Differential efficacy of cognitive behavioral therapy and psychodynamic therapy for major depression

A study of prescriptive factors

Driessen, E.; Smits, N.; Dekker, J.J.M.; Peen, J.; Don, F.J.; Kool, S.; Westra, D.; Hendriksen, M.; Cuijpers, P.; Van, H.L.

DOI
10.1017/S0033291715001853

Publication date
2016

Document Version
Final published version

Published in
Psychological Medicine

Psychological Medicine (2016), **46**, 731–744. © Cambridge University Press 2016
doi:10.1017/S0033291715001853

ORIGINAL ARTICLE

Differential efficacy of cognitive behavioral therapy and psychodynamic therapy for major depression: a study of prescriptive factors

E. Driessen[1,2]*, N. Smits[1,2], J. J. M. Dekker[1,2,3], J. Peen[3], F. J. Don[3,4], S. Kool[3], D. Westra[3], M. Hendriksen[3], P. Cuijpers[1,2] and H. L. Van[3]

[1] *Department of Clinical Psychology, VU University Amsterdam, Amsterdam, The Netherlands*
[2] *EMGO Institute for Health and Care Research, VU University and VU University Medical Center Amsterdam, The Netherlands*
[3] *Arkin Mental Health Care, Amsterdam, The Netherlands*
[4] *ProPersona Mental Health, Nijmegen, The Netherlands*

Background. Minimal efficacy differences have been found between cognitive behavioral therapy (CBT) and psychodynamic therapies for depression, but little is known about patient characteristics that might moderate differential treatment effects. We aimed to generate hypotheses regarding such potential prescriptive factors.

Method. We conducted *post-hoc* model-based recursive partitioning analyses alongside a randomized clinical trial comparing the efficacy of CBT and short-term psychodynamic supportive psychotherapy (SPSP). Severely depressed patients received additional antidepressant medication. We included 233 adults seeking treatment for a major depressive episode in psychiatric outpatient clinics, who completed post-treatment assessment. Post-treatment mean Hamilton Depression Rating Scale scores constituted the main outcome measure.

Results. While treatment differences (CBT *v.* SPSP) were minimal in the total sample of patients ($d = 0.04$), model-based recursive partitioning indicated differential treatment efficacy in certain subgroups of patients. SPSP was found more efficacious among moderately depressed patients receiving psychotherapy only who showed low baseline co-morbid anxiety levels ($d = -0.40$) and among severely depressed patients receiving psychotherapy and antidepressant medication who reported a duration of the depressive episode of $\geqslant 1$ year ($d = -0.31$), while CBT was found more efficacious for such patients reporting a duration <1 year ($d = 0.83$).

Conclusions. Our findings are observational and need validation before they can be used to guide treatment selection, but suggest that knowledge of prescriptive factors can help improve the efficacy of psychotherapy for depression. Depressive episode duration and co-morbid anxiety level should be included as stratification variables in future randomized clinical trials comparing CBT and psychodynamic therapy.

Received 18 February 2014; Revised 19 August 2015; Accepted 27 August 2015; First published online 11 January 2016

Key words: Depression, prescriptive factors, psychotherapies.

Introduction

While psychological treatments for depression have different theoretical backgrounds, studies of comparative efficacy have resulted in minimal efficacy differences (Cuijpers *et al.* 2008; Barth *et al.* 2013). Cognitive behavioral therapy (CBT), for example, aims at alleviating depressive symptoms by changing maladaptive thought schemata and errors in thinking in combination with engaging more in activities that affect mood positively, while psychodynamic therapies assume that gaining insight in (partly) unconscious emotions and relational functioning related to vulnerability for depression is curative. Notwithstanding these markedly different theoretical backgrounds, minimal differences have been found between CBT and psychodynamic therapies with regard depressive symptom reduction during short-term treatment (Leichsenring, 2001; Barth *et al.* 2013).

If minimal efficacy differences are found when CBT and psychodynamic therapies are compared across larger patient samples, the question can be raised whether smaller subgroups of patients can be identified that might benefit more from one of these treatments than the other and whether such patient characteristics can be used to guide treatment selection.

* Address for correspondence: Dr E. Driessen, Faculty of Psychology and Education, Department of Clinical Psychology, VU University Amsterdam, Transitorium 2B-57, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands.
(Email: e.driessen@vu.nl)

Little is known about prescriptive factors that are associated with differential efficacy to CBT and psychodynamic therapy for depression. We could retrieve only one study in this regard. Gallagher-Thompson & Steffen (1994) found that clinically depressed family caregivers who had been caregivers for less than 44 months improved more in psychodynamic therapy than in CBT, while patients who had been caregivers for at least 44 months improved more in CBT than in psychodynamic therapy.

Given the lack of research findings in this regard, the National Institute for Health and Clinical Excellence (NICE, 2009, p. 46) called for the examination of moderators of response to CBT and psychodynamic therapy in the treatment of moderate and severe depression as a research recommendation in order to improve patient care. Considering the paucity of findings, we aimed to conduct a hypothesis-generating (rather than a hypothesis-testing; Kraemer *et al.* 2002) study by means of conducting *post-hoc* analyses alongside a randomized clinical trial (RCT) comparing CBT and psychodynamic therapy in the outpatient treatment of depression (Driessen *et al.* 2007). Hypothesis-generating studies are often dismissed as 'fishing expeditions', but as argued by Kraemer *et al.* (2002, p. 882) such studies 'are necessary to foster stronger hypotheses for the next generation of hypothesis-testing studies and to provide the background information necessary to design such powerful studies'. Kraemer *et al.* (2002, p. 883), therefore, state that 'RCTs should routinely include and report such analyses'. Given the hypothesis-generating nature of our study, we chose to examine a large set of demographic, clinical and psychological patient characteristics as potential prescriptive variables.

Method

Design

This paper draws from data collected in the context of a RCT comparing the efficacy of CBT and psychodynamic therapy in the outpatient treatment of depression that included 341 patients. This intervention study was registered as ISRCTN31263312 with Current Controlled Trials (http://www.controlled-trials.com). The Dutch Union of Medical-Ethic Trial Committees for mental health organizations approved the study design and the study protocol was published (Driessen *et al.* 2007). Efficacy results of this study are reported elsewhere (Driessen *et al.* 2013, 2015); here we report an effort to identify prescriptive factors associated with differential treatment efficacy. Although these analyses are *post hoc*, the decision to conduct them was made *a priori* (Driessen *et al.* 2007).

Patients

Participants were referred by their general practitioner to one of three psychiatric outpatient clinics in Amsterdam, The Netherlands. Inclusion criteria were: (1) presence of a depressive episode according to DSM-IV criteria as assessed with the MINI-International Neuropsychiatric Interview – Plus (Sheehan *et al.* 1998); (2) Hamilton Depression Rating Scale (HAMD; Hamilton, 1960) scores $\geqslant 14$; (3) age between 18 and 65 years; (4) written informed consent after complete description of the study. Exclusion criteria are described elsewhere (Driessen *et al.* 2007).

