CHAPTER 7

Brief cognitive behavioural therapy compared to optimised general practitioners’ care for depression: a multicenter randomised trial

Kim D. Baas, Maarten W.J. Koeter, Henk C. van Weert, Peter Lucassen, Claudi L.H. Bockting, Karin A. Wittkampf, Jochanan Huyser, Aart H. Schene

Submitted
Abstract

Background How to treat major depressive disorder (MDD) in primary care? A brief form of psychotherapy, including both cognitive and behavioural ingredients has so far not been compared with optimized treatment by the general practitioner (GP).

Aims To investigate whether a protocolised 8 session brief cognitive behavioural therapy (CBT) is more effective than optimised GP care, based on current evidence based guidelines (GPC).

Method 121 patients with MDD, 18-70 years, from 40 Dutch general practices, were randomised to either GPC or brief CBT. Main outcomes: depressive symptomatology, response and remission on the Hamilton Depression Rating Scale-17 (HDRS-17) and the Patient Health Questionnaire-9 (PHQ-9). (Trial registration: ISRCTN65811640)

Results Both groups experienced a substantial reduction of mean HDRS-17 and PHQ-9 scores from baseline to 1 year follow-up. Number of treatment contacts and external referrals were not different.

Conclusions Brief CBT was not more successful than GPC. GPs prescribed antidepressants to 48% of GPC patients and to 11% of brief CBT patients.
**Introduction**

In the general population Major Depressive Disorder (MDD) is highly prevalent with 12 month prevalence rates ranging from 4 to 8%.\textsuperscript{1-3} Because of different health care systems, year prevalence rates of MDD in primary care (PC) vary across the world from 12 to 25%\textsuperscript{4} and so MDD is among the mental illnesses the general practitioner (GP) most frequently encounters. The majority of patients suffering from MDD is treated in primary care settings, mostly by their GP\textsuperscript{5} and about two thirds of them prefer psychological therapy to antidepressants (AD).\textsuperscript{6,7}

Treatment strategies differ for mild and moderate/severe MDD. For mild MDD treatment guidelines from the National Institute for Health and Clinical Excellence\textsuperscript{8} and the American Psychiatric Association\textsuperscript{9} recommend low intensity psychological therapy (e.g. computerized Cognitive Behavioural Therapy) and only in some cases AD (e.g. in case of a past history of more severe depression or inadequate response to initial interventions).

For moderate and severe MDD guidelines recommend AD, or psychological therapy, or a combination of both. Of the psychological interventions Cognitive Behavioural Therapy (CBT) has the best evidence base.\textsuperscript{8}

CBT can be delivered in brief forms (i.e. fewer than ten sessions).\textsuperscript{10} These brief forms are not only more suitable for primary care settings, they are also more acceptable for patients. From an explanatory trial perspective brief CBT forms are better for a comparison with non psychological treatments than classical 16 to 18 sessions CBT. In case of longer forms, differences in the total amount of attention may lead to differences in outcomes.

Effectiveness studies showed that for patients with MDD treatment as usual by the GP is less effective than brief CBT especially in the short term (i.e. four months).\textsuperscript{11-15} However, in most comparisons of treatment as usual with an experimental treatment, the first is not standardised according to clinical guideline recommendations. So the quality may vary as a result of differences between GPs and/or mental health workers to whom they refer. Consequently, this may result in most cases in an overestimation of the effect of the experimental treatment.

Protocolising treatment as usual and maximising adherence to it are possible solutions to this methodological problem.\textsuperscript{16-19} In this study we did so by summarising and manaulising the Dutch GP Practice Guideline for MDD and maximising adherence to this manual. We also asked GPs to refrain from referrals to other health care professionals.
We will refer to this intervention as optimised general practitioners’ care (GPC). In our study we investigated whether a protocolised brief CBT (8 sessions) was more effective than optimised GPC for primary care patients with MDD.

**Methods**

**Setting**

The study was conducted from January 2007 through April 2010 in patients of 40 general practices in two regions (Amsterdam and Nijmegen) in The Netherlands. The study protocol was approved by the institutional ethics review committees of the Academic Medical Center in Amsterdam and the University Medical Center St Radboud in Nijmegen. Details of the study design, including the content of the two treatments, can be found in Baas et al.²⁰

**Participants and procedure**

GPs referred patients with a possible MDD to the study centre. Patients who gave informed consent for the inclusion procedure were assessed within three days with a telephonic Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I)²¹,²² to determine eligibility. Inclusion criteria were: MDD and age between 18 and 70 years. Eligible patients were asked for a second informed consent form for the actual study. After baseline assessment they were randomised to optimised general practitioners’ care (GPC) or brief cognitive behavioural therapy (CBT).

**Randomisation**

After informed consent, GPs performed an internet based randomisation procedure. A centralised computer program generated a stratified block randomisation (block randomisation of 4 blocks stratified by gender and location (Nijmegen/ Amsterdam)) where the patient was randomised over the two treatment strategies with no preference option. The computer then generated an outcome on the GP’s computer screen and simultaneously sends an email of this outcome to the researcher. In this way, patients and investigators were not able to foresee the allocation.

