Journal of Affective Disorders 213 (2017) 78–85

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Research paper

A prognostic index (PI) as a moderator of outcomes in the treatment of depression: A proof of concept combining multiple variables to inform risk-stratified stepped care models

Lorenzo Lorenzo-Luaces[a,*], Robert J. DeRubeis[a], Annemieke van Straten[b], Bea Tiemens[c,d]

[a] Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA
[b] Department of Clinical Psychology, VU University Amsterdam, Amsterdam, The Netherlands
[c] Department of Clinical Psychology, Radboud University, Nijmegen, The Netherlands
[d] Pro Persona Research, Renkum, The Netherlands

ARTICLE INFO

Keywords:
Depression
Psychotherapy
Stepped care
Prognosis
Personalized medicine

ABSTRACT

Background: Prognostic indices (PIs) combining variables to predict future depression risk may help guide the selection of treatments that differ in intensity. We develop a PI and show its promise in guiding treatment decisions between treatment as usual (TAU), treatment starting with a low-intensity treatment (brief therapy (BT)), or treatment starting with a high-intensity treatment intervention (cognitive-behavioral therapy (CBT)).
Methods: We utilized data from depressed patients (N=622) who participated in a randomized comparison of TAU, BT, and CBT in which no statistically significant differences in the primary outcomes emerged between the three treatments. We developed a PI by predicting depression risk at follow-up using a LASSO-style bootstrap variable selection procedure. We then examined between-treatment differences in outcome as a function of the PI.
Results: Unemployment, depression severity, hostility, sleep problems, and lower positive emotionality at baseline predicted a lower likelihood of recovery across treatments. The PI incorporating these variables produced a fair classification accuracy (c=0.73). Among patients with a high PI (75% percent of the sample), recovery rates were high and did not differ between treatments (79–86%). Among the patients with the poorest prognosis, recovery rates were substantially higher in the CBT condition (60%) than in TAU (39%) or BT (44%).
Limitations: No information on additional treatment sought. Prospective tests needed.
Conclusion: Replicable PIs may aid treatment selection and help streamline stepped models of care. Differences between treatments for depression that differ in intensity may only emerge for patients with the poorest prognosis.

1. Introduction

There are a wide variety of treatments for depression. These treatments differ in how intense they are in terms of investment and resources required from patients and providers but the average superiority of higher-intensity treatments relative to lower-intensity ones appears to be small. For example, the combination of antidepressants and psychotherapy appears to be superior to either treatment as a monotherapy but that difference is small (Khan et al., 2012). Similarly, the superiority of antidepressants relative to placebos appears limited (Kirsch and Sapirstein, 1998) and the superiority of evidence-based psychotherapy relative to treatment as usual appears to be small as well (Flückiger et al., 2014; Wampold et al., 2011). Despite the availability and approximate equivalence of many treatments for depression, current models for the delivery of care are inadequate. At one end of the spectrum, many patients receive higher intensity interventions than they require to experience symptom relief (Lorenzo-Luaces et al., 2015; Lovell and Richards, 2000). Conversely, many do not receive the level of care that they might require to experience symptom relief (Kocsis et al., 2008; Lecrubier, 2007).

It is difficult to match patients to an appropriate level of care because depression is extremely heterogeneous in presentation and prognosis (Lorenzo-Luaces, 2015; Parker, 2005). We describe an approach, conceptually motivated by research on risk-stratified care in breast cancer (Akay et al., 2012; Chen et al., 2004; Chen et al., 2005; Huang et al., 2006), for combining variables to create a prognostic

* Correspondence to: Levine Building, 433 S. University Avenue, Philadelphia, PA 19104, USA.
E-mail address: lorenzl@sas.upenn.edu (L. Lorenzo-Luaces).

http://dx.doi.org/10.1016/j.jad.2017.02.010
Received 17 July 2016; Received in revised form 29 January 2017; Accepted 6 February 2017
Available online 07 February 2017
0165-0327/

index (PI) that can be used when selecting between treatments that differ in intensity. PIs can be thought of as predictive of patients' symptom status in a future time frame and they can also be used to determine the intensity level of care a patient should receive (see Delgadillo et al. (2016) and Garber et al. (2016)).

1.1. Prognosis in breast cancer

In the treatment of breast cancer after neoadjuvant chemotherapy, mastectomies are known to be a more aggressive treatment than breast-conservation therapy (BCT), which preserves a part of the breast tissue for cosmetic purposes. While these treatments clearly differ in intensity, on average there are only small (1–9%) differences in 5–10 year recurrence rates between the two treatments (Morris et al., 1997; van Dongen et al., 2000). Although this average difference is small, it is possible that BCT might be indicated for patients with less aggressive cancers whereas mastectomies, the more aggressive or higher intensity treatment, are better-suited for patients with more aggressive illnesses. This possibility was explored by Huang et al. (2006) who analyzed treatment differences according to a previously-developed PI (Chen et al., 2005). Their PI consisted of a score, ranging from 0 to 4, which indicated the number of risk factors for a cancer recurrence (see Chen et al., 2004). For patients with a low predicted risk of recurrence, recurrence rates were low and did not differ between the treatments (12% for BCT, 9% for mastectomy). For patients with a score of 3–4, however, recurrence rates were significantly higher for those treated with BCT (61%) than for those treated with mastectomy (19%). This pattern of results were subsequently replicated by Akay et al. (2012) who reported no differences between BCT and mastectomy for patients who were at a low risk of recurrence but a large difference in favor of mastectomy (6% vs. 32%) for patients with a high risk score. We propose that a similar phenomenon occurs in the treatment of depression and other disorders. Specifically, comparisons of treatments that appear to differ in intensity (e.g., combination treatment vs. psychotherapy alone or long term therapy vs. brief interventions) may produce small differences because many patients in treatment trials can be expected to benefit from minor interventions and it is only patients with an overall poorer prognosis that will evidence more benefit from the higher-intensity treatment.

1.2. Combining multiple variables in depression treatment

The idea of using baseline variables to predict response to depression treatments is not new. For example, the efficacy of antidepressants relative to placebos seems to be limited to patients with more severe depressions (Barbui et al., 2011; Fournier et al., 2010; Khan et al., 2002; Kirsch et al., 2008). Similar findings have been reported for psychotherapy (Driessen et al., 2010). Likewise, the efficacy of combination treatment relative to antidepressants alone or psychotherapy alone seems to be limited to patients with more severe depression (Hollon et al., 2014; Thase et al.al., 1997). Some stepped models of care for depression make use of findings such as these. In stepped models of care, most patients are started on lower-intensity treatment options before entering more intense care. Illness severity, however, can be used as an indicator to bypass lower intensity treatments for more intensive ones. For example, the NICE guidelines in the United Kingdom (National Institute for Clinical Excellence, 2004) do not endorse the use of combined treatment as the initial treatment strategy in mild depression and instead suggest further assessment, a low-intensity intervention, or a monotherapy.