Interventions

Both psychotherapies encompassed 16 individual sessions within 22 weeks and were conducted according to published treatment manuals (de Jonghe, 2005; Molenaar *et al.* 2009). CBT was based on the principles described by Beck (1976) and included behavioral activation and cognitive restructuring according to a session-by-session protocol with homework assignments. Short-term psychodynamic supportive psychotherapy (SPSP; de Jonghe *et al.* 2013) was used to represent the psychodynamic intervention. This modality involved an open patient–therapist dialog that used supportive and insight-facilitating techniques to address the emotional background of the depressive symptoms by discussing current relationships, internalized past relationships, and interpersonal patterns. Psychotherapists in both conditions were trained psychiatrists or psychologists. Although no formal assessments were conducted, manual fidelity was checked by means of bi-weekly supervision sessions, chaired by a study supervisor, in which audio-taped material was discussed.

Severely depressed (HAMD >24) patients at baseline (CBT: $n = 66$; SPSP: $n = 63$) and moderately depressed patients at baseline who developed severe symptoms during treatment (CBT: $n = 6$; SPSP: $n = 15$) were offered additional antidepressant medication. The number of patients starting pharmacotherapy at baseline or during psychotherapy did not differ significantly between the treatment conditions. Antidepressant medication was administrated by a psychiatrist (who was not the patient's psychotherapist) according to a protocol starting with extended-release venlafaxine 75 mg/day that could be raised to a maximum of 225 mg/day and switched to either citalopram or nortriptyline in case of intolerance or complete non-response. Pharmacotherapy consultations addressed symptom evaluation, side-effects and adherence.

Instruments

Post-treatment remission rate (HAMD ⩽7) constituted the RCT's primary outcome (Driessen *et al.* 2007). However, as the rather low remission rates found (Driessen *et al.* 2013) would likely result in empty cells when examining subgroups of patients, thereby compromising logistic analyses, we decided to use continuous post-treatment HAMD scores as primary outcome measure in this study. HAMD assessors were trained master-level clinical psychology students not blind to patient grouping, who assessed the HAMD according to the Dutch scoring manual (de Jonghe, 1994). Assessors engaged in 1-h peer supervision sessions bi-weekly, in which audio-taped interviews were discussed. The average intraclass correlation coefficient over 46 audio-taped assessments scored by multiple assessors was 0.97.

We aimed at examining a large set of potential prescriptive factors and we therefore wanted to examine all moderator variables included in the design of the RCT (for a description see Driessen *et al.* 2007). However, we were faced with missing data on a number of these variables due to participants not returning questionnaire booklets. We decided to examine a variable as a potential prescriptive factor only if ⩽10% of the cases were missing. This resulted in a total of 23 variables (12 demographic, nine clinical, two psychological) to be considered as prescriptive factors. These variables are written in italics below as well as listed in Table 1 with the classification options for categorical variables. All variables were assessed at baseline.

With regard to the 12 demographic characteristics, *age* and *gender* were extracted from the mental health clinic registration system, while a self-report demographic questionnaire designed by the investigators was used to assess *nationality, cultural background, marital status, living situation, religion, educational level, current profession, job status, breadwinnership, and breadwinner's main source of income*. Cultural background was included in this study, because we considered it a potential prescriptive factor of differential treatment efficacy (Ward, 2007). Participants classified themselves with regard to cultural background, with options (see Table 1) being defined by the investigators.

With regard to the nine clinical characteristics, we used the Beck Anxiety Inventory (Beck *et al.* 1988) total score to assess *co-morbid anxiety symptoms*, the Brief Symptom Inventory (de Beurs & Zitman, 2006) total score to assess (*co-morbid*) *general psychopathology*, and four items of the Trimbos/iMTA questionnaire for costs associated with psychiatric illness (Hakkaart-van Roijen, 2002) to assess healthcare use (*medication use* and *number of consultations with general practitioner, outpatient mental health care, and industrial medical officer in the last 4 weeks*). *Duration of the current depressive episode, previous treatment for the current depressive episode,* and *lifetime number of depressive episodes* were self-reported by the patient during the baseline clinical interview.

The two psychological characteristics included *cognitive reactivity to sad mood* (the relative ease with which maladaptive cognitions or cognitive styles are triggered by mild mood fluctuations) and *anxiety sensitivity* (a person's beliefs that anxiety experiences have negative somatic, psychological or social consequences), which were measured by total scores of, respectively, the Leiden Index of Depression Sensitivity (van der Does, 2002) and the Anxiety Sensitivity Index (Reiss *et al.* 1986).

Statistical analysis

Rather than using the intention-to-treat sample and imputing missing post-treatment HAMD scores, we examined potential prescriptive factors in the group of patients who completed post-treatment HAMD assessment. We chose to do so because loss to HAMD assessment was significantly associated with missing baseline data for a number of variables, such that for the patients who were lost to assessment and whose HAMD value was to be imputed, data on which the imputation should be based was missing as well and imputation would likely result in unreliable post-treatment HAMD values. We were mainly interested in the differential efficacy between CBT and SPSP, and therefore included all abovementioned 23 potential prescriptive pre-treatment factors in the analyses. In addition, we wanted to take into account whether or not patients received additional antidepressant medication. Therefore, we used this variable as an extra prescriptive factor. As the absence of treatment differences on the prescriptive factors is a precondition for estimating differential efficacy, we checked for treatment differences on all potential prescriptive factors in the group of patients who completed post-treatment HAMD assessment in an initial series of analyses.

In clinical trials, prescriptive factors are often examined by means of stepwise regression analysis, adding a main effect and a predictor×treatment interaction effect into the model for each potential prescriptive factor (Kraemer *et al.* 2002; for further explanation of the method see Aiken & West, 1991 and Cohen *et al.* 2013). However, the power of finding significant effects in such a model with a multitude of parameters may be low in a sample of modest size, the resulting model can be quite complex making it hard to interpret, and stepwise regression does not allow for modeling nonlinear relationships. A new method that can also be used in this regard, but does not suffer from

