients' adherence to the treatment program [21]. Furthermore, an asset of a blended treatment may be that it combines the advantages of both treatment forms [3,14,22,23]. For example, the face-to-face contact enables clinicians to individualize or tailor the treatment and to react in crisis situations, while providing online modules between sessions could promote patient engagement and enhance the translation of treatment into daily life (e.g., [24]). On the other hand, when online components are not used by the patients in blended treatments, reductions in the number of face-to-face sessions may lead to worse treatment outcomes [25]. Furthermore, therapists may raise concerns of overburdening depressed patients [18]. So far, it is not clear for which patients blended treatment may be a feasible option and for which patients a conventional treatment should be favored.

Patients with depression may differ substantially from each other, and evidence suggests that the diagnostic categories leave room for great diversity [26,27]. This results in differences with regard to patients' illness courses and individual treatment responses [28]. Research suggests that individual patient characteristics may moderate the efficacy of different treatments at an individual level [29]. It is, therefore, important to recognize that no single treatment is likely to be the best for everyone, even though, on a group level, it is efficacious for patients suffering from depression [26,30]. This is why more and more researchers move away from investigating treatment efficacy on a group level and instead focus on custom-tailoring the treatment to the individual patient [30,31]. In this sense, it may be a solution to increase the overall treatment response rates [32].

Precision medicine tries to tailor treatments to the specific needs of the patient [33]. More recently, in clinical psychology and psychotherapy, algorithms are used that predict from which treatment a patient benefits the most [34]. As an example, Becker and colleagues [35] introduce a conceptual framework that helps classifying applications of predictive modeling in mental health research. These authors try to bridge the gap between psychologists and predictive modelers with providing a common language for classifying predictive modeling mental health research. They suggest that e-mental health researchers should focus more on the validity of model predictions instead of solely focusing on identifying predictors. Another example is the personalized advantage index PAI) approach, which identifies patients with a certain disorder (e.g., major depression) who benefit more from one treatment than another [30]. Using the personalized advantage index (PAI), it is possible to identify the treatment that predicts a better treatment outcome for a given patient if there are two comparably effective treatments to choose from [36]. The PAI estimates how much a specific treatment is better for an individual patient than another, and its feasibility and relevance have been shown in several studies on the treatment of depression [36–40]. Baseline patient characteristics can be divided into two types of predictors: a prognostic variable predicts treatment outcome irrespective of treatment condition, whereas a prescriptive variable predicts a differential treatment response to two or more treatment modalities [29,30]. Up to today, different treatments have been compared using the PAI, and its values range from 1.4 when comparing CBT to CBT with integrated exposure and emotion-focused elements [38] up to 8.9 when comparing cognitive therapy to IPT [40]. Higher absolute values of the PAI stand for stronger predicted benefits of one treatment over another. Being able to identify the best treatment for an individual with depression is essential, because it may make health care delivery more efficient [32]. Even though predictive modeling is still very young in the field of e-mental health [41], Bremer and colleagues [42] were able to predict clinical outcomes and costs of patients with depression prior to starting blended psychotherapy in a subsample of the current study using machine learning techniques.

In the current study, treatment as usual (TAU), i.e., regular face-to-face psychotherapy, was compared to blended treatment for patients with major depressive disorder (MDD) in secondary care. The present study set out to answer the following research questions: (1) What are the most important predictors determining optimal treatment allocation to TAU or blended treatment? and (2) Would model-determined treatment allocation using this predictive information and the PAI-approach result in better treatment outcomes? To the best of our knowledge, this is the first study comparing different treatment delivering formats, i.e., traditional face-to-face CBT versus blended CBT, by using the PAI-approach.

2. Materials and Methods

Data used in the present study was drawn from the European project "European COMPARative effectiveness research on blended depression treatment" (E-COMPARED, February 2018) [43]. The E-COMPARED project included a randomized, controlled, noninferiority trial that examined the clinical and cost-effectiveness of blended treatment compared to treatment as usual in routine care in nine European countries. Adult patients diagnosed with MDD were recruited in primary or in specialized mental health care. The current study uses the data of the four countries that recruited patients in specialized mental health care (France, the Netherlands, Switzerland, and Denmark). The following inclusion criteria were met by participants: (1) being 18 years of age or older, (2) meeting

Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) diagnostic criteria for MDD as confirmed by a telephone-administered MINI international neuropsychiatric interview (M.I.N.I.) version 5.0 [44], and (3) minimal to severe symptoms of depression based on a score of 5 or above on the patient health questionnaire-9 (PHQ-9) [45]. Exclusion criteria for participating in the study were: (1) high risk for suicide according to the M.I.N.I.; (2) psychiatric comorbidity such as substance dependence, bipolar affective disorder, psychotic illness, or obsessive compulsive disorder; (3) currently receiving another psychological treatment for depression; (4) being unable to comprehend the spoken and written language of the country where the study is conducted; (5) not having access to a computer with fast internet connection; and (6) not having a smartphone that is compatible with the mobile component of the intervention that is offered. Patients were randomized to blended treatment or TAU using an allocation scheme with a computerized random number generator at an allocation ratio of 1:1 and between 8 and 14 allocations per block. Details about the study design are described elsewhere [43]. For Switzerland, this study was approved by the cantonal ethics committees Bern and Zurich (registration number: 001/15; date: 18 March 2015); for Denmark, the study was approved by the Ethics Committee of the Region of Southern Denmark (registration number S-20150150; date: 18 November 2015); for France, the study was approved by the "Comité de protection des personnes", Ile de France V (registration number 15033-$n°$ 2015-A00565-44; date: 2 June 2015); and, for the Netherlands, the study was approved by the METC VUMC (registration number 2015.078; date: 8 May 2015). Furthermore, all participants provided written informed consent and gave permission to all E-COMPARED partners to use their anonymized data.

2.1. Sample

The current study has a sample size of *n* = 251. The sample consists of 83 participants from the Netherlands (33.9%), 79 participants from France (32.2%), 44 from Switzerland (18.0%), and 39 participants from Denmark (15.9%). The mean age at baseline was 41.0 years (SD = 13.7), and 68.2% of the participants were female. The majority were either single (33.5%) or married (31.8%), and 21.2% of participants were living together and 12.8% were divorced. In the TAU condition, 57.9% of the sample suffered from a recurrent depression, 45.2% from a current melancholic depressive episode, 7.9% from a comorbid dysthymia, and 46.8% from a comorbid anxiety disorder. The number of patients are taking antidepressant medication at the time of the baseline measurement was 53.2%. In the blended treatment condition, 53.8% of the participants suffered from a recurrent depression, 34.5% from a current melancholic depressive episode, 5.9% from a comorbid dysthymia, and 61.3% had a comorbid anxiety disorder. Half of the participants (49.6%) were taking antidepressant medication at baseline.