Research suggests that a multitude of variables, other than symptom severity, account for treatment response (see Kessler et al. (2016a)). For example, in a large dataset pooling comparisons of antidepressants vs. placebos, Nelson et al. (2013) reported that there were no statistically significant differences in outcomes for patients whose depression was non-chronic. However, a large difference (d=0.70) emerged for the subset of patients who had chronic and severe depression. Similarly, Thase et al. (1997) reported that, for individuals with mild depression, combining an antidepressant and cognitive-behavioral therapy (CBT) or interpersonal psychotherapy (IPT) was not superior to psychotherapy alone. However, for patients with severe and recurrent depression, combination treatment yielded superior recovery rates (60% vs. 19%; see also Hollon et al., 2014). Kessler et al. (2016a) conducted a review of patient self-reported variables that have been replicated at least once as predictors or moderators of treatment outcomes. According to these authors, predictors of depression can broadly be grouped into demographics (e.g., age, unemployment), features of depression (e.g., severity, prior episodes), co-morbidities (e.g., anxiety, sleep problems), stress history (e.g., childhood maltreatment), personality features (e.g., high negative affect), and other features (e.g., impairment). Broadly speaking, it appears as if the variables that predict overall depression treatment response are variables that predict the persistence and severity of depression. These are usually variables that relate to underlying vulnerabilities to depression and social-interpersonal functioning. However, an issue that obscures the interpretation of the existing literature is that most studies only explore single variables (DeRubeis et al., 2014a; Kessler et al., 2016b).

Kraemer (2013) asserted that "if there are multiple [moderators] related to the same underlying construct, these ... should be combined in order both to increase the reliability of the measurement of that construct and to avoid problems associated with multicollinearity in combining them." (p. 1969) DeRubeis et al. (2014b) argued for the existence of one such construct when they discussed patient response profiles. According to these authors, patients differ in the extent to which they can benefit from the active effects of treatments and processes. Patients who are likely to improve much irrespective of interventions, as is characteristic of samples of patients with depression, are unlikely to reveal specific intervention effects. Preliminary evidence supports the use of combined moderator variables like PIs in guiding treatment selection in mental health treatment (see Cloitre et al., 2016; Delgadillo et al., 2016). In the context of a clinical trial in which there were no differences in overall outcome between treatment as usual (TAU), stepped care starting with brief therapy (BT) or stepped care starting with CBT, we hypothesized that prognostic status would moderate the treatment differences. Our study is meant as a proof of concept in the treatment of depression. We hypothesize that among patients who, based on pre-treatment characteristics, are predicted to do well, few if any differences in outcome will emerge between the treatments we studied. However, among patients with a poorer prognosis, the more intensive CBT should outperform TAU and BT.

2. Methods

The aim of the trial from which these data were drawn was to compare TAU to each of two stepped care regimens. One of the regimens began with a low-intensity treatment (i.e., BT) and the other began with a high-intensity treatment (i.e., CBT; Van Straten et al., 2006).

The trial was designed so as to mimic conditions found in routine care settings. Patients were sampled from a representative subsample of 7 of the 47 regional mental health care centers (MHCs) that provide mental health care in the Netherlands. Exclusion criteria were: the presence of psychotic, manic, or thought disorder symptoms, dependence on hard drugs (patients with alcohol abuse or dependence were not excluded), high suicide risk, or poor command of Dutch. Patients not excluded on the basis of these criteria were screened for mood or anxiety disorders (Goldberg et al., 1988; Tiemens, 1999) with those who screened positive (n=1608) being followed up for an at-home interview with the Composite International Diagnostic Interview (CIDI; Wittchen et al., 1993). Of these patients, 214 could not be

Table 1
Baseline demographic, personality, and clinical characteristics of subjects randomized to treatment as usual (TAU, n=234), brief therapy (BT, n=179), or cognitive-behavioral therapy (CBT, n=208).

	TAU		BT		CBT				%
	%	(n)	%	(n)	%	(n)	χ^2	*p*	*miss.*
On antidepressants	42%	(85)	25%	(50)	33%	(67)	2.91	0.23	9%
Female [a]	66%	(154)	62%	(110)	59%	(123)	2.17	0.34	
Dutch immigrant [a,b]	6%	(13)	10%	(16)	14%	(27)	7.51	0.02	10%
Problematic drinking [b]	11%	(24)	13%	(21)	10%	(19)	0.69	0.71	9%
Unemployed||	36%	(77)	30%	(48)	31%	(57)	2.04	0.36	9%
Educational attainment||							6.47	0.37	10%
None	12.6%	(26)	9.9%	(16)	11.7%	(22)			
Lower	39.1%	(84)	43.8%	(71)	46.3%	(87)			
Middle	38.1%	(82)	30.2%	(49)	30.9%	(58)			
High	10.7%	(23)	16.0%	(26)	11.2%	(21)			
Somatic illnesses (#)||							4.12	0.85	8%
1.00	30.1%	(65)	32.7%	(53)	30.3%	(57)			
2.00	16.7%	(36)	20.4%	(33)	18.6%	(35)			
3.00	9.7%	(21)	8.60%	(14)	9.6%	(18)			
4+	12.5%	(27)	7.4%	(12)	8.5%	(16)			
Anxiety co-morbidity||	44.9%	(105)	45.8%	(82)	48.6%	(101)	0.63	0.73	
Severe MDD (CIDI)||	39.5%	(90)	38.2%	(66)	41.8%	(84)	1.97	0.74	3%
Recurrent MDD (CIDI)||	57.7%	(135)	49.2%	(88)	60.6%	(126)	5.43	0.07	
	M	SD	M	SD	M	SD	F	*p*	
Age||	35.98	(10.29)	36.53	(10.27)	36.63	(9.9)	0.26	0.77	
Symptom Checklist 90									
Sleep||	9.22	(3.85)	9.26	(3.85)	9.29	(3.66)	0.02	0.98	9%
Agoraphobia [a,c]	13.29	(6.31)	13.96	(7.03)	14.27	(6.77)	1.13	0.32	9%
Anxiety [c]	26.06	(8.85)	26.18	(8.97)	27.41	(8.77)	1.35	0.26	9%
Depression [c]	48.58	(12.35)	47.99	(14.24)	49.15	(12.68)	0.34	0.71	9%
Hostility||	12.95	(5.37)	12.49	(5.28)	12.44	(5.36)	0.57	0.57	9%
Insufficiency [a,c]	24.93	(7.51)	24.11	(7.95)	25.33	(7.57)	1.14	0.32	9%
Interpersonal sensitivity [c]	42.48	(15.56)	40.87	(15.06)	43.87	(14.96)	1.69	0.19	9%
Somatic[c]	29.51	(9.55)	28.24	(9.52)	28.34	(9.14)	1.12	0.33	9%
Other [a,c]	19.32	(5.87)	18.79	(5.92)	19.47	(6.16)	0.61	0.54	9%
NEO									
Conscientiousness [a]	3.25	(1.95)	2.79	(1.7)	3.11	(1.84)	2.22	0.11	9%
Agreeableness||	4.43	(2.06)	4.27	(2.)	4.21	(2.)	0.49	0.62	9%
Extraversion||	2.23	(1.83)	3.22	(1.94)	3.12	(1.84)	0.18	0.84	9%
Neuroticism||	7.77	(1.13)	1.89	(1.07)	7.88	(1.11)	0.55	0.58	9%
Openness||	4.99	(1.87)	4.80	(2.07)	5.39	(1.9)	3.22	0.04	9%