Table 1. *Baseline characteristics of the study sample*

Variable	Total sample ($n = 233$)		CBT ($n = 111$)		SPSP ($n = 122$)		Test statistic (df)	p
Additional antidepressant medication? (n, %)							$\chi^2(1) = 0.28$	0.34
Psychotherapy only (HAMD $\leqslant$24)	151	64.8	68	61.3	83	68.0		
Combined treatment (HAMD >24)	82	35.2	43	38.7	39	32.0		
Pretreatment HAMD score (mean, S.D.)	23.03	5.20	23.17	5.04	22.90	5.35	$t(231) = 0.39$	0.69
Potential prescriptive factors								
Demographic characteristics								
Age, years (mean, S.D.)	39.98	10.33	39.46	10.16	40.46	10.50	$t(231) = -0.74$	0.46
Gender (n, %)							$\chi^2(1) = 0.04$	0.89
Male	75	32.2	35	31.5	40	32.8		
Female	158	67.8	76	68.5	82	67.2		
Nationality (n, %)							$\chi^2(1) = 0.05$	1.00
Dutch	208	91.2	99	91.7	109	90.8		
Non-Dutch	20	8.8	9	8.3	11	9.2		
Cultural background (n, %)							Fisher's exact = 3.00	0.95
North-west European	140	60.1	68	61.3	72	59.0		
South European	2	0.9	1	0.9	1	0.8		
Caribbean	27	11.6	11	9.9	16	13.1		
North African	1	0.4	1	0.9	0	0.0		
Moroccan	23	9.9	12	10.8	11	9.0		
Turkish	18	7.7	7	6.3	11	9.0		
Asian	5	2.1	3	2.7	2	1.6		
Other	17	7.3	8	7.2	9	7.4		
Marital status (n, %)							Fisher's exact = 9.66	0.02
Married	62	26.6	38	34.2	24	19.7		
Divorced	46	19.7	24	21.6	22	18.0		
Widowed	6	2.6	1	0.9	5	4.1		
Never married	119	51.1	48	43.2	71	58.2		
Living situation (n, %)							Fisher's exact = 2.11	0.38
Living with at least one other person	147	63.4	74	66.7	73	60.3		
Living alone	77	33.2	35	31.5	42	34.7		
Other	8	3.4	2	1.8	6	5.0		
Religion (n, %)							Fisher's exact = 1.94	1.00
Christian	53	24.4	26	24.5	27	24.3		
Muslim	43	19.8	22	20.8	21	18.9		
Hindu/Buddhist	2	0.9	1	0.9	1	0.9		
Other	45	20.7	22	20.8	23	20.7		
Educational level (n, %)							$\chi^2(2) = 2.04$	0.38
Low	49	21.2	23	20.9	26	21.5		
Intermediate	102	44.2	44	40.0	58	47.9		
High	80	34.6	43	39.1	37	30.6		
Current profession (n, %)							Fisher's exact = 6.50	0.49
No profession	43	19.9	19	18.6	24	21.1		
Houseman/housewife	34	15.7	17	16.7	17	14.9		
Unskilled labor	23	10.6	9	8.8	14	12.3		
Skilled labor	61	28.2	26	25.5	35	30.7		
Low-level white-collar work	6	2.8	4	3.9	2	1.8		
Small company entrepreneur	14	6.5	5	4.9	9	7.9		
Middle-level white collar executive/high-level white-collar non-executive	32	14.8	20	19.6	12	10.5		

Table 1 (*cont.*)

Variable	Total sample ($n = 233$)		CBT ($n = 111$)		SPSP ($n = 122$)		Test statistic (df)	p
Large company entrepreneur/high-level white-collar executive	3	1.4	2	2.0	1	1.6		
Job status (n, %)							Fisher's exact = 5.10	0.41
Currently working	92	40.0	42	38.2	50	41.7		
Sickness benefits	41	17.8	25	22.7	16	13.3		
Social security benefits	41	17.8	17	15.5	24	20.0		
Disability benefits	25	10.9	12	10.9	13	10.8		
Student	10	4.3	3	2.7	7	5.2		
Other	21	9.1	11	10.0	10	8.3		
Family breadwinner (n, %)							Fisher's exact = 1.74	0.83
Yes	153	69.5	70	67.3	83	71.6		
No: partner is breadwinner	48	21.8	23	22.1	25	21.6		
No: parent is breadwinner	4	1.8	3	2.9	1	0.9		
No: two-earner household	11	5.0	6	5.8	5	4.3		
Other	4	1.8	2	1.9	2	1.7		
Breadwinner's main source of income (n, %)							Fisher's exact = 5.20	0.53
Salary	141	65.9	74	71.2	67	60.9		
Pension	2	0.9	1	1.0	1	0.9		
Welfare	26	12.1	13	12.5	13	11.8		
Disability benefit	15	7.0	6	5.8	9	8.2		
Unemployment benefit	10	4.7	4	3.8	6	5.5		
Study subsidy	2	0.9	1	1.0	1	0.9		
Other	18	8.4	5	4.8	13	11.8		
Clinical characteristics								
Duration present episode (n, %)							Fisher's exact = 8.25	0.08
<6 months	58	24.9	32	28.8	26	21.3		
6 months to 1 year	57	24.5	30	27.0	27	22.1		
1–2 years	32	13.7	18	16.2	14	11.5		
>2 years	78	33.5	27	24.3	51	41.8		
Unknown	8	3.4	4	3.6	4	3.3		
Prior treatment for current depressive episode (n, %)							$\chi^2(1) = 1.49$	0.27
No	150	64.4	67	60.4	83	68.0		
Yes	83	35.6	44	39.6	39	32.0		
Number of prior depressive episodes (n, %)							$\chi^2(2) = 0.68$	0.73
0	70	30.3	36	32.7	34	28.1		
1	41	17.7	18	16.4	23	19.0		
⩾2	120	51.9	56	50.9	64	52.9		
Medication use (n, %)							$\chi^2(1) = 3.15$	0.08
Yes	110	50.5	58	56.9	52	44.8		
No	108	49.5	44	43.1	64	55.2		
Co-morbid anxiety (Beck Anxiety Inventory; mean, S.D.)	23.59	12.97	23.84	13.07	23.37	12.93	$t(218) = 0.27$	0.79
Co-morbid psychopathology (Brief Symptom Inventory; mean, S.D.)	1.83	0.76	1.89	0.78	1.78	0.74	$t(193) = 1.10$	0.27
Consultations with general practitioner in the last 4 weeks (mean, S.D.)	0.87	1.19	0.85	1.01	0.88	1.34	$t(221) = -0.20$	0.85
Consultations with mental health care in the last 4 weeks (mean, S.D.)	0.75	1.12	0.71	1.04	0.78	1.20	$t(221) = -0.43$	0.67
Consultations with industrial medical officer in the last 4 weeks (mean, S.D.)	0.39	0.72	0.48	0.78	0.31	0.65	$t(202) = 1.75$	0.08

Table 1 (*cont.*)

Variable	Total sample ($n = 233$)		CBT ($n = 111$)		SPSP ($n = 122$)		Test statistic (df)	p
Psychological characteristics								
Cognitive reactivity to sad mood (LEIDS; mean S.D.)	54.55	16.62	54.26	15.94	54.80	17.26	$t(219) = 0.03$	0.98
Anxiety sensitivity (Anxiety Sensitivity Index; mean, S.D.)	36.83	13.58	36.86	13.09	36.81	14.07	$t(212) = -0.24$	0.81

CBT, Cognitive behavioral therapy, HAMD, Hamilton Depression Rating Scale; LEIDS, Leiden Index of Depression Sensitivity; S.D., standard deviation; SPSP, short-term psychodynamic supportive psychotherapy.

these issues is model-based recursive partitioning (MOB; Zeileis *et al.* 2008, 2012; Strobl *et al.* 2009). MOB is a method for finding subgroups for which a specified basic model has different parameters. In the case of our study, this basic model is an ANOVA model including post-treatment HAMD score as dependent variable and a dichotomous treatment factor (CBT or SPSP) as independent variable. This ANOVA model has two parameters: an intercept (equaling the average post-treatment HAMD score of all subjects) and a treatment parameter (equaling the average difference in post-treatment HAMD score between the treatment conditions). This latter parameter is considered the most interesting as it reflects differential treatment efficacy.