2.2. Interventions

Individual face-to-face CBT was combined with internet-based CBT elements delivered through a platform for blended treatment. Three different online platforms were used across the participating countries. Switzerland used "Deprexis" [46], whereas Denmark used NoDep [47] and France and the Netherlands used the platform "Moodbuster" [3]. The most important components of the treatment were cognitive restructuring, behavioral activation, psychoeducation, and relapse prevention, which were delivered over 11–20 sessions. In the blended treatment, a smaller number of face-to-face sessions is offered, and some sessions are replaced by online modules. Treatment was provided by CBT therapists who received special training on how to deliver the blended treatment. In Switzerland and the Netherlands, the therapists were either licensed CBT therapists or CBT therapists who were supervised by an experienced licensed CBT therapist. In Denmark, the treatment was delivered by either licensed psychologists or psychologists under supervision of licensed psychologists. In France, the blended treatment was provided by licensed psychotherapists [43].

The TAU treatment was defined as the routine care that patients received in specialized mental health care when they were diagnosed with depression. In practice, this meant that the TAU group received regular fact-to-face CBT.

2.3. Measures

2.3.1. Primary Outcome

The primary outcome measure for this study was the PHQ-9 [45] assessed after 12 weeks. The PHQ-9 consists of nine questions which are based upon the DSM-IV criteria for the diagnosis of depressive disorders. It is used as a diagnostic instrument and as a severity measure for depression. A 5-point difference in PHQ-9 scores is seen as clinically significant [48]. The validity and sensitivity to change of the PHQ-9 were satisfactory in previous studies [49,50]. Cronbach's alpha in the present study was 0.78.

2.3.2. Predictor Variables

We used an exploratory approach and included a total of 28 potential predictors measured at baseline in the analysis. All variables of the baseline assessment that did not exceed a number of acceptable missing values (<50%) were included. We classified the variables into four categories: (1) sociodemographic variables, (2) symptomatology and quality of life, (3) healthcare utilization, and (4) patient expectancy.

The sociodemographic variables included age, marital status, education, gender, and country and were assessed with single item questions. Variables related to symptomatology and treatment history were recurrent depression, therapy preference, dysthymia, melancholic depressive episodes, comorbid anxiety, and current use of antidepressants. Quality of life was measured with the EQ-5D [51]. Healthcare utilization was assessed with the TiC-P [52]. The TiC-P examines healthcare consumption and productivity losses as a consequence of a mental disorder via a self-report questionnaire. The questions include contacts within the healthcare sector and the use of medication. All the questions aim at the period of the last four months before the start of the treatment (see Table 1 for the items of the TiC-P that have been included). Patient expectancy was measured by the credibility and expectancy questionnaire (CEQ; [53]).

Table 1. TiC-P items included in the analysis.

"How many times did you consult a general practitioner?" "How many times did you consult a psychologist?" "How many times did you consult a psychotherapist?", "How many times did you consult a psychiatrist?", and "How many times did you consult a professional from an ambulatory mental health institution?" "How many times did you consult a professional from a clinic for alcohol or drugs?" and "How many times did you consult self-help groups?"	"How many days did you spend in a day-time treatment program in a regular hospital?" "How many days did you spend in a day-time treatment program in a psychiatric hospital?" and "How many days did you use outpatient psychotherapeutic services in addition to your psychotherapy?"	"How many admissions to a regular hospital did you have?" and "How many admissions to a psychiatric hospital did you have?" "How many admissions to a rehabilitation clinic did you have?" "Do you have a paid job?" "Did health problems oblige you to call in sick from work at any time?"

2.4. Data Analytical Strategy

Regarding the predictor variables, a bottom-up approach was followed, which means that even though some variables might have a particular relevance to one treatment or the other, the predictors are treated equally in the data analysis.

2.4.1. Missing Data

In line with previous research using the PAI approach, we included those participants for which PHQ-9 scores after 12 weeks were available [36,38,40]. This left us with *n* = 245, representing 97.6% of the total sample. Distributed over the two conditions, there were 126 patients in the TAU condition

and 119 patients in the blended treatment condition. With regard to the baseline measures, missing values were found in the dimensions credibility and expectancy of the CEQ (3.7% respectively 4.1%), the EQ-5D (1.6%), antidepressants (1.2%), prior psychotherapy (40.4%), comorbid melancholic episodes (14.3%), comorbid dysthymia (5.3%), comorbid anxiety (2.4%), and some items of the TiC-P (2.0–40.8%). We imputed these missings in the baseline measures with the R *missForest* [54]. Here, missing values are predicted on the basis of a random forest approach, trained on observed values of the available data. An advantage of *missForest* is that it imputes categorical and continuous variables simultaneously [55]. It has been shown to be highly accurate and to outperform other imputation methods due to its small imputation error [56]. The imputed data set is the basis for all analyses that follow.

2.4.2. Bayesian Model Averaging (BMA)

There are different data analysis approaches that can be used to identify baseline variables that predict outcome in one treatment versus another. Data analysis approaches that rely on one model only can only include a limited number of baseline measures that could be predictors. In that case, the most common rule that is used states that at least 10 observations per predictor are necessary to not exceed a level of bias which is acceptable [57,58]. If we had followed this approach, we could have only included a small number of predictors. Furthermore, the problem with relying upon a single selected model is that it may result in overconfidence in the conclusions drawn regarding quantified associations. The problem is that there may be alternative models that have different subsets of predictors that fit the data just as well as the one selected model [59,60]. Overfitted models cause uncertainty regarding the actual value of findings, because they may not replicate in future samples [61]. Bayesian model averaging (BMA) is a method that can account for model uncertainty while providing a better predictive ability. The BMA method has two advantages. First, it results in predictions that are less risky and. second, BMA provides simpler model choice criteria, because it uses the Bayesian inference for model prediction and selection [62]. With BMA, a posterior probability is estimated on the basis that each considered model is the correct one. This included the aforementioned model uncertainty in the estimates for the parameter and inferences. That means that BMA averages over all possible predictor sets and delivers model choice criteria that help to identify the most probable model. With using BMA, sharper predictions can be derived from the data, especially in cases with many possible predictors but a limited sample size. Several studies have supported BMA's predictive performance [60,63–66].