Note. % miss – percentage missing data at baseline. MDD – major depressive disorder. Problematic drinking – as assessed by two affirmative responses to the CAGE questionnaire. CIDI – Composite International Diagnostic Interview. NEO- NEO Five Factor Inventory. || variable explored as a predictor of outcomes. a – variable not explored as a predictor of outcomes because no prior research suggests it is predictive of outcomes in depression treatment, b – variable not explored as a predictor of outcomes because there was low variability, c – variable not explored as a predictor of outcomes because it was co-linear with some variables and represented by other variables

reached. Patients in the original trial were eligible if they met the criteria for any of the following DSM-IV (American Psychiatric Association, 2000) disorders: major depressive disorder (MDD), dysthymia, panic, social anxiety, or generalized anxiety disorder. Of the 1394 patients reached for interview, 214 did not meet full criteria for mood or anxiety disorder, 17 had a poor command of Dutch, and 396 declined to participate. Thus, in total, 702 patients consented to participate in the parent study. In the present paper, we focus on the 622 patients who met criteria for MDD. Prior to the start of treatment, patients were randomized to one of three treatment conditions:

2.1. Treatment as usual (TAU)

The TAU condition consisted of matched care as it was conducted in the Netherlands at the time: an interdisciplinary mental health care team reviewed each case and patients were assigned to the treatment that they were expected to benefit the most from. Treatments varied by type (e.g., CBT, interpersonal, psychodynamic), format (e.g., group, online), and intensity (i.e., duration). Thus, ultimately the TAU condition consisted of various treatments of different treatment intensity.

2.2. Brief therapy(BT)

In the 1980s, brief therapy (BT) was introduced in the Netherlands as a remedy for lengthy waiting lists. In this study, BT was provided for a total of 5 sessions with a maximum of 2 booster sessions in the six-month period following treatment completion. The aim of BT in this study was to create hope by clarifying problems and emphasizing and strengthening the patient's own competence and coping skills.

2.3. Cognitive-behavioral therapy (CBT)

In this study, CBT consisted of five modules spanning 11–15 sessions: a) introduction (one session), b) psychoeducation and assessing of cognitions (three sessions); iii) changing cognitions by challenging them (three sessions); iv) behavior experiments to challenge cognitions (three sessions); v) integrating new behavior in patients' lives by additional behavior experiments (one to five sessions).

BT and CBT were considered as first steps in a stepped-care model. Therefore, all patients were allowed to switch treatments, during or after treatment completion, if either the patient or the therapist deemed appropriate. In other words, although the BT and CBT conditions had a set protocol and numbers of sessions, patients were

allowed to 'step up' from these treatments, receive more treatment sessions, or to switch treatment modalities or formats. Thus, patients in the BT condition could ultimately receive CBT. Because treatment was conducted in a naturalistic setting, the randomization rule was sometimes overturned by clinicians prior to the start of treatment. This was somewhat more likely to occur in the BT condition but was not associated with outcomes in any of the conditions (ps > 0.30). No information was available in terms of the timing, frequency, or types of shifts that occurred in any of the conditions.

Patients who met the criteria for severe depression in any of the three treatment conditions were allowed to receive antidepressant medication in addition to the psychological treatment. There were no statistically significant differences in the proportion of patients taking medication in TAU (42%, n=85), BT (25%, n=50), and CBT (33%, n=67). Medication use did not predict outcome either by itself or in interaction with any of the treatment condition (ps > 0.63).

2.1 Outcomes and missing data

Patients were interviewed at baseline and then every 3 months, irrespective of the timing of treatment initiation and termination. The primary outcome for the current analyses was recovery, defined by the absence of MDD status, at the 18–24 month follow-up. This final follow-up interview occurred at least 18 months after enrollment in the study. The first 59 patients who entered the study were followed for 24 months; subsequent enrollees were followed for 21 months (n=105) or 18 months (n=256). We also examined recovery at an earlier (12-month) assessment.

At baseline, rates of missing data on co-variates were low (all < 10%, see table 1) and most (86%. n=536) participants in our sample had information on all baseline variables. At the end of the study, 68% of the participants (n=420) were available to be interviewed. There were no statistically significant differences in lost-to-follow-up (LTFU) between the three treatments (ps > 0.66). Most individuals (n=416) who were followed up in the 18–24 month assessment provided data on baseline co-variates. To address missing data, including missing outcome assessments, we used a non-parametric missing value imputation procedure using random forests with the R package missForest (Stekhoven and Bühlmann, 2012). The variables in the imputation model were the baseline co-variates reported in Table 1, treatment condition, number of sessions, medication status, and follow-up time frame (i.e., 18, 21, 24). Imputation via random forests has been shown to yield a lower imputation error than the more commonly-known approach of multiple imputations via chained equations (MICE; see Stekhoven and Bühlmann, 2012; Waljee et al., 2013). To check for the potential that missing data imputation influenced our results, we re-ran the analyses described below with the listwise-deleted version of the dataset (n=417). The results were, by and large, quite similar so we report the results obtained with the imputed data.

2.2 Analytic approach

Analyses were conducted using the R programming language (R Core Team, 2014). A total of 23 variables were available for analysis. These included demographics, clinical variables, personality traits as assessed by the NEO Five-Factor Inventory (Costa and MacCrae, 1992), and subscales of the Symptom Checklist 90 (Derogatis, 1983; Table 1). In choosing which variables to explore as predictors of treatment outcomes, we cross-referenced a recent review by Kessler et al. (2016a) on predictors of depression treatment outcomes that have been replicated at least once. When removing variables, we took into account multicollinearity in the context of redundancy or multiple indicators (e.g., both objective and subjective levels of somatic symptoms); lack of prior research suggesting variable predicts depression outcome (i.e., 6 out of 23), and observed variability. Thirteen of the 23 variables were thus retained (Table 1).

To determine which of the 13 variables would be included in our PI, we utilized the 'SparseLearner' R package developed by Guo et al. (2015). This package includes a regression-weight ranking procedure with a k-fold cross-validation across bootstrap samples using LASSO (least absolute shrinkage and selection operator) penalties. LASSO approaches belong to the family of penalized regression models (Tibshirani, 1996, 2011) and have been recommended and used in other efforts to predict MDD status (Kessler et al., 2016b). The procedure converges on a final solution in which the variables are given a set of weights that are more conservative, and thus more likely to replicate than what would be obtained with other regression approaches. We ran the analyses to predict recovery status at the 18–24 month follow-up in 5000 bootstrapped samples using 10-fold cross-validation. Thus, in each of the 5000 bootstrapped samples, recovery status at the 18–24 month follow-up was derived by a regression equation of the form:

$$\text{Recovery 18–24 mo.} = \beta 0 + (\beta 1 * \text{Var1}) + (\beta 2 * \text{Var2}) \ldots (\beta 13 * \text{Var13})$$

In each of the 5000 bootstrapped samples, the β estimates correspond to the values that minimize 10-fold cross-validation prediction errors. Final weights are computed by aggregating β estimates as described in the formulas by Guo et al. (2015). Each patient's predicted likelihood of recovery can then be estimated by entering their baseline values on the relevant variables into the equation produced by this procedure. We refer to this estimate as the patient's value on the PI. Because the PI is derived for a prediction of a binary variable, it is in the form of log odds. To facilitate interpretation, we convert it to an estimate of the probability of recovery (PI%).