MOB uses potential prescriptive variables (e.g. demographic characteristics) to identify groups of patients for which the ANOVA model parameters differ. If the treatment parameter differs between subgroups, this means that the average difference between the two treatments is different for these subgroups. Consequently, the predictor(s) defining these groups can be considered prescriptive factors. For example, if MOB would show a positive treatment parameter for males and a negative one for females, this would mean that, on average, the one treatment gives better outcomes for males, whereas the other treatment gives better outcomes for females and gender could be considered a prescriptive factor.

To identify these subgroups, a binary tree is used in MOB. In the first node of the tree, all observations are used to estimate the basic ANOVA model. Next, for each potential prescriptive variable, the hypothesis that the basic model's parameters are stable over the variable's (ordered or nominal) values is tested, which results in a p value. Next, it is checked if the lowest p value is lower than some pre-specified α, and if so, the predictor associated with this minimal value is chosen to split up the observations into two subgroups (child nodes) using the cut-score which shows the highest structural change (i.e. instability). Next, in both the left and the right child node, the subgroup's data are used to refit the ANOVA model, and the abovementioned procedure is repeated. The process of splitting nodes into child nodes is continued until the lowest p value fails to be significant.

In this study, MOB was conducted in R (R Development Core Team, 2010) using the party library (Hothorn *et al.* 2012). As described above, we chose an ANOVA model as the basic model. We decided not to include pre-treatment HAMD score as a covariate in this basic model, because pre-treatment HAMD score was used as a decision variable to add antidepressant medication to psychotherapy. Adding it as a covariate would be problematic, as this would result in a spurious relationship between the medication factor and the covariate's regression parameter in the basic model. We changed party's default value for minimal observations in the end nodes from 20 to 30 (about 15 in each therapy group), which we deemed the minimal for estimating the ANOVA. We used an α of 10%, because of the theory-generating nature of our study (Kraemer *et al.* 2002, pp. 881–882; Stevens, 2002) and because sample size was reduced by the exclusion of patients not completing post-treatment HAMD assessment and might be further decreased by missing values among the prescriptive variables.

In standard MOB, all predictors are entered in the model simultaneously, but we preferred a forward selection approach for two reasons. First, MOB as implemented in R's party library uses complete observations on all variables entered. Although the algorithm usually only selects a small subset of the potential predictors, observations with any missing value on the full set of variables are removed before the actual analysis. As missing value patterns varied over variables, entering all variables simultaneously would result in a large number of excluded cases, while this was not the case with the forward selection approach. Second, patients with a baseline HAMD scores >24 received additional antidepressant medication. Because all other predictors should be nested under this design factor, it should be in the first node of the binary tree. The only way to incorporate this into the model is to force this factor as

the first predictor in the forward approach, and then expand the model with additional predictors in each of the child nodes. The forward selection method was performed manually by repeating the analyses and keeping track of the p value of each moderator. Our data-analytic strategy bares a strong resemblance with the methods reported by Dusseldorp & Meulman (2004, Dusseldorp & Van Mechelen, 2014), which have been applied in a study of prescriptive factors associated with differential efficacy of cognitive therapy and antidepressants for panic disorder patients (Dusseldorp *et al.* 2007), but the current method can be considered a more general approach.

Ethical standards

The authors assert that all procedures contributing to this work comply with the ethical standards of relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Results

Patient sample

Of the 341 patients randomized, 233 (68.3%) completed post-treatment HAMD assessment and no significant differences were found between treatment conditions with regard to the number of patients lost to this assessment [CBT: 32.3%; SPSP: 31.1%; $\chi^2(1) = 0.06$, $p = 0.81$]. The group of patients who completed post-treatment HAMD assessment is described in Table 1 with regard to all potential prescriptive variables examined. With the exception of marital status, no significant differences between conditions were found. In Table 2, the baseline characteristics of the patients who completed post-treatment HAMD assessment are compared with those of the patients who were lost to this assessment. It can be seen from this Table that significant differences were found between these two groups on a large number of baseline characteristics.

Model-based recursive partitioning analysis

The results of the model-based recursive partitioning analysis are represented in Fig. 1 and treatment outcomes in each node of the tree are listed in Table 3. In Table 3, the standardized mean difference between the treatments (CBT *v.* SPSP) is expressed as Cohen's d (Cohen, 1988). A positive value indicates that CBT resulted in a better outcome (i.e. a lower post-treatment HAMD score) than SPSP, while a negative value indicates that SPSP resulted in a better outcome than CBT.

Node 1 includes all patients and treatment differences in this group are minimal ($d = 0.04$; Table 3). Node 1's p value ($p = 0.004$; Fig. 1) indicates that the parameters of the basic model are significantly different for the two subgroups in which this node is divided: left-hand node 2 representing the subgroup of patients with baseline HAMD scores $\leqslant 24$ who received psychotherapy only and right-hand node 3 representing the subgroup of patients with baseline HAMD scores >24 who received psychotherapy and antidepressant medication. While minimal treatment differences ($d = -0.07$; Table 3) are found in the first group, larger treatment differences ($d = 0.36$) are apparent in the latter group.

Left-hand node 2 represents the low-severity patients receiving psychotherapy only and its p value ($p = 0.051$; Fig. 1) indicates that the parameters of the basic model are significantly different for the two subgroups in which this node is divided: node 4 including patients with baseline co-morbid anxiety scores of $\leqslant 15$ and node 5 including patients with co-morbid anxiety scores >15. In the first, post-treatment depression scores are lower in the SPSP condition than in the CBT condition ($d = -0.40$; Table 3), while in the latter post-treatment depression scores differ minimally ($d = 0.04$).