The data frame was divided into two subsets: the TAU condition and the blended treatment condition. Using the R package *BAS*, for each condition, a separate linear regression model was computed [67]. Then, Bayesian adaptive sampling (BAS) without replacement for variable selection in linear models using the function *bas.lm* with treatment outcome was applied (PHQ-9 score after 12 weeks) as the dependent variable. The relative importance of each variable was evaluated using the posterior probabilities that were calculated for each potential predictor. The marginal posterior inclusion probabilities functioned as the criteria for determining the importance of the potential predictors. Values above 0.5 point out that the predictor has been incorporated in more than half the models; thus, in the present study, in over 15,000 models. The nominal variables were split into their different categories, which enabled a precise interpretation of the results. The model performance was further evaluated using posterior probabilities. The appropriate *bas.lm* function was chosen based on the following considerations: a Laplace approximation to the Jeffreys-Zellner-Siow (JZS) prior for the integration of alpha = 1 was used as the criterion for the priors, which is called the "ZS-null". The JZS prior uses the Zellner-Siow Cauchy prior on the coefficients and the Jeffreys prior on sigma. The squared scale of the prior, where the default is alpha = 1, can be controlled using the optional parameter "alpha" [38,67]. Marginal inclusion probabilities were calculated with the "MCMC+BAS" method, which runs an initial Markov chain Monte Carlo (MCMC) algorithm and then samples without replacement, as in BAS. Compared to the BAS alone, the "MCMC+BAS" method is the preferred option, because it provides estimates with low bias [68]. The number of models was set to 30,000 assuming that each additional model would add only a small increment to

the cumulative probability, i.e., not leading to essential differences in posterior distributions. Due to the limited sample size, the models have been built and tested in the same dataset.

2.4.3. Personalized Advantage Index (PAI)

When predicting the therapy outcome for each patient, applying a leave-one-out approach ("jackknife") to estimate regression models is an essential beginning step in generating the PAI [36,38,69]. In this procedure, overfitting can be avoided by excluding each target patient for whom the PAI prediction is estimated from the model. For each patient, two regression models were built using the treatment-specific predictors identified with BMA. For each patient, a factual prediction (PHQ-9 score at 12 weeks of the treatment the patient has received) and a counterfactual prediction (PHQ-9 score at 12 weeks of the intervention the patient did not receive) were estimated. In the next step, those two predictions were compared, and the prediction that resulted in the best outcome for the patient was defined to be the optimal treatment for that patient. When comparing the observed change scores, the size of the predicted difference of receiving the treatment with the greater predicted benefit is ultimately the PAI [36]. The higher the absolute values of the PAI, the stronger is the predicted benefit of one treatment over another. The interpretation of the PAI can be demonstrated with a recent study that used the PHQ-9 as the primary outcome and found a PAI of 2.5 [70]. This means that if patients had received their "optimal" treatment (out of the two), their PHQ-9 score at 12 weeks would have been 2.5 points lower than if they had obtained their nonoptimal treatment.

3. Results

Henceforth, we firstly report the five best models for each treatment condition, and, secondly, we report the PAI results. The best models are defined based on the highest posterior probability and the lowest Bayesian information criterion (BIC).

3.1. Variables Predicting Outcome in TAU

The five best models predicting depression severity at 12 weeks in the TAU condition are displayed in Table 2. The Bayes factor, number of predictors, R^2, log marginal likelihood, and the posterior probabilities are provided for each model. Model 1 has the largest Bayes factor and the largest posterior probability (0.02) and, thus, seems to fit the data best. As a result, Model 1 was selected as our final predictive model of the PHQ-9 score at 12 weeks in the TAU condition.

Table 2. Five best models for treatment as usual (TAU).

Fit Indices	Model 1	Model 2	Model 3	Model 4	Model 5
Bayes Factor	1	0.934	0.597	0.572	0.510
Number of Variables	6	6	6	5	6
R^2	0.428	0.445	0.441	0.440	0.439
Log Marginal Likelihood	22.729	22.659	22.213	22.170	22.056
Posterior Probabilities	0.021	0.019	0.012	0.012	0.011

While the selected model includes six variables in total, the strongest predictors of the PHQ-9 score at 12 weeks in the TAU condition included the pretreatment PHQ-9 score (Prob = 100%), CEQ expectancy (Prob = 97%), "How many days did you use outpatient psychotherapeutic services in addition to your psychotherapy?" (Prob = 95%), "How many times did you consult a psychiatrist?" (Prob = 64%), Denmark (Prob = 58%), and "How many days did you attend a day-time treatment program in a psychiatric hospital?" (Prob = 51%). A higher pretreatment score, more consultations with a psychiatrist, and more days in a day-time treatment program in a psychiatric hospital prior to treatment predicted a higher PHQ-9 score at 12 weeks. Higher expectancy scores, receiving TAU in Denmark, and more days using outpatient psychotherapeutic services in addition to the psychotherapy prior to treatment predicted lower PHQ-9 scores at 12 weeks. The effects of other

variables appeared minimal due to their small posterior probabilities. See Appendix 1 in the Supplemental Online Material for the complete list of variables and their inclusion probabilities.

3.2. Variables Predicting Outcome in the Nlended Treatment

Table 3 gives an overview of the five best models to predict treatment outcome in the blended treatment condition. Based on the posterior probabilities and the Bayes factor, Model 1 was rated the best model. Thus, Model 1 was selected as the final predictive model for the blended treatment condition.

Table 3. Five best models for the blended treatment condition.

Fit Indices	Model 1	Model 2	Model 3	Model 4	Model 5
Bayes Factor	1	0.887	0.577	0.525	0.516
Number of Variables	6	5	4	6	6
R^2	0.398	0.416	0.392	0.429	0.446
Log Marginal Likelihood	19.852	19.731	19.301	19.207	19.190
Posterior Probabilities	0.012	0.011	0.007	0.006	0.006

Based on the posterior probabilities, the most important predictors for treatment outcome in the blended treatment were the pretreatment PHQ-9 score (Prob = 99.9%), regular hospital admissions (Prob = 99.9%), EQ-5D quality of life (Prob = 74.6%), CEQ expectancy (Prob = 72.3%), consulting self-help groups (Prob = 70.0%), and being widowed (Prob = 49.7%). CEQ credibility reached a posterior probability of 42.9%. A higher pretreatment PHQ-9 score, being widowed, more hospital admissions, and consulting self-help groups predicted higher PHQ-9 scores at 12 weeks. A higher expectancy for improvement and a higher quality of life predicted lower PHQ-9 scores after 12 weeks. See Appendix 2 in the Supplemental Online Material for the posterior probabilities of all variables measured at baseline.

3.3. Personalized Advantage Index

Using the treatment specific predictors described above, the prediction of a patient's PHQ-9 score after 12 weeks was computed separately for each treatment condition. The true error of the PHQ-9 score predictions at 12 weeks was 4.16, representing the average absolute difference between the predicted and actual, observed scores across all patients. Patients who were categorized as having received their optimal treatment had a mean PHQ-9 score of 9.67 (n = 124) at 12 weeks, whereas patients who were classified as having received their suboptimal treatment had a mean PHQ-9 score of 12.00 (n = 121). Figure 1 shows the frequency of predicted PHQ-9 scores at 12 weeks for every patient in both the optimal and suboptimal treatments.