We used the traditional interpretation of the c-statistic/area under the curve to evaluate the performance of the PI in predicting depression recovery (Harrell, 2001). To determine whether outcomes varied between the three treatment conditions, we ran binary logistic regressions predicting outcomes at follow-up with dummy variables for each of the two active treatment conditions (i.e., CBT vs. TAU and BT vs. TAU), the PI, and the respective interactions between the PI and treatment condition:

$$\text{Recovery 18–24 mo.} = \beta 0 + (\beta 1 * \text{PI}) + (\beta 2 * \text{CBT}) + (\beta 3 * \text{BT}) + (\beta 4 * \text{CBT} * \text{PI}) + (\beta 5 * \text{BT} * \text{PI})$$

In this model, the terms for β4 and β5 indicate whether the coefficient for the interaction between the PI and CBT and the PI and BT is significantly different than zero. If the tests associated with these coefficients are statistically significant, it indicates that outcomes vary between the treatment across levels of the PI. We conducted a similar logistic regression comparing CBT vs. BT.

In all analyses, we controlled for medication status, the timing of the follow-up (i.e., 18, 21, or 24 months), and the number of treatment sessions attended, although there was no indication that either variable by itself, or in any of the treatment conditions, was related to recovery status. To probe significant interactions between the PI and treatment condition, we used the Johnson-Neyman technique (Hayes and Matthes, 2009), which gives a value of the moderator at which the significance of the predictor on outcome changes. In our analyses, the Johnson-Neyman would help us establish a point in the prognostic index at which treatment effects begin to be evident.

3. Results

Descriptive statistics at baseline are presented in Table 1. There were no statistically significant differences in the rate of MDD recovery at the 18–24 month follow-up between TAU (68.4%) and either of the stepped care conditions (BT=75.4%, OR=1.33, 95% CI=0.84 −2.07, B =0.28, SE=0.23, $\chi 2$=1.47, p=0.23; CBT=74.5%, OR=1.38, 95% CI=0.91 −2.11, B =0.32, SE=0.22, $\chi 2$=2.25, p=0.13). Five of the 13 potential predictor variables submitted to the LASSO procedure were

Table 2
LASSO solution for predictor variables used in prognostic index (PI) predicting recovery at follow-up and model estimates from logistic regression.

	LASSO	B	SE	OR	95% CI
Unemployment status	−0.39	−0.57	0.21	0.57	(0.38–0.84)
Depression severity	−0.15	−0.20	0.11	0.82	(0.67–1.01)
Hostility (SCL)	−0.38	−0.45	0.10	0.64	(0.52–0.77)
Sleep complaints (SCL)	−0.25	−0.31	0.11	0.74	(0.59–0.91)
Extraversion (NEO)	0.34	0.43	0.11	1.53	(1.23–1.90)

Note: Vertical lines separates results of LASSO solution from results of logistic regression. 95% CIs SCL: Symptom Checklist 90-R subscale. NEO: NEO Five Factor Inventory subscale.

retained. Being unemployed, having more severe symptoms of depression, higher levels of hostility, having more sleep problems, and having lower levels of extraversion/positive emotionality predicted a lower likelihood of recovery. Table 2 shows the standardized weights assigned to these variables by the LASSO model, along with the results obtained when the five variables were entered into a logistic regression predicting the likelihood of recovery. The resulting PI, generated from the LASSO, evidenced fair predictive accuracy (c=0.73, 95% CI=0.68 – 0.77, $p < 0.001$).

The PI was developed by predicting treatment outcomes ignoring the main effect of treatment condition on outcomes, which, albeit small and not statistically significant, was in the expected direction. Before examining whether it was related to outcomes across the treatments, we carried out a series of analyses to rule out the possibility that there was a systematic influence of treatment condition on the PI. First, we evaluated the treatment comparisons in a regression that contained the five variables selected previously: unemployment status, depression severity, hostility, sleep problems, and extraversion. The inclusion of these effects in a model containing all the prognostic variables did not affect their statistical significance or the strength of their predictive relation vis a vis outcome. We then examined whether the PIs themselves differed between the treatment conditions and found no differences ($F(2, 618)=0.95$, $p=0.39$). Taken together, these results suggest that treatment assignment did not influence the PI.

In the primary analyses predicting recovery at the 18–24 month follow-up, the test of the interaction of the PI and the CBT vs. TAU contrast was significant. The direction of the effect indicated that, consistent with our hypothesis: the poorer the overall prognosis, the greater the advantage of CBT relative to TAU (OR=1.90, 95% CI: 1.21–3.26, B=0.69, SE=0.25 $\chi2=7.45$, $p=0.008$) or BT (OR =1.88, 95% CI: 1.09–3.26, B=0.63, SE=0.28 $\chi2=5.12$, $p=0.024$). The BT/TAU contrast, however, did not interact with the PI in the prediction of outcome (OR=1.05, 95% CI: 0.601–1.85, B=0.06, SE=0.28, $\chi2=0.04$, $p=0.85$). This pattern was not evident in the data from the earlier 12-month assessment (all ps > 0.35).

The Johnson-Neyman technique was employed to explore the significant interaction of the PI with the CBT/TAU contrast at follow-up. The result of this procedure suggested a cutoff on the PI that reflected a 64% predicted likelihood of recovery, and which divided the sample into the 25% with the worst prognoses (i.e., with PI% < 64%) and the 75% with better prognoses (i.e., with PI% > 64%). Of the patients with a better prognosis, the observed likelihood of recovery ranged from 79% to 86% and there were no differences between the conditions in recovery rates (see Fig. 1). By contrast, among the patients with poorer prognoses, a higher percentage recovered in CBT (60.0%) relative to TAU (38.8%; OR =2.53, 95% CI: 1.15–3.22, B=0.93, SE=0.40, Z=2.30, p=0.02) and BT (44.4%; OR =2.03, 95% CI: 1.49–4.43, B=0.71, SE=0.40, Z=1.76, p=0.08).

To assess whether any single variable unduly influenced the PI, we recalculated it five times, each time removing one of the five variables. The results remained the same: there was an interaction between the PI and the CBT vs. TAU contrast (ORs > 2.1, ps < 0.01). This interaction was not present in the BT vs. TAU contrasts (ps > 0.62). We also tested the interactions between treatment condition and each of the five patient variables separately. None of the interaction effects was significant (all ps > 0.09).

4. Discussion

We described a procedure that yields a prognostic index that can be used in determining which patients are most likely to benefit more from a high-intensity intervention, relative to a lower intensity one. We tested this procedure in the context of a randomized trial comparing TAU, BT, and CBT. Despite the trial being adequately powered to detect treatment differences and the fact that the treatments differed in intensity, on average, there were only small, nonsignificant differences in outcomes. Using our prognostic index, we identified a subgroup of patients for whom the effects of CBT were substantially greater than those of TAU or BT. It is noteworthy that large differences between the interventions were observed in the subset of patients with poorer prognoses because this study was conducted in a naturalistic context; patients in the BT condition were allowed to step up their treatment and the TAU was a strong control group in that few restrictions were placed on the level and type of care. To put the findings in context, the differences reported between CBT and TAU are outside the 95% CI of estimates of comparisons of CBT vs. TAU reported in the latest meta-analysis exploring long-term outcomes between psychotherapies (Karyotaki et al., 2016).