Right-hand node 3 represents the group of high-severity patients who received psychotherapy combined with antidepressant medication and its p value ($p = 0.032$; Fig. 1) indicates that the parameters of the basic model are significantly different for the two subgroups in which this node is divided: node 6 including patients with a duration of the current depressive episode <12 months and node 7 including patients who report an episode duration of $\geqslant 12$ months. In the first, post-treatment depression scores are lower in the CBT condition ($d = 0.83$; Table 3), while in the latter post-treatment depression scores are lower in the SPSP condition ($d = -0.31$).

Sensitivity analyses

As pre-treatment HAMD score was not included as a covariate in the basic model, one might argue that the post-treatment differences in HAMD scores between conditions found in nodes 4, 6, and 7 might be the consequence of chance differences in pre-treatment HAMD scores between the treatment conditions within these nodes. Such that, for instance, the moderately severe patients with low anxiety levels in the SPSP condition in node 4 coincidently had lower pre-treatment HAMD scores than these patients in the CBT condition in this node and that this chance finding resulted in the lower post-treatment HAMD scores found in the SPSP condition. We, therefore,

Table 2. *Differences between patients completing post-treatment HAMD assessment and patients not completing this assessment, who were excluded from the analyses*

Variable	Total sample ($n = 341$)		Completers ($n = 233$)		Non-completers ($n = 108$)		Test statistic (df)	P
Additional antidepressant medication? (n, %)								
Psychotherapy only (HAMD $\leqslant$24)	212	62.2	151	64.8	61	56.5	$\chi^2(1) = 2.18$	0.15
Combined treatment (HAMD >24)	129	37.8	82	35.2	47	43.2		
Pretreatment HAMD score (mean, S.D.)	23.40	5.35	23.03	5.20	24.19	5.59	$t(339) = -1.88$	0.06
Demographic characteristics								
Age, years (mean, S.D.)	38.91	10.29	39.98	10.33	36.58	9.88	$t(339) = 2.87$	<0.01
Gender (n, %)							$\chi^2(1) = 1.82$	0.20
Male	102	29.9	75	32.2	27	25.0		
Female	239	70.1	158	67.8	81	75.0		
Nationality (n, %)							$\chi^2(1) = 4.70$	0.04
Dutch	291	88.7	208	91.2	83	83.0		
Non-Dutch	37	11.3	20	8.8	17	17.0		
Cultural background (n, %)							Fisher's exact = 22.15	<0.01
North-west European	186	55.0	140	60.1	46	43.8		
South European	2	0.6	2	0.9	0	0.0		
Caribbean	46	13.6	27	11.6	19	18.1		
North African	6	1.8	1	0.4	5	4.8		
Moroccan	46	13.6	23	9.9	23	21.9		
Turkish	22	6.5	18	7.7	4	3.8		
Asian	7	2.1	5	2.1	2	1.9		
Other	23	6.8	17	7.3	6	5.7		
Marital status (n, %)							Fisher's exact = 9.23	0.04
Married	80	23.7	62	26.6	18	17.1		
Divorced	69	20.4	46	19.7	23	21.9		
Widowed	10	3.0	6	2.6	4	3.8		
Never married	176	52.1	119	51.1	57	54.3		
Other	3	0.9	0	0.0	3	2.9		
Living situation (n, %)							Fisher's exact = 1.17	0.58
Living with at least one other person	220	65.3	147	63.4	73	69.5		
Living alone	106	31.5	77	33.2	29	27.6		
Other	11	3.3	8	3.4	3	2.9		
Religion (n, %)							Fisher's exact = 12.34	0.10
Christian	76	25.5	53	24.4	23	28.4		
Muslim	66	22.1	43	19.8	23	28.4		
Hindu/Buddhist	6	2.0	2	0.9	4	4.9		
Other	60	20.1	45	20.7	15	18.3		

Educational level (*n*, %)							$\chi^2(2) = 6.01$	0.05
Low	71	21.2	49	21.2	22	21.2		
Intermediate	161	48.1	102	44.2	59	56.7		
High	103	30.7	80	34.6	23	22.1		
Current profession (*n*, %)							Fisher's exact = 16.10	0.02
No profession	73	24.3	43	19.9	30	35.3		
Houseman/housewife	39	13.0	34	15.7	5	5.9		
Unskilled labor	35	11.6	23	10.6	12	14.1		
Skilled labor	87	28.9	61	28.2	26	30.6		
Low-level white-collar work	7	2.3	6	2.8	1	1.2		
Small company entrepreneur	16	5.3	14	6.5	2	2.4		
Middle-level white-collar executive/ high-level white collar non-executive	39	13.0	32	14.8	7	8.2		
Large company entrepreneur/high-level white-collar executive	5	1.7	3	1.4	2	2.4		
Job status (*n*, %)							Fisher's exact = 8.16	0.14
Currently working	130	38.8	92	40.0	38	36.2		
Sickness benefits	55	16.4	41	17.8	14	13.3		
Social security benefits	74	22.1	41	17.8	33	31.4		
Disability benefits	32	9.6	25	10.9	7	6.7		
Student	14	4.2	10	4.3	4	3.8		
Other	30	9.0	21	9.1	9	8.6		
Family breadwinner (*n*, %)							Fisher's exact = 14.99	<0.01
Yes	210	69.8	153	69.5	57	70.4		
No: partner is breadwinner	56	18.6	48	21.8	8	9.9		
No: parent is breadwinner	11	3.7	4	1.8	7	8.6		
No: two-earner household	15	5.0	11	5.0	4	4.9		
Other	9	3.0	4	1.8	5	6.2		
Breadwinner's main source of income (*n*, %)							Fisher's exact = 26.25	<0.01
Salary	177	60.6	141	65.9	36	46.2		
Pension	6	2.1	2	0.9	4	5.1		
Welfare	37	12.7	26	12.1	11	14.1		
Disability benefit	19	6.5	15	7.0	4	5.1		
Unemployment benefit	27	9.2	10	4.7	17	21.8		
Study subsidy	4	1.4	2	0.9	2	2.6		
Other	22	7.5	18	8.4	4	5.1		
Clinical characteristics								
Duration present episode (*n*, %)							Fisher's exact = 2.46	0.66
<6 months	84	25.1	58	24.9	26	25.7		
6 months to 1 year	89	26.6	57	24.5	32	31.7		
1–2 years	43	12.9	32	13.7	11	10.9		

Table 2 (*cont.*)