In the first step, an individual PAI was calculated for each patient. Secondly, the average PAI was calculated as the mean difference in PHQ-9 scores between the optimal and the suboptimal treatments for each patient. The average PAI of the current study was 2.33. The PAI can be read as follows: if patients had received the treatment that is "optimal" for them, their PHQ-9 score at 12 weeks would have been 2.33 points lower than if they had received the treatment that is suboptimal for them. In Figure 2, the frequencies of the individual PAIs are shown. A PAI that is five or greater would mean that a substantial difference was predicted between the two treatments, because 5 points on the PHQ-9 stands for a minimal clinically meaningful difference for individual change [71]. This was the case for 29% of the patients in this sample.
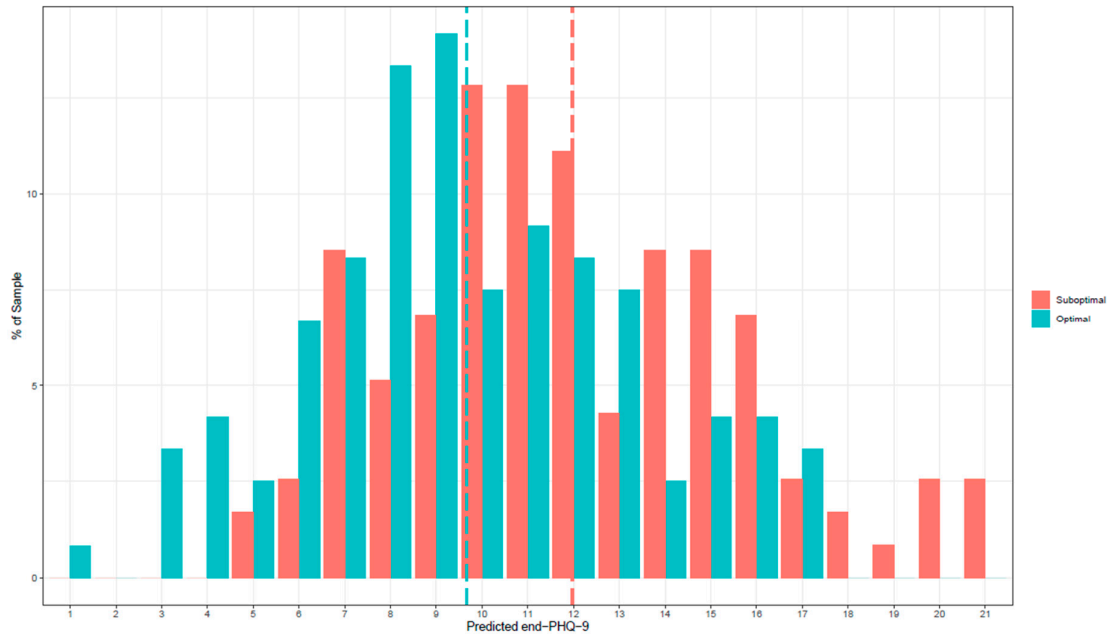
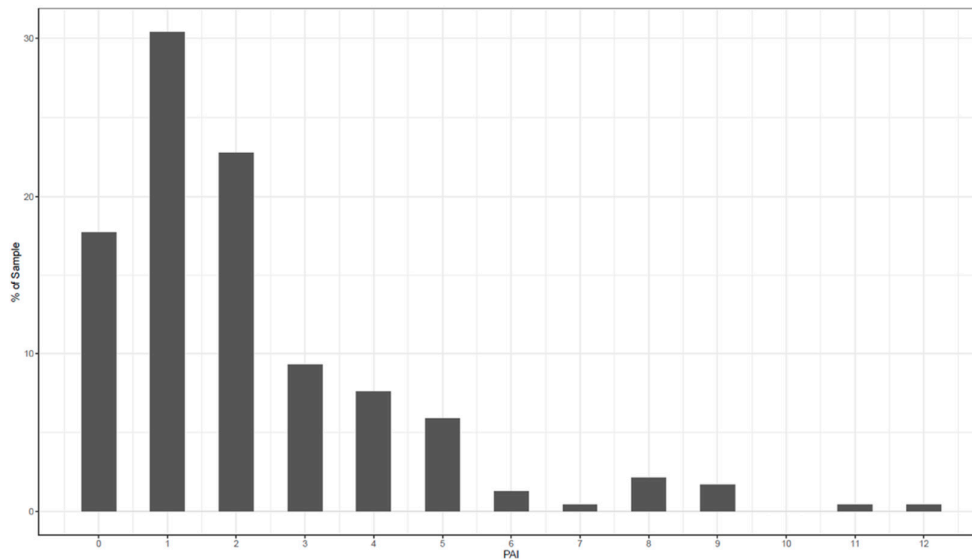**Figure 1.** Frequency of predicted PHQ-9 scores at 12 weeks.

Figure 2. Frequencies of individual personalized advantage indexes (PAIs).

4. Discussion

Regarding the predictors of treatment outcome at 12 weeks in each of the interventions, different relevant predictors were identified for TAU and the blended treatment, respectively. In the TAU condition, a lower pretreatment PHQ-9 score, less consultations with a psychiatrist and less days in a day-time treatment program in a psychiatric hospital, higher expectancy, receiving TAU in Denmark, and more days using outpatient psychotherapeutic services in the four months prior to the study predicted a better treatment outcome, i.e., a lower PHQ-9 score (at 12 weeks). In contrast, in the blended treatment condition, a lower pretreatment PHQ-9 score, not being widowed, less hospital admissions and consulting self-help, a higher expectancy for improvement, and a higher EQ-5D score predicted lower scores at 12 weeks; thus, a better treatment outcome.

To offer an initial interpretation of our findings, the distinction between prescriptive and prognostic predictors is used. Prognostic variables predict treatment outcomes regardless of

treatment conditions [30,38]. In contrast, prescriptive variables may support differential indications by predicting whether a patient will benefit more from one treatment in comparison to another. In the present study, the pretreatment depressive symptomatology and treatment expectancy are the only prognostic predictors, i.e., the only variables that predict treatment outcome in both conditions. This is in line with previous research that has found pretreatment symptomatology and expectancy to be important predictors of treatment outcome, in the sense that higher symptomatology before treatment predicts worse end-state symptomatology [30,34,36,38,72,73] and higher expectancy for improvement predicts better treatment outcome [74,75]. Interestingly, for internet-based treatments, higher baseline symptomatology is not necessarily a negative predictor of treatment outcome. More often the opposite is found, i.e., that higher depressive symptomatology pretreatment predicts better treatment outcomes [76–78]. This might be partly explained by the efficacy nature of previous randomized controlled trials (RCTs) in comparison to the routine care and effectiveness nature of the current study.

Regarding the prescriptive predictors, our findings partly corroborate findings from previous studies predicting treatment outcomes for patients with depression. With regard to prescriptive predictors of the blended treatment condition, a lower quality of life and being widowed predicted worse treatment outcomes. In contrast to the present result, the study by Huibers and colleagues [40] found a higher quality of life to be a prognostic predictor, i.e., to predict favorable outcomes irrespective of treatment conditions.