Before interpreting our study findings, several limitations are worth noting. First, the present data were collected over 10 years ago in a naturalistic context in which randomization was not guaranteed and assessments of treatment integrity were not conducted. We chose this specific dataset because of its strengths. It included a treatment arm that had a high-intensity empirically supported intervention (i.e., CBT), a high-intensity intervention that was not protocolized so as to be supported by research (i.e., TAU), and a low-intensity intervention that has arguably less empirical support (i.e., BT). However, the fact that we only have information on the number of sessions attended and not on the type of psychotherapy or whether the patients stepped up is a major limitation. In spite of the age of the study, design features that are threats to internal validity (e.g., lack of fidelity assessments), this is still a uniquely-designed study. The fact that patterns treatment differences were obtained in this secondary data analysis suggests that the groups of patients received treatments that differed at least somewhat. In either case, it is difficult to envision how the age of the study, the failure to randomize, or the lack of fidelity assessments accounted for the patterns we reported. Moreover, it is worth considering that our study is not intended to change clinical practice but to illustrate a principle and stimulate further research. The findings need validation in an alternative dataset and testing in prospective studies so as to influence clinical practice. It is also worth observing that present trial was a multi-site study. Controlling for site-level effects would have lead to a less biased estimation of outcomes. Given the complexity of our analyses and the low likelihood that this accounts for the observed pattern of results, which was predicted a priori, we chose not to account for the nesting of the data.

We have interpreted our findings to mean that a higher intensity of care is needed for patients with a poorer prognosis. However, it is possible that patients with a better prognosis benefit as much from non-specific interventions such as BT as they do from empirically-supported treatments like CBT but patients with a poorer prognosis require treatments with more active ingredients (e.g., CBT) to experience treatment benefit (see DeRubeis et al. (2014b)). It is possible that the degree to which a presumably active treatment is superior to a control is a function of both the presence of interventions targeted to specific psychological processes and to the frequency with which the intervention is delivered. This is not inconsistent with the ideas we have discussed. In other words, if any differences between treatments

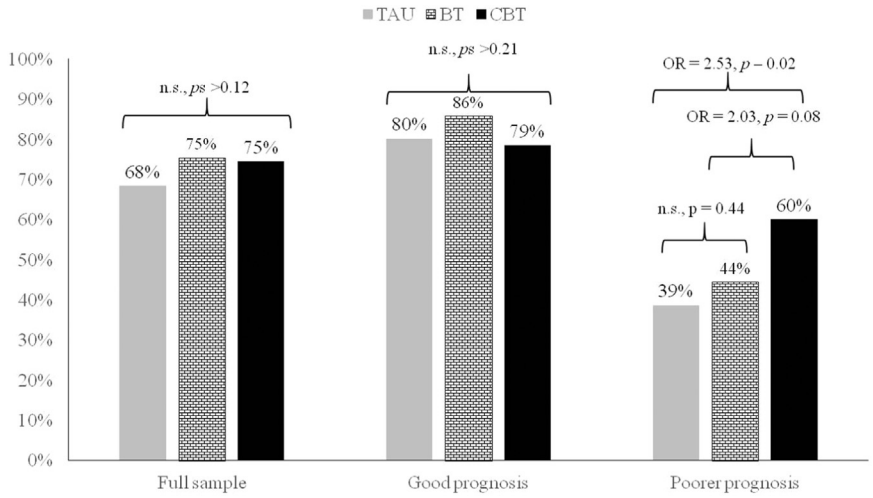

Fig. 1. Recovery rates at 18–24 follow-up in treatment as usual (TAU, n=234), stepped care starting with brief therapy (BT, n=179), and cognitive-behavioral therapy (CBT, n=208), as a function of patient status on the prognostic index.

are supposed to emerge in regards to the intensity (as indexed by number of sessions) or specificity, they are likely to emerge among those patients who have a poorer prognosis (see DeRubeis et al. (2014b)). We chose to limit the number of predictor variables to those that have been previously related to treatment outcome, that had enough variability in our sample and were represented by other variables. In all likelihood, results would have been different had different variables (e.g., chronicity) been available.

There has been great interest in the use of patient characteristics to match patients to the treatments that might be most suitable to them, often in the context of treatment options of similar high intensity that are known to be approximately equally effective (e.g., CBT, IPT, and antidepressant medications; DeRubeis et al., 2014a; Huibers et al., 2015; Kang et al., 2014; Kraemer, 2013; Wallace et al., 2013). Somewhat less attention has been paid to the use of information about patient characteristics to assist in the matching of patients to the level of intensity of care that is appropriate to them (but, see Delgadillo et al., 2016). We have demonstrated how a statistical method could be used to achieve this aim. The first step is to develop a prognostic algorithm estimating the likelihood MDD recovery. We accomplished this by modeling MDD status at follow-up, ignoring treatment assignment. Our approach parallels the work of Huang et al. (2006) and Akay et al. (2012) who used the PI developed by Chen et al. (2005) to predict treatment differences. The similarity of this finding with our finding as well as those of Hollon et al. (2014), Nelson et al. (2013), and Thase et al. (1997) suggest that across areas of health care, prognostic status may serve to moderate the efficacy of higher versus lower intensity interventions (see DeRubeis et al., (2014b)). For most patients, there might be little if any advantage of engaging in the higher intensity treatments. Even so, a small and potentially identifiable subgroup of patients could experience large benefits from higher intensity treatments. An alternative approach to developing PIs, as suggested by Kessler et al. (2016a), is to develop risk models based on variables that are known from naturalistic studies to predict treatment response. This risk estimate could then be tested as a moderator of outcomes in a comparative clinical trial.

In actual clinical practice, a PI could inform decision algorithms to guide the level of care. In this application of a PI using patient information, factors such as the ease with which information can be obtained need to be considered. For example, symptom severity, unemployment status, and the presence of sleep problems are variables that are easier to ascertain relative to hostility or positive emotionality. There may be settings in which the practical application of PI or other decision-making tools requires that the more easily obtainable variables be used, either to the exclusion of the more difficult to ascertain variables or in conjunction with rough proxies. Variables that are considered to be trait level (e.g., positive emotionality) or that will not change (e.g., whether the patient has experienced depression in the past) can be collected and stored in electronic medical records (EMRs) which might facilitate the collection of state-level variables (e.g., severity, sleep problems).

Of the variables included in our index, symptom severity (Driessen et al., 2010; Fournier et al., 2010; Khan et al., 2002; Kirsch et al., 2008), unemployment status (Delgadillo et al., 2016; Jarrett et al., 2013; Rush et al., 2008; Trivedi et al., 2006), and sleep complaints (Andreescu et al., 2008; Dew et al., 1997; Troxel et al., 2011) have been directly implicated as predictors of outcome in depression. The various constructs captured by the measures of hostility and extraversion – high negative affect, low positive emotionality, difficulty in interpersonal relationships, and overall maladaptive personality traits— have also been reported to predict outcomes across various studies (see Kessler et al., 2016a). These variables share as commonalities that they capture vulnerability to depression, severity, and impairment associated with MDD. For those with a greater vulnerability to MDD, higher severity, and more impairment, lower intensity treatments like BT and TAU may not be sufficiently potent to produce sustained symptom change.

It is widely accepted that evidence-based psychotherapies for depression are equally efficacious (Barth et al., 2013; Cuijpers et al., 2014, 2008). However, insofar as comparisons of psychotherapies for depression have been conducted with populations that include a high proportion of patients who are expected to improve irrespective of treatment type, relative differences in the potencies of psychotherapies may have been obscured. Providing support for this conjecture, in several RCTs in which between-treatment differences were not found in the full sample, differences were identified in subsamples that comprised the more severely depressed patients (Dimidjian et al., 2006; Driessen et al., 2016; Elkin et al., 1995; Luty et al., 2007).