Variable	Total sample ($n = 341$)		Completers ($n = 233$)		Non-completers ($n = 108$)		Test statistic (df)	P
>2 years	108	32.3	78	33.5	30	29.7		
Unknown	10	3.0	8	3.4	2	2.0		
Prior treatment for current depressive episode (n, %)							$\chi^2(1) = 0.27$	0.62
No	218	65.3	150	64.4	68	67.3		
Yes	116	34.7	83	35.6	33	32.7		
Number of prior depressive episodes (n, %)							$\chi^2(2) = 6.12$	<0.05
0	103	31.1	70	30.3	33	33.0		
1	69	20.8	41	17.7	28	28.0		
⩾2	159	48.0	120	51.9	39	39.0		
Medication use in the last 4 weeks (n, %)							$\chi^2(1) = 0.03$	0.89
Yes	145	50.2	110	50.5	35	49.3		
No	144	49.8	108	49.5	36	50.7		
Co-morbid anxiety (Beck Anxiety Inventory; mean, S.D.)	24.46	13.16	23.59	12.97	27.13	13.47	$t(290) = -1.99$	0.05
Co-morbid psychopathology (Brief Symptom Inventory; mean, S.D.)	1.86	0.76	1.83	0.76	1.93	0.77	$t(260) = -0.94$	0.35
Consultations with general practitioner in the last 4 weeks (mean, S.D.)	0.85	1.21	0.87	1.19	0.80	1.26	$t(303) = 0.39$	0.70
Consultations with mental health care in the last 4 weeks (mean, S.D.)	0.73	1.14	0.75	1.12	0.67	1.21	$t(303) = 0.53$	0.60
Consultations with industrial medical officer in the last 4 weeks (mean, S.D.)	0.40	0.78	0.39	0.72	0.43	0.95	$t(290) = -0.38$	0.71
Psychological characteristics								
Cognitive reactivity to sad mood (LEIDS; mean, S.D.)	55.09	16.88	54.55	16.62	56.84	17.72	$t(287) = -1.03$	0.30
Anxiety sensitivity (Anxiety Sensitivity Index; mean, S.D.)	37.31	13.99	36.83	13.58	38.83	15.25	$t(278) = -0.97$	0.34

HAMD, Hamilton Depression Rating Scale; LEIDS, Leiden Index of Depression Sensitivity; S.D., standard deviation.

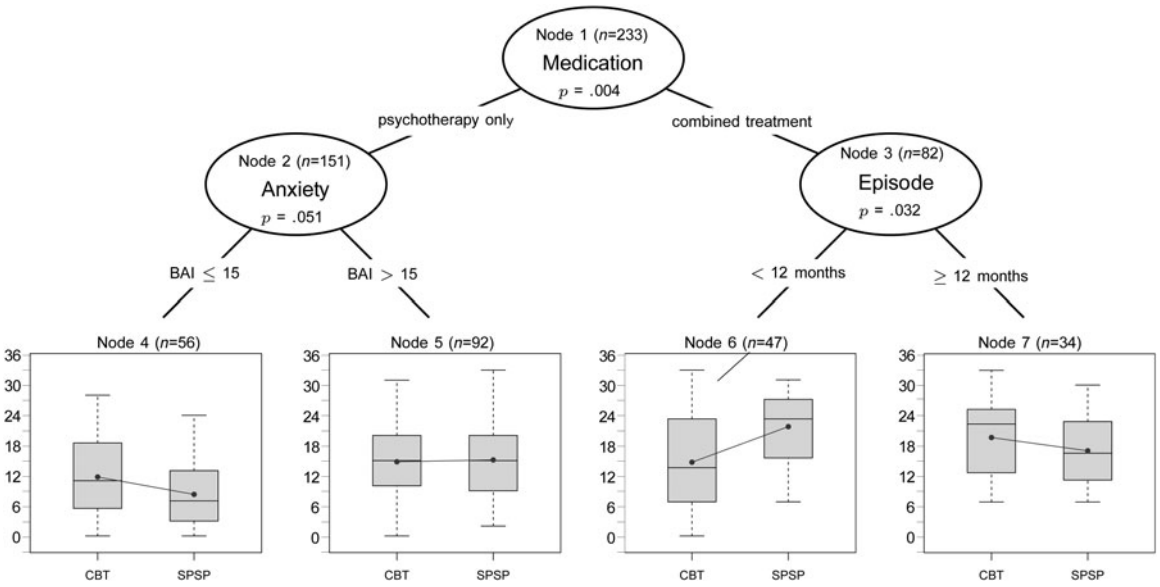

Fig. 1. ANOVA-based tree. BAI, Beck Anxiety Inventory; CBT, cognitive behavioral therapy; SPSP, short-term psychodynamic supportive psychotherapy. The model predicts HAMD score at post-treatment using the factor Therapy (CBT *v*. SPSP). Boxplots represent scores for both groups under the ANOVA model; the dots connected by a line represent the averages in the two groups. Due to missing values, the number of observations in the child nodes do not always add up to those in the parent node.

Table 3. *Outcomes in each node of the model-based recursive partitioning tree*

	CBT			SPSP			Contrast	
Node	*N*	Mean	S.D.	*N*	Mean	S.D.	Cohen's *d**	*p*
1	111	14.91	8.80	122	15.25	8.72	0.04	0.77
2	68	13.43	8.08	83	12.82	8.20	−0.07	0.65
3	43	17.26	9.45	39	20.44	7.50	0.36	0.10
4	27	11.74	7.66	29	8.28	6.95	−0.40	0.08
5	39	14.77	8.28	53	15.13	7.84	0.04	0.83
6	25	15.12	9.80	22	22.41	7.20	0.83	0.01
7	18	20.22	8.32	16	17.50	7.36	−0.31	0.32

CBT, Cognitive behavioral therapy; SPSP, short-term psychodynamic supportive psychotherapy.

Pooled standard deviation is 8.754.

* The standardized mean difference between the treatments (CBT *v*. SPSP) is expressed as Cohen's *d* (Cohen, 1988). A positive value indicates that CBT resulted in a better outcome (i.e. a lower post-treatment HAMD score) than SPSP, while a negative value indicates that SPSP resulted in a better outcome than CBT.

checked for mean differences in pre-treatment HAMD scores between treatment conditions within nodes 4, 6, and 7, but found none, suggesting that chance differences in mean pre-treatment HAMD scores within these nodes do not provide an alternative explanation for our findings.

We assessed the amount of bias in the effect sizes resulting from the MOB procedure using the bootstrapping procedure that has been described in detail by Dusseldorp & Van Mechelen (2014, online supplementary material C.2). With 1000 bootstraps, the resulting estimated bias in the range of effect sizes (Cohen's *d*; maximum-minimum) was 0.68. The range of effect sizes in the tree that resulted from the MOB analyses in our study was 1.23 ($d_{min} = -0.40$; $d_{max} = 0.83$). Thus, when using the tree resulting from this study in a new sample, the range of effect sizes could expected to be 1.23–0.68.

Competing splits

Because of the hierarchical nature of MOB, covariates and splitting values earlier in the tree impact the selection of those lower in the tree; small changes in the data may therefore result in different tree structures. For this reason, it is instructive to inspect the size and significance of alternative splitting variables.