For the TAU condition, more consultations with a psychiatrist and more days in a day-time treatment program in a psychiatric hospital predicted worse treatment outcomes. This could mean that patients' symptomatology and patients' general functioning is too severe to be able to profit from TAU. Furthermore, to our knowledge, there are no international studies regarding psychotherapy of depression, indicating country as a relevant predictor of outcome.

Healthcare utilization within the four months prior starting treatment was found to be a prescriptive predictor in both conditions. Healthcare uptake may be a proxy for a higher somatic or mental burden and/or may represent a more severe symptomatology of depression. In previous studies, more complex cases (e.g., with chronic symptoms and psychiatric comorbidities) or more severe depressive symptomatology predicted a worse therapy outcome [29,38,79–81]. Interestingly, more hospital admissions only predicted worse treatment outcomes in the blended treatment condition. A possible interpretation of this finding is the fact that the online modules in the blended treatment protocol are highly standardized to target depressive symptomatology. As a result, they may have not sufficiently addressed comorbid symptoms. Somatic comorbidities in patients with depression are not a scarcity and have an influence on individual treatment response and illness course, because they can complicate treatment [81]. Furthermore, the number of hospital admissions may also be a proxy for case complexity and higher mental burden, which in turn may have a negative impact on treatment outcome.

This study's results demonstrate that BMA makes it possible to use a limited set of baseline variables to predict treatment outcome. This is in line with a recent study by Bremer and colleagues [42] that showed the feasibility of providing personalized treatment recommendations at baseline regarding clinical and cost-effectiveness using a subsample of the current study by evaluating various machine learning techniques. Moreover, the present study showed that despite sharing the same diagnosis, patients might benefit more from different treatments. The current study found a PAI of 2.33, indicating that patients could have a PHQ-9 score at 12 weeks that is, on average, more than two points lower if they receive their model-determined optimal treatment in comparison to the suboptimal treatment. This value is in the range of other studies that have found PAIs ranging from 1.35 [38] up to 8.9 [40]. Importantly, for almost one-third (29%) of the patients in the present study, a substantial difference was predicted between the two treatment modalities as the individual PAI was 5 or greater. This result is in line with increasing evidence suggesting that differential treatment responses are not rare and might play an important role for an individual patient and the health care system [30].

The current study has several limitations. First of all, the relatively small sample size did not allow us to build and test the models in separate samples. Using the same sample for testing and building the model might lead to a potential risk of overconfidence. Nevertheless, if studies are designed to develop and validate prescriptive prediction scores that can be tested in future hypothesis-driven confirmatory studies, a smaller sample size might be legitimate [82]. Secondly, the results are based solely on self-reports, and future studies should also include observer ratings. Third, the restricted set of baseline measures is another limitation. Constructs such as personality traits or the familiarity with computers have not been assessed but may have influenced engagement with the online component. Relatedly, people with a low socioeconomic status or senior citizens may not have been well represented in the present study sample. Such groups may not have the opportunity to benefit from the blended treatment, because they may not have access to a computer or a smartphone and/or lack the knowledge to use them. As a consequence, the restricted sample in the present study limits the generalizability of the results. Furthermore, the current study predicts treatment outcome after 12 weeks. Future studies should predict long-term treatment outcome [83]. Finally, we have followed a data-driven approach instead of a theory-driven approach. Although the two methods should be considered complementary [84], a disadvantage of data-driven research is that it is not experiential, and relying on the data alone might not capture the whole picture. Relatedly, the predictors found in the current study need to be validated and replicated in future hypothesis-driven studies.

In spite of the limitations above, the current study is promising to contribute to the further understanding of treatment for depression, because it investigates implications for the use of the blended treatment for patients with depression. Clinical practice should consider factors found to play a role in the treatment and processes of change to provide the optimal treatment for each individual. For example, the quality of life should be taken into account, as these patients may need a more intense treatment protocol that integrates face-to-face interventions with web-based technologies. This interpretation is in line with the notion that more severely depressed patients see the availability of an online program between face-to-face sessions as an advantage of the blended treatment [85]. Furthermore, healthcare utilization should be evaluated prior to treatment selection, because it can give valuable information about the patients' needs, treatment history, and course of illness. In addition, the predictors found to be important in this study and in previous studies could be taken into account to make an informed treatment recommendation in clinical practice. However, future studies with larger samples and more advanced techniques are necessary to validate the current findings, and the identified predictors have to be tested within clinical routine treatment settings. Moreover, prospective studies need to integrate the PAI in the diagnostics process at the beginning of a treatment.

5. Conclusions

To conclude, with the first aim of the study, two prognostic predictors, namely, pretreatment symptomatology and treatment expectancy, were found. Furthermore, several prescriptive predictors were found, predicting the treatment outcome respective of each of the two conditions. Some of our findings are in line with previous research, but other variables, such as baseline healthcare utilization, have not been investigated in this context. The interpretations regarding the prognostic and prescriptive predictors need to be tested empirically, because they are somewhat speculative. Furthermore, this study showed an advantage of model-determined treatment allocation to the TAU or blended treatment, as one-third of the participants had a PAI larger than 5, which means they would have improved significantly if they had received their "optimal" treatment. Although the results need to be validated in future hypothesis-driven studies, the predictors found to be important in the current study should be taken into account to make an informed treatment recommendation in clinical practice.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1: Appendix 1: Bayesian model averaging (BMA) results based on the best 30,000 models in the treatment as usual