It is worth noting that the differences we reported emerged only in the follow-up phase of the study and were limited to a relatively small proportion of the sample. It is possible that differential effects of treatments are most evident in the long-term (Bell et al., 2013; Bockting et al., 2015; Cloitre et al., 2016) in that shorter term outcomes may index nonspecific effects of therapy. Longer term outcomes may reflect whether patients' underlying vulnerability to psychopathology was addressed or whether the patient acquired compensatory skills (Bell et al., 2013). In other instances in which significant differences have been reported in subgroups but not in full samples, the subgroups have tended to constitute a minority of the sample (e.g., 18% for Nelson et al., 2013).

Interventions of low-intensity, perhaps even less intensive than the

BT implemented in the present study, should be investigated further. Exercise (Lawlor and Hopker, 2001), unguided self-help (Cuijpers et al., 2011), and internet-based psychotherapies (Cuijpers et al., 2010) all are promising interventions in this regard. Although there is little evidence that these interventions are superior to high-intensity interventions for patients with mild to moderate MDD, they might be preferable in that they achieve similar outcomes with lower costs. In any effort to optimize care for depression, the consideration of individual patient features is likely to lead to improved outcomes.

We have illustrated the combination of several intake variables to construct a PI that can be used to assign patients to treatments of different intensities. As a proof of concept, our analyses suggest that most patients benefit from minimal interventions. However, for patients with a poorer prognosis, more intensive treatments are required to achieve positive outcomes. Given the difficulty in replicating prediction efforts, it is clear that more research, especially in the way of prospective tests, is needed. Although the recovery rate in patients with a poor prognosis was higher in CBT than in BT or TAU, there is still room for improvement. The judicious use of low-intensity interventions in stepped care conserves resources and thus facilities the use of higher intensity interventions (e.g., antidepressant therapy with lithium augmentation, electroconvulsive therapy) for those who most need them. Future research studies should explore depression predictors across the full range of depression treatments.

References

Akay, C.L., Meric-Bernstam, F., Hunt, K.K., Grubbs, E.G., Bedrosian, I., Tucker, S.L., Buchholz, T.A., 2012. Evaluation of the MD Anderson Prognostic Index for local-regional recurrence after breast conserving therapy in patients receiving neoadjuvant chemotherapy. Ann. Surg. Oncol. 19, 901–907. http://dx.doi.org/10.1245/s10434-011-2006-7.

American Psychiatric Association, 2000. Diagnostic and Statistical Manual Of Mental Disorders. American Psychiatric Association, Washington, DC, (revised 4th ed.).

Andreescu, C., Mulsant, B.H., Houck, P.R., Whyte, E.M., Mazumdar, S., Dombrovski, A.Y., Reynolds III, M., Charles, F., 2008. Empirically derived decision trees for the treatment of late-life depression. Am. J. Psychiatry 165, 855–862. http://dx.doi.org/10.1176/appi.ajp.2008.07081340.

Barbui, C., Cipriani, A., Patel, V., Ayuso-Mateos, J.L., van Ommeren, M., 2011. Efficacy of antidepressants and benzodiazepines in minor depression: systematic review and meta-analysis. Br. J. Psychiatry 198, 11–16. http://dx.doi.org/10.1192/bjp.bp.109.076448.

Barth, J., Munder, T., Gerger, H., Nüesch, E., Trelle, S., Znoj, H., Cuijpers, P., 2013. Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. PLoS Med. 10 (5), e1001454. http://dx.doi.org/10.1371/journal.pmed.1001454.

Bell, E.C., Marcus, D.K., Goodlad, J.K., 2013. Are the parts as good as the whole? A meta-analysis of component treatment studies. J. Consult. Clin. Psychol. 81, 722–736. http://dx.doi.org/10.1037/a0033004.

Bockting, C.L., Smid, N.H., Koeter, M.W., Spinhoven, P., Beck, A.T., Schene, A.H., 2015. Enduring effects of Preventive Cognitive Therapy in adults remitted from recurrent depression: A 10 year follow-up of a randomized controlled trial. J. Affect. Disord. 185, 188–194. http://dx.doi.org/10.1016/j.jad.2015.06.048.

Chen, A.M., Meric-Bernstam, F., Hunt, K.K., Thames, H.D., Outlaw, E.D., Strom, E.A., Singletary, S.E., 2005. Breast conservation after neoadjuvant chemotherapy. Cancer 103, 689–695. http://dx.doi.org/10.1002/cncr.20815.

Chen, A.M., Meric-Bernstam, F., Hunt, K.K., Thames, H.D., Oswald, M.J., Outlaw, E.D., Singletary, S.E., 2004. Breast conservation after neoadjuvant chemotherapy: the MD Anderson cancer center experience. J. Clin. Oncol. 22, 2303–2312. http://dx.doi.org/10.1200/JCO.2004.09.062.

Cloitre, M., Petkova, E., Su, Z., Weiss, B., 2016. Patient characteristics as a moderator of post-traumatic stress disorder treatment outcome: combining symptom burden and strengths. Br. J. Psychiatry Open 2, 101–106. http://dx.doi.org/10.1192/bjpo.bp.115.000745.

Costa, P.T., MacCrae, R.R., 1992. Revised NEO Personality Inventory (NEO PI-R) and NEO Five-factor Inventory (NEO FFI): Professional Manual. Psychological Assessment Resources, Odessa, FL.

Cuijpers, P., van Straten, A., Andersson, G., van Oppen, P., 2008. Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies. J. Consult. Clin. Psychol. 76, 909–922. http://dx.doi.org/10.1037/a0013075.

Cuijpers, P., Donker, T., van Straten, A., Li, J., Andersson, G., 2010. Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. Psychol. Med. 40, 1943–1957. http://dx.doi.org/10.1017/S0033291710000772.

Cuijpers, P., Donker, T., Johansson, R., Mohr, D.C., van Straten, A., Andersson, G., 2011. Self-guided psychological treatment for depressive symptoms: a meta-analysis. PLoS One 6, e21274. http://dx.doi.org/10.1371/journal.pone.0021274.

Cuijpers, P., Karyotaki, E., Weitz, E., Andersson, G., Hollon, S.D., van Straten, A., 2014. The effects of psychotherapies for major depression in adults on remission, recovery and improvement: a meta-analysis. J. Affect. Disord. 159, 118–126. http://dx.doi.org/10.1016/j.jad.2014.02.026.

Delgadillo, J., Moreea, O., Lutz, W., 2016. Different people respond differently to therapy: a demonstration using patient profiling and risk stratification. Behav. Res. Ther. 79, 15–22. http://dx.doi.org/10.1016/j.brat.2016.02.003.

Derogatis, L.R., 1983. Symptom Checklist-90-R Administration, Scoring and Procedures Manual II. Clinical Psychometric Research, Towson, MD.

DeRubeis, R.J., Gelfand, L.A., German, R.E., Fournier, J.C., Forand, N.R., 2014b. Understanding processes of change: how some patients reveal more than others – and some groups of therapists less – about what matters in psychotherapy. Psychother. Res. 24, 419–428. http://dx.doi.org/10.1080/10503307.2013.838654.