The second best splitting variable for node 2 ($p = 0.077$) was breadwinner's main source of income. A split was made into two groups: (1) Salary/Pension/Study subsidy, and (2) Welfare/Disability benefit/Unemployment benefit/Other. In both groups SPSP resulted in a lower mean post-treatment HAMD score than CBT, but the difference in HAMD points was rather small in the first group (0.62) and markedly larger in the second group (2.44). The difference between the p value of the first split (Anxiety: $p = 0.051$) and this second best split ($p = 0.077$) was not that large, suggesting that co-morbid level of anxiety is only a marginally stronger moderator of differential efficacy than breadwinner's main source of income in the group of low-severity patients receiving psychotherapy only.

By contrast, for node 3 there was no clear competing split. The second best splitting variable for this node was cultural background. However, the p value of this split was not significant ($p = 0.172$). Thus, episode duration was the strongest moderator of differential efficacy in the group of high-severity patients receiving combined treatment.

Discussion

We conducted a hypothesis-generating study of prescriptive factors associated with differential efficacy of CBT and SPSP in a sample of patients seeking treatment for a depressive episode in psychiatric outpatient clinics. Minimal differences between treatments were found when examining the total sample of patients who completed post-treatment assessment, but model-based recursive partitioning analyses suggested differential efficacy among subgroups of patients that differed with regard to episode duration and co-morbid anxiety level. These findings are in line with another study that found a specific patient factor (length of caregiving) associated with differential efficacy of CBT and psychodynamic therapy for depression (Gallagher-Thompson & Steffen, 1994).

p Values within some of the nodes, relating to treatment differences in the group of patients represented in that node, were not significant at an $\alpha = 0.05$ level. Therefore, one might argue that no significant treatment differences are found in these subgroups and that, consequently, the factors that define these subgroups cannot be considered prescriptive factors. However, Kraemer *et al.* (2002) have argued that moderators of treatment should not be defined by p values ('because moderator status ... would then change with sample size' p. 881), but by population parameters instead. Effect sizes within these nodes did show relevant treatment differences, which can be considered large enough to be of importance.

We found co-morbid anxiety to be associated with differential treatment efficacy in low-severity patients receiving psychotherapy only. No treatment differences were found among patients who presented higher co-morbid anxiety levels and SPSP was found to be more efficacious than CBT for patients with low anxiety levels. We speculate that patients with low anxiety levels might benefit from SPSP, as they may be better capable to enter the more open relational- and insight-oriented SPSP dialog and feel less comfortable with the more structured and protocolized CBT.

Duration of the current depressive episode was associated with differential treatment efficacy in high-severity patients who received psychotherapy combined with antidepressant medication. CBT was more efficacious than SPSP for patients reporting a duration of <1 year, while SPSP was more efficacious than CBT for such patients reporting a duration of $\geqslant 1$ year. We speculate that patients with longer episode durations have depressive symptoms that are more influenced by their personality structure resulting in more complex working alliances and transference feelings; psychodynamic therapists are trained to elaborate on these therapeutic relational aspects if necessary. At the same time, we think that it is harder for patients with longer episode duration to start doing homework and to change their lifestyle and activities, which is what CBT calls for. Similarly, patients with relative short episode durations might benefit less from a psychodynamic approach, while the combination of antidepressant medication and behavioral activation resulting from completed homework assignments in the first part of CBT might lead to alleviation of depressive symptoms in this group.

This study has a number of strengths. First, it is one of few studies examining prescriptive factors associated with differential efficacy of CBT and psychodynamic therapy for depression, which refers to a clinically relevant but yet rather unexplored research topic related to a highly prevalent mental disorder. It also is the first study that considered demographic and illness characteristic as potential prescriptive factors. Furthermore, it is the first study to apply model-based recursive partitioning in this regard, which provides the benefits of reducing the number of prescriptive factors and selecting the most important ones, resulting in models that are easily interpretable. Additional strengths of this study

include its efforts to increase external validity, such as providing treatment in general psychiatric outpatient clinics by a large number of therapists with different experience levels and including patients with relatively low social-economic statuses (Driessen *et al.* 2013).

However, this study also has a number of limitations (see Driessen *et al.* 2007, 2013 for further discussion of the RCT's limitations). First, we examined prescriptive factors of CBT and psychodynamic therapy efficacy in a group of patients restricted to those that completed post-treatment assessment. Comparative analyses suggested that this group of patients differed from the group of patients lost to this assessment. Related, there was a baseline difference in marital status between the CBT and SPSP conditions in the group of patients who completed post-treatment assessment; a larger proportion of married persons received CBT. Moreover, bootstrapping suggested that the resulting estimated bias in the range of the effect sizes was 0.68. For these reasons, we need to be cautious in generalizing this study's results, particularly the effect sizes from Table 3, to the entire population of depressed patients. Second, HAMD assessors were not blind to patient grouping. Therefore, we cannot rule out observer bias.

The most important limitation of our study is, however, its *post-hoc* nature. Our findings are observational and can be the consequence of chance findings in the sample studied. The use of an alpha level of 0.10 (which allowed for a larger power for detecting prescriptive factors) as well as the interpretation of treatment differences within some of the Nodes that would not be significant under more stringent alpha levels, increases the possibility of chance findings. Our findings should, therefore, be validated before being used to guide treatment selection. Validation should occur by including episode duration and comorbid anxiety level (and possibly breadwinner's main source of income as well) as stratification variables in a future RCT comparing CBT and psychodynamic therapy in the treatment of depression and conduct a formal test of a moderator-by-treatment interaction (Kraemer *et al.* 2002). We also encourage replication of our study by means of applying model-based recursive partitioning in samples of other studies that compare different treatment options for depression in order to identify further potential prescriptive factors to test in such a future RCT.

In sum, although generally similar efficacy is found for CBT and psychodynamic therapy across larger patient samples, efficacy differences can be apparent in specific subgroups of patients. We identified comorbid anxiety level and depressive episode duration as possible prescriptive factors associated with differential efficacy of CBT and psychodynamic therapy. If validated, these variables might be used to guide treatment selection, which might improve the efficacy of psychotherapy for depression.

Acknowledgements

This study was financed by an unrestricted research grant by Wyeth Pharmaceuticals, The Netherlands. Arkin Mental Health Care, The Netherlands, financially supported research logistics and the contributions of E.D., J.D., F.J.D., J.P., S.K., D.W., M.H., and H.L.V. F.J.D.'s contributions were also financially supported by ProPersona Mental Health, The Netherlands. VU University, Faculty of Psychology and Education, Department of Clinical Psychology, The Netherlands, financially supported E.D.'s, N.S.'s and P.C.'s contributions to the study. None of the sponsors had a role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; nor in the preparation, review, or approval of the manuscript.