(TAU) condition, Appendix 2: BMA results based on the best 30,000 models in the blended treatment condition.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. World Health Organization. Depression. 2018. Available online: https://www.who.int/news-room/fact-sheets/detail/depression (accessed on 21 November 2019).
2. World Health Organization. The Global Burden of Disease: 2004 Update. 2008. Available online: https://www.who.int/healthinfo/global_burden_disease/2004_report_update/en/ (accessed on October 3rd, 2019).
3. Kemmeren, L.; Van Schaik, D.J.; Riper, H.; Kleiboer, A.M.; Bosmans, J.E.; Smit, J.H. Effectiveness of blended depression treatment for adults in specialised mental healthcare: Study protocol for a randomised controlled trial. *BMC Psychiatry* **2016**, *16*, 113.
4. Karasu, T.B.; Gelenberg, A.; Merriam, A.; Wang, P. Treatment of patients with major depressive disorder. In *American Psychiatric Association Practice Guidelines for the Treatment of Psychiatric Disorders Compendium*; 2000; American Psychiatric Publishing, Washington, DC. pp. 413–495.
5. Cuijpers, P. Psychotherapies for adult depression: Recent developments. *Curr. Opin. Psychiatry* **2015**, *28*, 24–29.
6. Carlbring, P.; Andersson, G.; Cuijpers, P.; Riper, H.; Hedman-Lagerlöf, E. Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: An updated systematic review and meta-analysis. *Cognit. Behav. Ther.* **2018**, *47*, 1–18.
7. Arnberg, F.K.;Linton, S.J.; Hultcrantz, M.; Heintz, E.; Jonsson, U. Internet-delivered psychological treatments for mood and anxiety disorders: A systematic review of their efficacy, safety, and cost-effectiveness. *PLoS ONE* **2014**, *9*, e98118.
8. Gerhards, S.; De Graaf, L.; Jacobs, L.; Severens, J.; Huibers, M.; Arntz, A.; Riper, H.; Widdershoven, G.; Metsemakers, J.F.M.; Evers, S.M.A.A. Economic evaluation of online computerised cognitive–behavioural therapy without support for depression in primary care: Randomised trial. *Br. J. Psychiatry* **2010**, *196*, 310–318.
9. Richards, D.; Richardson, T. Computer-based psychological treatments for depression: A systematic review and meta-analysis. *Clin. Psychol. Rev.* **2012**, *32*, 329–342.
10. Krieger, T.; Meyer, B.; Sude, K.; Urech, A.; Maercker, A.; Berger, T. Evaluating an e-mental health program ("deprexis") as adjunctive treatment tool in psychotherapy for depression: Design of a pragmatic randomized controlled trial. *BMC Psychiatry* **2014**, *14*, 285.
11. Hickie, I.B.; Davenport, T.A.; Luscombe, G.M.; Moore, M.; Griffiths, K.M.; Christensen, H. Practitioner-supported delivery of internet-based cognitive behaviour therapy: Evaluation of the feasibility of conducting a cluster randomised trial. *Med. J. Aust.* **2010**, *192*, S31–S35.
12. Høifødt, R.S.; Lillevoll, K.R., Griffiths, K.M.; Wilsgaard, T.; Eisenmann, M.; Waterloo, K.; Kolstrup, N.. The clinical effectiveness of web-based cognitive behavioral therapy with face-to-face therapist support for depressed primary care patients: Randomized controlled trial. *J. Med. Int. Res.* **2013**, *15*, e153.
13. Kessler, D.; Lewis, G.; Kaur, S.; Wiles, N.; King, M.; Weich, S.;Sharp, D.J.; Araya, R.; Hollinghurst, S.; Peters, T.J. Therapist-delivered Internet psychotherapy for depression in primary care: A randomised controlled trial. *Lancet* **2009**, *374*, 628–634.
14. Kooistra, L.C.; Wiersma, J.E.; Ruwaard, J.; van Oppen, P.; Smit, F.; Lokkerbol, J.; Cuijpers, P.; Riper, H. Blended vs. face-to-face cognitive behavioural treatment for major depression in specialized mental health care: Study protocol of a randomized controlled cost-effectiveness trial. *BMC Psychiatry* **2014**, *14*, 290.
15. Robertson, L.; Smith, M.; Castle, D.; Tannenbaum, D. Using the Internet to enhance the treatment of depression. *Aust. Psychiatry* **2006**, *14*, 413–417.

16. Berger, T.; Krieger, T.; Sude, K.; Meyer, B.; Maercker, A. Evaluating an e-mental health program ("deprexis") as adjunctive treatment tool in psychotherapy for depression: Results of a pragmatic randomized controlled trial. *J. Affect. Disord.* **2018**, *227*, 455–462.

17. Thase, M.E.; Wright, J.H.; Eells, T.D.;Barrett, M.S.; Wisniewski, S.R.; Balasubramani, G.K.; McCrone, P.; Brown, G.K. Improving the efficiency of psychotherapy for depression: Computer-assisted versus standard CBT. *Am. J. Psychiatry* **2017**, *175*, 242–250.

18. Zwerenz, R.; Becker, J.; Knickenberg, R.J.; Siepmann, M.; Hagen, K.; Beutel, M.E. Online Self-Help as an Add-On to Inpatient Psychotherapy: Efficacy of a New Blended Treatment Approach. *Psychother. Psychosom.* **2017**, *86*, 341–350.

19. Erbe, D.; Eichert, H.-C.; Riper, H.; Ebert, D.D. Blending face-to-face and internet-based interventions for the treatment of mental disorders in adults: Systematic review. *J. Med. Int. Res.* **2017**, *19*, e306.

20. Wright, J.H.; Wright, A.S.; Albano, A.M.; Basco, M.R.; Goldsmith, L.J.; Raffield, T.; Otto, M.W. Computer-assisted cognitive therapy for depression: Maintaining efficacy while reducing therapist time. *Am. J. Psychiatry* **2005**, *162*, 1158–1164.

21. Clough, B.A.; Casey, L.M. Technological adjuncts to increase adherence to therapy: A review. *Clin. Psychol. Rev.* **2011**, *31*, 697–710.

22. Riper, H.; Van Ballegooijen, W.; Kooistra, L.; De Wit, J.; Donker, T. *Preventie & eMental-Health-Prevention & eMental-Health*; Vrije Universiteit Commissioned by ZonMw: Amsterdam, The Netherlands, 2013.

23. van der Vaart, R.; Witting, M.; Riper, H.; Kooistra, L.; Bohlmeijer, E.T.; van Gemert-Pijnen, L.J. Blending online therapy into regular face-to-face therapy for depression: Content, ratio and preconditions according to patients and therapists using a Delphi study. *BMC Psychiatry* **2014**, *14*, 355.

24. Wentzel, J.; van der Vaart, R.; Bohlmeijer, E.T.; van Gemert-Pijnen, J.E. Mixing online and face-to-face therapy: How to benefit from blended care in mental health care. *JMIR Mental Health* **2016**, *3*, e9.

25. Ruwaard, J.; Kok, R. Wild West eHealth: Time to hold our horses? *Eur. Health Psychol.* **2015**, *17*, 45–49.

26. Zilcha-Mano, S. Major developments in methods addressing for whom psychotherapy may work and why. *Psychother. Res.* **2018**, *29*, 693–708.

27. Kazdin, A.E. Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *Am. Psychol.* **2008**, *63*, 146.

28. Cuijpers, P.; Van Straten, A.; Bohlmeijer, E.; Hollon, S.D.; Andersson, G. The effects of psychotherapy for adult depression are overestimated: A meta-analysis of study quality and effect size. *Psychol. Med.* **2010**, *40*, 211–223.

29. Fournier, J.C.; DeRubeis, R.J.; Shelton, R.C.; Hollon; S.D.; Amsterdam, J.D.; Gallop, R. Prediction of response to medication and cognitive therapy in the treatment of moderate to severe depression. *J. Consult. Clin. Psychol.* **2009**, *77*, 775.

30. Cohen, Z.D.; DeRubeis, R.J. Treatment selection in depression. *Ann. Rev. Clin. Psychol.* **2018**, *14*, 209–236.

31. Fisher, A.J.; Boswell, J.F. Enhancing the personalization of psychotherapy with dynamic assessment and modeling. *Assessment* **2016**, *23*, 496–506.