DeRubeis, R.J., Cohen, Z.D., Forand, N.R., Fournier, J.C., Gelfand, L.A., Lorenzo-Luaces, L., 2014a. The Personalized advantage index: translating research on prediction into individualized treatment recommendations. A demonstration. PLoS One 9, e83875. http://dx.doi.org/10.1371/journal.pone.0083875.

Dew, M.A., Reynolds, C.F., Houck, P.R., Hall, M., Buysse, D.J., Frank, E., Kupfer, D.J., 1997. Temporal profiles of the course of depression during treatment: predictors of pathways toward recovery in the elderly. Arch. Gen. Psychiatry 54, 1016–1024. http://dx.doi.org/10.1001/archpsyc.1997.01830230050007.

Dimidjian, S., Hollon, S.D., Dobson, K.S., Schmaling, K.B., Kohlenberg, R.J., Addis, M.E., Gollan, J.K., 2006. Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. J. Consult. Clin. Psychol. 74, 658–670. http://dx.doi.org/10.1037/0022-006X.74.4.658.

Driessen, E., Cuijpers, P., Hollon, S.D., Dekker, J.J., 2010. Does pretreatment severity moderate the efficacy of psychological treatment of adult outpatient depression? A meta-analysis. J. Consult. Clin. Psychol. 78, 668–680. http://dx.doi.org/10.1037/a0020570.

Driessen, E., Smits, N., Dekker, J., Peen, J., Don, F., Kool, S., Van, H., 2016. Differential efficacy of cognitive behavioral therapy and psychodynamic therapy for major depression: a study of prescriptive factors. Psychol. Med., 731–744. http://dx.doi.org/10.1017/S0033291715001853.

Elkin, I., Gibbons, R.D., Shea, M.T., Sotsky, S.M., Watkins, J.T., Pilkonis, P.A., Hedeker, D., 1995. Initial severity and differential treatment outcome in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. J. Consult. Clin. Psychol. 63, 841. http://dx.doi.org/10.1037/0022-006X.63.5.841.

Flückiger, C., Del Re, A.C., Munder, T., Heer, S., Wampold, B.E., 2014. Enduring effects of evidence-based psychotherapies in acute depression and anxiety disorders versus treatment as usual at follow-up—a longitudinal meta-analysis. Clin. Psychol. Rev. 34, 367–375. http://dx.doi.org/10.1016/j.cpr.2014.05.001.

Fournier, J.C., DeRubeis, R.J., Hollon, S.D., Dimidjian, S., Amsterdam, J.D., Shelton, R.C., Fawcett, J., 2010. Antidepressant drug effects and depression severity. JAMA 303, 47–53. http://dx.doi.org/10.1001/jama.2009.1943.

Garber, J., Weersing, V., Hollon, S., Porta, G., Clarke, G., Dickerson, J., Shamseddeen, W., 2016. Prevention of depression in at-risk adolescents: moderators of long-term response. Prev. Sci., 1–10. http://dx.doi.org/10.1007/s11121-015-0626-z.

Goldberg, D., Bridges, K., Duncan-Jones, P., Grayson, D., 1988. Detecting anxiety and depression in general medical settings. BMJ 297, 897–899. http://dx.doi.org/10.1136/bmj.297.6653.897.

Guo, P., Zeng, F., Hu, X., Zhang, D., Zhu, S., Deng, Y., Hao, Y., 2015. Improved variable selection algorithm using a LASSO-type penalty, with an application to assessing Hepatitis B infection relevant factors in community residents. PLoS ONE 10, e0134151. http://dx.doi.org/10.1371/journal.pone.0134151.

Harrell, F.E., Jr, 2001. Regression Modeling Strategies: with Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. Springer, New York, NY.

Hayes, A.F., Matthes, J., 2009. Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations. Behav. Res. Methods 41, 924–936. http://dx.doi.org/10.3758/BRM.41.3.924.

Hollon, S.D., DeRubeis, R.J., Fawcett, J., Amsterdam, J.D., Shelton, R.C., Zajecka, J., Gallop, R., 2014. Effect of cognitive therapy with antidepressant medications vs antidepressants alone on the rate of recovery in major depressive disorder: a randomized clinical trial. JAMA Psychiatry 71, 1157–1164. http://dx.doi.org/10.1001/jamapsychiatry.2014.

Huang, E.H., Strom, E.A., Perkins, G.H., Oh, J.L., Chen, A.M., Meric-Bernstam, F., Buchholz, T.A., 2006. Comparison of risk of local-regional recurrence after mastectomy or breast conservation therapy for patients treated with neoadjuvant chemotherapy and radiation stratified according to a prognostic index score. Int. J. Radiat. Oncol. Biol. Phys. 66, 352–357. http://dx.doi.org/10.1016/j.ijrobp.2006.04.046.

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, e0140771. http://dx.doi.org/10.1371/journal.pone.0140771.

Jarrett, R.B., Minhajuddin, A., Kangas, J.L., Friedman, E.S., Callan, J.A., Thase, M.E., 2013. Acute phase cognitive therapy for recurrent major depressive disorder: who drops out and how much do patient skills influence response? Behav. Res. Ther. 51, 221–230. http://dx.doi.org/10.1016/j.brat.2013.01.006.

Kang, C., Janes, H., Huang, Y., 2014. Combining biomarkers to optimize patient treatment recommendations. Biometrics 70, 695–707. http://dx.doi.org/10.1111/biom.12191.

Karyotaki, E., Smit, Y., Beurs, D.P., Henningsen, K.H., Robays, J., Huibers, M.J., Cuijpers, P., 2016. The long-term efficacy of acute-phase psychotherapy for