Declaration of Interest

H.L.V. and M.H. have received training fees from RINO, Amsterdam, and VU University, Amsterdam, for respectively lecturing students and clinicians short-term psychodynamic supportive psychotherapy, and H.L.V. is president of the Dutch Association of Psychoanalytic Psychotherapy. F.D. serves on the board of the mood disorders section of the Dutch Association of Cognitive and Behavior Therapy (unpaid). F.D. and J.J.M.D. receive royalties from Springer Media, The Netherlands, for authorship of the published CBT treatment manual. The remaining authors report no financial relationships with commercial interests.

References

Aiken LS, West SG (1991). *Multiple Regression: Testing and Interpreting Interactions*. Sage: Newbury Park.

Barth J, Munder T, Gerger H (2013). Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. *PLoS Medicine* **10**, e1001454.

Beck AT (1976). *Cognitive Therapy and the Emotional Disorders*. International Universities Press: New York.

Beck AT, Epstein N, Brown G, Steer RA (1988). An inventory for measuring clinical anxiety: psychometric properties. *Journal of Consulting and Clinical Psychology* **56**, 893–897.

Cohen J (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn. Lawrence Erlbaum Associates: Mahwah, New Jersey, USA.

Cohen J, Cohen P, West SG, Aiken LS (2013). *Applied Multiple Regression/Correlation Analysis for the Behavioral*

Sciences, 3rd edn. Lawrence Erlbaum Associates: Mahwah, New Jersey, USA.

Cuijpers P, van Straten A, Andersson G, van Oppen P (2008). Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies. *Journal of Consulting and Clinical Psychology* **76**, 909–922.

de Beurs E, Zitman F (2006). De Brief Symptom Inventory (BSI) De betrouwbaarheid en validiteit van een handzaam alternatief voor de SCL-90. [The Brief Symptom Inventory (BSI). Reliability and validity of a manageable SCL-90 alternative]. *Maandblad Geestelijke Volksgezondheid* **61**, 120–141.

de Jonghe F (1994). *Leidraad voor het scoren van de Hamilton Depression Rating Scale [Hamilton Depression Rating Scale scoring manual]*. Benecke Consultants: Amsterdam.

de Jonghe F (2005). *Kort en Krachtig. Kortdurende Psychoanalytische Steungevende Psychotherapie [Short and snappy. Short-term psychoanalytical supportive psychotherapy]*. Benecke N.I.: Amsterdam.

de Jonghe F, de Maat S, Van R, Hendriksen M, Kool S, van Aalst G, Dekker J (2013). Short-term Psychoanalytic Supportive Psychotherapy for depressed patients. *Psychoanalytic Inquiry* **33**, 614–625.

Driessen E, Van HL, Don FJ, Peen J, Kool S, Westra D, Hendriksen M, Cuijpers P, Twisk J, Dekker JJM (2013). The efficacy of cognitive behavioral therapy and psychodynamic therapy in the outpatient treatment of major depression: a randomized clinical trial. *American Journal of Psychiatry* **170**, 1041–1050.

Driessen E, Van HL, Peen J, Don FJ, Kool S, Westra D, Hendriksen M, Cuijpers P, Twisk J, Dekker JJM (2015). Therapist-rated outcomes in a randomized clinical trial comparing cognitive behavioral therapy and psychodynamic therapy for major depression. *Journal of Affective Disorders* **170**, 112–118.

Driessen E, Van HL, Schoevers RA, Cuijpers P, van Aalst G, Don FJ, Hendriksen M, Molenaar P, Dekker JJM (2007). Cognitive behavioral therapy versus short psychodynamic psychotherapy in the outpatient treatment of depression: a randomized controlled trial. *BMC Psychiatry* **7**, 58.

Dusseldorp E, Meulman J (2004). The regression trunk approach to discover treatment covariate interaction. *Psychometrika* **69**, 355–374.

Dusseldorp E, Spinhoven P, Bakker A, van Dyck R, van Balkom AJLM (2007). Which panic disorder patients benefit from which treatment: cognitive therapy or antidepressants? *Psychotherapy and Psychosomatics* **76**, 154–161.

Dusseldorp E, Van Mechelen I (2014). Qualitative interaction trees: a tool to identify qualitative treatment-subgroup interactions. *Statistics in Medicine* **33**, 219–237.

Gallagher-Thompson D, Steffen AM (1994). Comparative effects of cognitive-behavioral and brief psychodynamic psychotherapies for depressed family caregivers. *Journal of Consulting and Clinical Psychology* **62**, 543–549.

Hakkaart-van Roijen L (2002). *Handleiding Trimbos/iMTA questionnaire for Costs associated with Psychiatric Illness (TiC-P)*. Institute for Medical Technology Assessment, Erasumus Universiteit Rotterdam: Rotterdam.

Hamilton M (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry* **23**, 56–62.

Hothorn T, Hornik K, Strobl C, Zeileis A (2012). *Party: a laboratory for recursive partytioning (version 1.0–2) [Computer software manual]*. Available from http://cran.r-project.org/

Kraemer HC, Wilson T, Fairburn CG, Agras WS (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry* **59**, 877–883.

Leichsenring F (2001). Comparative effects of short-term psychodynamic psychotherapy and cognitive-behavioral therapy in depression: a meta-analytic approach. *Clinical Psychology Review* **21**, 401–419.

Molenaar PJ, Don F, van den Bout J, Sterk F, Dekker J (2009). *Cognitieve gedragstherapie bij depressie [Cognitive Behavioral Therapy for Depression]*. Bohn Stafleu van Loghum: Houten, The Netherlands.

NICE (2009). Depression: the treatment and management of depression in adults (update). (http://guidance.nice.org.uk/CG90/). National Institute for Health and Clinical Excellence. Accessed 21 December 2012.

R Development Core Team (2010). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Available from http://www.R-project.org

Reiss S, Peterson RA, Gursky DM, McNally RJ (1986). Anxiety sensitivity, anxiety frequency and the predictions of fearfulness. *Behavior Research and Therapy* **24**, 1–8.

Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Herqueta T, Baker R, Dunbar GC (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry* **59** (Suppl. 20), 22–33.

Stevens J (2002). *Applied Multivariate Statistics for the Social Sciences*, 4th edn. Lawrence Erlbaum: London.

Strobl C, Malley J, Tutz G (2009). An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods* **14**, 323–348.

van der Does W (2002). Cognitive reactivity to sad mood: structure and validity of a new measure. *Behavior Research and Therapy* **40**, 105–120.

Ward EC (2007). Examining differential treatment effects for depression in racial and ethnic minority women: a qualitative systematic review. *Journal of the National Medical Association* **99**, 265–274.

Zeileis A, Hothorn T, Hornik K (2008). Model-based recursive partitioning. *Journal of Computational and Graphical Statistics* **17**, 492–514.

Zeileis A, Hothorn T, Hornik K (2012). Party with the mob: model-based recursive partitioning in R (Tech. Rep.). (http://cran.r-project.org/web/packages/party/vignettes). Accessed 19 November 2012.