32. Hollon, S.D.; Cohne, Z.D.; Singla, D.R.; Andrews, P. Recent Developments in the Treatment of Depression. *Behav. Ther.* **2019**, *50*, 257–269.

33. Hamburg, M.A.; Collins, F.S. The path to personalized medicine. *N. Engl. J. Med.* **2010**, *363*, 301–304.

34. van Bronswijk, S.C.; Lemmens, L.H.; Keefe, J.R.; Huibers, M.J.; DeRubeis, R.J.; Peeters, F.P . A prognostic index for long-term outcome after successful acute phase cognitive therapy and interpersonal psychotherapy for major depressive disorder. *Depress. Anxiety* **2018**, *36*, 252–261.

35. Becker, D.;van Breda, W.; Funk, B.; Hoogendoorn, M.; Ruwaard, J.; Riper, H. Predictive modeling in e-mental health: A common language framework. *Internet Interv.* **2018**, *12*, 57–67.

36. DeRubeis, R.J.; Cohen, Z.D.; Forand, N.R.; Fournier, J.C.; Gelfand, L.A.; Lorenzo-Luaces, L. The Personalized Advantage Index: Translating research on prediction into individualized treatment recommendations. A demonstration. *PLoS ONE* **2014**, *9*, e83875.

37. Cohen, Z.D.; Kim, T.; Van, H.; Dekker, J.; Driessen, E. Recommending Cognitive-Behavioral Versus Psychodynamic Therapy for Mild to Moderate Adult Depression. *Psychotherapy Research* **2017,** 30:2, 137-150.

38. Friedl, N.; Berger, T.; Krieger, T.; Caspar, F.; Grosse Holtforth, M. Using the Personalized Advantage Index for individual treatment allocation to cognitive behavioral therapy (CBT) or a CBT with integrated

exposure and emotion-focused elements (CBT-EE). *Psychother. Res.* **2019**. doi:10.1080/10503307.2019.1664782.

39. DeRubeis, R.J.; Hollon, S.D.; Amsterdam, J.D.; Shelton, R.C.; Young, P.R.; Salomon, R.M.; O'Reardon, J.P.; Lovett, M.L.; Gladis, M.M.; Brown, L.L. Cognitive therapy vs medications in the treatment of moderate to severe depression. *Arch. Gen. Psychiatry* **2005**, *62*, 409–416.

40. Huibers, M.J.; Cohen, Z.D.; Lemmens, L.H.; Arntz, A.; Peeters, F.P.; Cuijpers, P.; DeRubeis, R.J. Predicting optimal outcomes in cognitive therapy or interpersonal psychotherapy for depressed individuals using the personalized advantage index approach. *PLoS ONE* **2015**, *10*, e0140771.

41. van Breda, W.; Bremer, V.; Becker, D.; Hoogendoorn, M.; Funk, B.; Ruwaard, J.; Riper, H. Predicting therapy success for treatment as usual and blended treatment in the domain of depression. *Internet Interv.* **2018**, *12*, 100–104.

42. Bremer, V.; Becker, D.; Kolovos, S.; Funk, B.; Van Breda, W.; Hoogendoorn, M.; Riper, H. Predicting therapy success and costs for personalized treatment recommendations using baseline characteristics: Data-driven analysis. *J. Med. Internet Res.* **2018**, *20*, e10275.

43. Kleiboer, A.; Smit, J.; Bosmans, J.; Ruwaard, J.; Andersson, G.; Topooco, N.; Berger, T.; Krieger, T.; Botella, C.; Banos, R. European COMPARative Effectiveness research on blended Depression treatment versus treatment-as-usual (E-COMPARED): Study protocol for a randomized controlled, non-inferiority trial in eight European countries. *Trials* **2016**, *17*, 387.

44. Sheehan, D.V.; Lecrubier, Y.; Sheehan, K.H.; Amorim, P.; Janavs, J.; Weiller, E.; Hergueta, T.; Baker, R.; Dunbar, G.C. The Mini-International Neuropsychiatric Interview (MINI): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J. Clin. Psychiatry* **1998**, *59*, 22–33.

45. Kroenke, K.; Spitzer, R.L.; Williams, J.B. The PHQ-9: Validity of a brief depression severity measure. *J. Gen. Intern. Med.* **2001**, *16*, 606–613.

46. Meyer, B.; Berger T.; Caspar, F.; Beevers, C.G.; Andersson, G.; Weiss, M. Effectiveness of a Novel Integrative Online Treatment for Depression (Deprexis): Randomized Controlled Trial. *J. Med. Internet Res.* **2009**, *11*, e15.

47. Mathiasen, K.; Andersen, T.E.; Riper, H.; Kleiboer, A.A.; Roessler, K.K. Blended CBT versus face-to-face CBT: A randomised non-inferiority trial. *BMC Psychiatry* **2016**, *16*, 432.

48. Kroenke, K. Enhancing the clinical utility of depression screening. *CMAJ Can. Med. Assoc. J.* **2012**, *184*, 281–282.

49. Kroenke, K.; Spitzer, R.L. The PHQ-9: A new depression diagnostic and severity measure. *Psychiatr. Ann.* **2002**, *32*, 509–515.

50. Manea, L.; Gilbody, S.; McMillan, D. Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): A meta-analysis. *Can. Med. Assoc. J.* **2012**, *184*, E191–E196.

51. EuroQol. EuroQol—A new facility for the measurement of health-related quality of life. *Health Policy* **1990**, *16*, 199–208.

52. Bouwmans, C.; De Jong, K.; Timman, R.; Zijlstra-Vlasveld, M.; Van der Feltz-Cornelis, C.; Tan, S.S., Hakkaart-van Roijen, L. Feasibility, reliability and validity of a questionnaire on healthcare consumption and productivity loss in patients with a psychiatric disorder (TiC-P). *BMC Health Serv. Res.* **2013**, *13*, ,217.

53. Devilly, G.J.; Borkovec, T.D. Psychometric properties of the credibility/expectancy questionnaire. *J. Behav. Ther. Exp. Psychiatry* **2000**, *31*, 73–86.

54. Stekhoven, D.J. *missforest: Nonparametric Missing Value Imputation Using Random Forest*; R Package Version, 1; 2013.

55. Stekhoven, D.J.; Buhlmann, P. MissForest: Non-parametric missing value imputation for mixed-type data. *Bioinformatics* **2012**, *28*, 112–118.

56. Waljee, A.K.; Mukherjee, A.; Singal, A.G.; Zhang, Y.; Warren, J.; Balis, U.; Marrero, J.; Zhu, J.; Higgins, P.D. Comparison of imputation methods for missing laboratory data in medicine. *BMJ Open* **2013**, *3*, e002847.

57. Peduzzi, P.; Concato, J.; Feinstein, A.R.; Holford, T.R. Importance of events per independent variable in proportional hazards regression analysis II. Accuracy and precision of regression estimates. *J. Clin. Epidemiol.* **1995**, *48*, 1503–1510.