depression: a meta-analysis of randomized trials. Depress. Anxiety 33, 370–383. http://dx.doi.org/10.1002/da.22491.
Kessler, R., van Loo, H., Wardenaar, K., Bossarte, R., Brenner, L., Ebert, D., Sampson, N., 2016b. Using patient self-reports to study heterogeneity of treatment effects in major depressive disorder. Epidemiology *and* psychiatric sciences. Adv. Online Publ.. http://dx.doi.org/10.1017/S2045796016000020.
Kessler, R., van Loo, H., Wardenaar, K., Bossarte, R., Brenner, L., Cai, T., de Jonge, P., 2016a. Testing a machine-learning algorithm to predict the persistence and severity of major depressive disorder from baseline self-reports. Molecular psychiatry. Adv. Online Publ.. http://dx.doi.org/10.1038/mp.2015.198.
Khan, A., Leventhal, R.M., Khan, S.R., Brown, W.A., 2002. Severity of depression and response to antidepressants and placebo: an analysis of the Food and Drug Administration database. J. Clin. Psychopharmacol. 22, 40–45. http://dx.doi.org/10.1097/00004714-200202000-00007.
Khan, A., Faucett, J., Lichtenberg, P., Kirsch, I., Brown, W.A., 2012. A systematic review of comparative efficacy of treatments and controls for depression.. PloS ONE 7, e41778.
Kirsch, I., Deacon, B.J., Huedo-Medina, T.B., Scoboria, A., Moore, T.J., Johnson, B.T., 2008. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. PLoS Med. 5, e45. http://dx.doi.org/10.1371/journal.pmed.0050045.
Kirsch, I., Sapirstein, G., 1998. Listening to Prozac but hearing placebo: A meta-analysis of antidepressant medication. Prev. T. 2a.. http://dx.doi.org/10.1037/1522-3736.1.1.12a.
Kocsis, J.H., Gelenberg, A.J., Rothbaum, B., Klein, D.N., Trivedi, M.H., Manber, R., Thase, M.E., 2008. Chronic forms of major depression are still undertreated in the 21st century: systematic assessment of 801 patients presenting for treatment. J. Affect. Disord. 110, 55–61. http://dx.doi.org/10.1016/j.jad.2008.01.002.
Kraemer, H.C., 2013. Discovering, comparing, and combining moderators of treatment on outcome after randomized clinical trials: a parametric approach. Stat. Med. 32, 1964–1973. http://dx.doi.org/10.1002/sim.5734.
Lawlor, D.A., Hopker, S.W., 2001. The effectiveness of exercise as an intervention in the management of depression: systematic review and meta-regression analysis of randomised controlled trials. BMJ 322 (7289), 763. http://dx.doi.org/10.1136/bmj.322.7289.763.
Lecrubier, Y., 2007. Widespread underrecognition and undertreatment of anxiety and mood disorders: results from 3 European studies. J. Clin. Psychiatry 68 (suppl 2), (478-441).
Lorenzo-Luaces, L., 2015. Heterogeneity in the prognosis of major depression: from the common cold to a highly debilitating and recurrent illness. Epidemiol. Psychiatr. Sci. 24, 466–472. http://dx.doi.org/10.1017/S2045796015000542.
Lorenzo-Luaces, L., DeRubeis, R.J., Bennett, I.M., 2015. Primary care physicians' selection of low-intensity treatments for patients With depression. Fam. Med. 47, 511–516.
Lovell, K., Richards, D., 2000. Multiple access points and levels of entry (maple): ensuring choice, accessibility and equity for CBT services. Behav. Cogn. Psychother. 28, 379–391.
Luty, S.E., Carter, J.D., McKenzie, J.M., Rae, A.M., Frampton, C.M., Mulder, R.T., Joyce, P.R., 2007. Randomised controlled trial of interpersonal psychotherapy and cognitive–behavioural therapy for depression. Br. J. Psychiatry 190, 496–502. http://dx.doi.org/10.1192/bjp.bp.106.024729.
Morris, A.D., Morris, R.D., Wilson, J.F., White, J., Steinberg, S., Okunieff, P., Van Dongen, J.A., 1997. Breast-conserving therapy vs mastectomy in early-stage breast cancer: a meta-analysis of 10-year survival. Cancer J. Sci. Am. 3, 6–12. http://dx.doi.org/10.1016/j.cpr.2014.05.001.
National Institute for Clinical Excellence Depression, 2004. Management of Depression in Primary and Secondary Care. National Institute for Clinical Excellence.
Nelson, J.C., Delucchi, K.L., Schneider, L.S., 2013. Moderators of outcome in late-life depression: a patient-level meta-analysis. Am. J. Psychiatry 170, 651–659. http://dx.doi.org/10.1176/appi.ajp.2012.12070927.
Parker, G., 2005. Beyond major depression. Psychol. Med. 35, 467–474. http://dx.doi.org/10.1017/S0033291704004210.
Rush, A., Wisniewski, S.R., Warden, D., et al., 2008. Selecting among second-step antidepressant medication monotherapies: predictive value of clinical, demographic, or first-step treatment features. Arch. Gen. Psychiatry 65, 870–880. http://dx.doi.org/10.1001/archpsyc.65.8.870.
Stekhoven, D.J., Bühlmann, P., 2012. MissForest—non-parametric missing value imputation for mixed-type data. Bioinformatics 28, 112–118. http://dx.doi.org/10.1093/bioinformatics/btr597.
Thase, M.E., Greenhouse, J.B., Frank, E., Reynolds, C.F., Pilkonis, P.A., Hurley, K., Kupfer, D.J., 1997. Treatment of major depression with psychotherapy or psychotherapy-pharmacotherapy combinations. Arch. Gen. Psychiatry 54, 1009–1015. http://dx.doi.org/10.1001/archpsyc.1997.01830230043006.
Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B (Methodol.) 58, 267–288. http://dx.doi.org/10.1111/j.1467-9868.2011.00771.x.
Tibshirani, R., 2011. Regression shrinkage and selection via the lasso: a retrospective. J. R. Stat. Soc.: Ser. B (Stat. Methodol.) 73, 273–282. http://dx.doi.org/10.1111/j.1467-9868.2011.00771.x.
Tiemens, B.G., 1999. Management of mental health problems in primary care. The doctor, the Patient, and the Medical Model. University of Groningen.
Trivedi, M.H., Rush, A.J., Wisniewski, S.R., Nierenberg, A.A., Warden, D., Ritz, L., McGrath, P.J., 2006. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR* D: implications for clinical practice. Am. J. Psychiatry. http://dx.doi.org/10.1176/appi.ajp.163.1.28.
Troxel, W.M., Kupfer, D.J., Reynolds, C.F., III, Frank, E., Thase, M.E., Miewald, J.M., Buysse, D.J., 2011. Insomnia and objectively measured sleep disturbances predict treatment outcome in depressed patients treated with psychotherapy or psychotherapy-pharmacotherapy combinations. J. Clin. Psychiatry 73, 478–485. http://dx.doi.org/10.4088/JCP.11m07184.
van Dongen, J.A., Voogd, A.C., Fentiman, I.S., Legrand, C., Sylvester, R.J., Tong, D., Bartelink, H., 2000. Long-term results of a randomized trial comparing breast-conserving therapy with mastectomy: European Organization for Research and Treatment of Cancer 10801 trial. J. Natl. Cancer I. 92, 1143–1150. http://dx.doi.org/10.1093/jnci/92.14.1143.
Van Straten, A., Tiemens, B., Hakkaart, L., Nolen, W., Donker, M., 2006. Stepped care vs. matched care for mood and anxiety disorders: a randomized trial in routine practice. Acta Psychiatr. Scand. 113, 468–476. http://dx.doi.org/10.1111/j.1600-0447.2005.00731.x.
Waljee, A.K., Mukherjee, A., Singal, A.G., Zhang, Y., Warren, J., Balis, U., Higgins, P.D., 2013. Comparison of imputation methods for missing laboratory data in medicine. BMJ OPEN 3, e002847. http://dx.doi.org/10.1136/bmjopen-2013-002847.
Wallace, M.L., Frank, E., Kraemer, H.C., 2013. A novel approach for developing and interpreting treatment moderator profiles in randomized clinical trials. JAMA Psychiatry 70, 1241–1247. http://dx.doi.org/10.1001/jamapsychiatry.2013.1960.
Wampold, B.E., Budge, S.L., Laska, K.M., Del Re, A.C., Baardseth, T.P., Flückiger, C., Gunn, W., 2011. Evidence-based treatments for depression and anxiety versus treatment-as-usual: A meta-analysis of direct comparisons. Clin. Psychol. Rev. 31 (8), 1304–1312. http://dx.doi.org/10.1016/j.cpr.2011.07.012.
Wittchen, H.-U., Robins, L.N., Semler, W.J., Cottler, L., World Health Organization, 1993. Composite International Diagnostic Interview (CIDI): interviewer's Manual. World Health Organization, Geneva, Switzerland.