58. Vittinghoff, E.; McCulloch, C.E. Relaxing the rule of ten events per variable in logistic and Cox regression. *Am. J. Epidemiol.* **2007**, *165*, 710–718.

59. Raftery, A.E.; Madigan, D.; Hoeting, J.A. Bayesian model averaging for linear regression models. *J. Am. Stat. Assoc.* **1997**, *92*, 179–191.

60. Stock, E.M.; Kimbrel, N.A.; Meyer, E.C.; Copeland, L.A.; Monte, R.; Zeber, J.E.; Gulliver, S.B.; Morissette, S.B. A Bayesian model averaging approach to examining changes in quality of life among returning Iraq and Afghanistan Veterans. *Int. J. Methods Psychiatry Res.* **2014**, *23*, 345–358.

61. Babyak, M.A. What you see may not be what you get: A brief, nontechnical introduction to overfitting in regression-type models. *Psychosom. Med.* **2004**, *66*, 411–421.

62. Fragoso, T.M.; Bertoli, W.; Louzada, F. Bayesian model averaging: A systematic review and conceptual classification. *Int. Stat. Rev.* **2018**, *86*, 1–28.

63. Fernandez, C.; Ley, E.; Steel, M.F. Benchmark priors for Bayesian model averaging. *J. Econom.* **2001**, *100*, 381–427.

64. Fernandez, C.; Ley, E.; Steel, M.F. Model uncertainty in cross-country growth regressions. *J. Appl. Econom.* **2001**, *16*, 563–576.

65. Ley, E.; Steel, M.F.J. On the effect of prior assumptions in Bayesian model averaging with applications to growth regression. *J. Appl. Econom.* **2009**, *24*, 651–674.

66. Steel, M.F. Bayesian model averaging and forecasting. *Bull. Inflat. Macroecon. Anal.* **2011**, *200*, 30–41.

67. Clyde, M.A. BAS: Bayesian Variable Selection and Model Averaging using Bayesian. In *Adaptive Sampling*; 2018; R Package Version 1.5.3; Available online: https://cran.r-project.org/web/packages/BAS/BAS.pdf (accessed on September 15th, 2019).

68. Clyde, M.A.; Ghosh, J.; Littman, M.L. Bayesian adaptive sampling for variable selection and model averaging. *J. Comput. Gr. Stat.* **2011**, *20*, 80–101.

69. Efron, B.; Gong, G. A leisurely look at the bootstrap, the jackknife, and cross-validation. *Am. Stat.* **1983**, *37*, 36–48.

70. Deisenhofer, A.K.; Delgadillo, J.; Rubel, J.A.; Böhnke, J.R.; Zimmermann, D.; Schwartz, B.; Lutz, W. Individual treatment selection for patients with posttraumatic stress disorder. *Depress. Anxiety* **2018**, *35*, 541–550.

71. Lowe, B.; Unutzer, J.; Callahan, C.M.; Perkins, A.J.; Kroenke, K. Monitoring depression treatment outcomes with the patient health questionnaire-9. *Med. Care* **2004**, *42*, 1194–201.

72. Dinga, R.; Marquand, A.F.; Veltman, D.J.; Beekman, A.T.F.; Schoevers, R.A.; van Hemert, A.M.; Penninx, B.; Schmaal, L. Predicting the naturalistic course of depression from a wide range of clinical, psychological, and biological data: A machine learning approach. *Transl. Psychiatry* **2018**, *8*, 241.

73. Hamilton, K.E.; Dobson, K.S. Cognitive therapy of depression: Pretreatment patient predictors of outcome. *Clin. Psychol. Rev.* **2002**, *22*, 875–893.

74. Constantino, M.J.; Arnkoff, D.B.; Glass, C.R.; Ametrano, R.M.; Smith, J.Z. Expectations. *J. Clin. Psychol.* **2011**, *67*, 184–192.

75. Greenberg, R.P.; Constantino, M.J.; Bruce, N. Are patient expectations still relevant for psychotherapy process and outcome? *Clin. Psychol. Rev.* **2006**, *26*, 657–678.

76. Spek, V.; Nyklíček, I.; Cuijpers, P.; Pop, V. Predictors of outcome of group and internet-based cognitive behavior therapy. *J. Affect. Disord.* **2008**, *105*, 137–145.

77. Andersson, G.; Hedman, E. Effectiveness of guided internet-based cognitive behavior therapy in regular clinical settings. *Verhaltenstherapie* **2013**, *23*, 140–148.

78. Bower, P.; Kontopantelis, E.; Sutton, A.; Kendrick, T.; Richards, D.A.; Gillbody, S.; Knowles, S.; Cuijpers, P.; Andersson, G.; Christensen, H.; et al. Influence of initial severity of depression on effectiveness of low intensity interventions: Meta-analysis of individual patient data. *BMJ (Clin. Res.)* **2013**, *346*, f540.

79. Thase, M.E.; Reynolds, C.F.; Frank, E.; Simons, A.D. Response to cognitive-behavioral therapy in chronic depression. *J. Psychother. Pract. Res.* **1994**, *3*, 304–314.

80. Blom, M.B.; Spinhoven, P.; Hoffman, T.; Jonker, K.; Hoencamp, E.; Haffmans, P.J.; van Dyck, R. Severity and duration of depression, not personality factors, predict short term outcome in the treatment of major depression. *J. Affect. Disord.* **2007**, *104*, 119–126.

81. Delgadillo, J.; Huey, D.; Bennett, H.; McMillan, D. Case complexity as a guide for psychological treatment selection. *J. Consult. Clin. Psychol.* **2017**, *85*, 835–853.

82. Luedtke, A.; Sadikova, E.; Kessler, R.C. Sample Size Requirements for Multivariate Models to Predict Between-Patient Differences in Best Treatments of Major Depressive Disorder. *Clin. Psychol. Sci.* **2019**, *7*, 445–461.

83. van Bronswijk, S.C.; DeRubeis, R.J.; Lemmens, L.H.; Peeters, F.P.; Keefe, J.R.; Cohen, Z.D.; Huibers, M. Precision medicine for long-term depression outcomes using the Personalized Advantage Index approach:

Cognitive therapy or interpersonal psychotherapy? *Psychol. Med.* **2019**, 1–11. doi:10.1017/S0033291719003192.

84. Gillan, C.M.; Whelan, D.R. What big data can do for treatment in psychiatry. *Curr. Opin. Behav. Sci.* **2017**, *18*, 34–42.

85. Urech, A.; Krieger, T.; Möseneder, L.; Biaggi, A.; Vincent, A.; Poppe, C.; Meyer, B.; Riper, H.; Berger, T. A patient post hoc perspective on advantages and disadvantages of blended cognitive behaviour therapy for depression: A qualitative content analysis. *Psychother. Res.* **2019**, *29*, 986–998.