**Interventions**

*General practitioners’ care (GPC)*

The GPC treatment protocol (developed by HvW and AS and available on request) was based on the Dutch College of General Practitioners Practice Guideline (NHG-Standaard) for depressive disorder.²³ The GPC had a duration of 12 weeks and consisted of supportive contacts with psycho-education, life style advices and monitoring...
of symptoms which, in accordance with the guidelines, could be combined with an antidepressant. So ADs were one of the treatment options in the GP guidelines. GPs were free to choose whether they prescribed an AD, it was finally their decision. Minimum frequency was one contact every two weeks during the first six weeks and one telephonic contact and one face-to-face evaluation contact during the next 6 weeks. This could be increased in case of severity of the symptoms and/or complaints and/or lack of social support (for an overview of the contents of the GPC see Baas et al.20) During these 12 weeks GPs were asked to refrain from referring the patient (for example to secondary care) and to contact the researcher if they wanted to divert from the protocol.

Brief Cognitive Behavioural Therapy (brief CBT)
The brief CBT protocol (developed by CLHB and available on request) consisted of 8 fifty minutes sessions within a 12 week period. Treatment included behavioural activation and cognitive interventions including identification and challenging of negative thoughts and underlying attitudes and schema (for an overview of the contents of the brief CBT see Baas et al.20) All therapists were skilled psychologists connected to one of the participating general practices and specifically trained in this form of therapy. They all had a Master in Clinical Psychology, a 4 year post academic education in Behavioural Therapy and were members of the Association of Behavioural and Cognitive Therapy (VGct).

All brief CBT sessions were audio taped and treatment integrity was assessed by checking whether the essential ingredients of the intervention were present (behavioural activation, i.e. identification and expanding of potentially pleasant activities and identifying and challenging negative thoughts/formulating rational thoughts) in a random sample of 10% of each therapists tapes. Therapists had intervision sessions (an organized meeting between colleagues in which performed sessions and related problems can be discussed) and supervision sessions (an organized meeting between the therapist and a supervisor, a fully trained cognitive behavioural therapist with a Master in Clinical Psychology and a 4-year post academic education in Behavioural Therapy and Clinical Psychology). Both may lead to a quality improvement.

During the treatment period therapists were asked to refrain from referring the patient to the GP for medication or referral to another health care professional and to contact the researcher if they wanted to divert from the protocol.

Measurements
Patients were assessed at three occasions: before randomisation (baseline); at the end of the treatment period (12 weeks) and at one year follow-up (52 weeks).
Eligibility assessment
The Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) is a semi-structured interview designed for diagnosing mental disorders according to the DSM-IV criteria. Throughout the study all SCID-I interviewers participated in ongoing training sessions and monthly consensus meetings supervised by an expert psychiatrist (JH) to maximize accuracy and consistency in the administration.

Outcome measures
As primary outcome we used both a clinician-rated scale (Hamilton Depression Rating Scale-17; HDRS-17) and a patient-rated scale (Patient Health Questionnaire-9; PHQ-9). For both scales we measured continuous outcomes. As stated in Baas et al. we also measured dichotomous outcomes (response and remission) with the HDRS-17 and we have added here response and remission for the PHQ-9.

The Hamilton Depression Rating Scale-17 (HDRS-17)
The 17-item HDRS is a well known clinician rated, semi-structured clinical interview for assessing the severity of depressive disorder. In this study it was administered by telephone. Three independent interviewers administered the HDRS-17. They were trained in administration of the HDRS by an expert psychiatrist (JH). To maximize accuracy and consistency in the administration we used the same procedure as with the SCID-I. The HDRS-17 total score ranges from 0 to 52 points and can be categorised as: 8-13, mild; 14-18, moderate; 19-52, severe. The HDRS-17 has acceptable psychometric properties.

The Patient Health Questionnaire-9
The PHQ-9 is the 9 item patient rated subscale of the Primary Care Evaluation of Mental Disorders (PRIME-MD). It evaluates the presence of the nine DSM-IV criteria for a depressive episode. The sum score ranges from 0 to 27 and can be categorised as: 5-9, mild; 10-14, moderate; 15-19, moderately severe; 20-27, severe. The PHQ-9 also offers a diagnostic algorithm. A positive outcome on the algorithm requires that five or more of the nine depressive symptom criteria of the DSM-IV are present more than half the days in the past 2 weeks (suicidal thoughts count if present at all), and at least one of these 5 or more symptoms has to be depressed mood or anhedonia. The PHQ-9 connects well to Primary Care.

Dichotomous outcomes
Response was defined as a total score reduction of >50% on the HDRS-17 or the PHQ-9. Using the HDRS-17, remission was defined as a total score of 7 or less.
there is no consensus on the definition of remission for the PHQ-9, we applied two definitions. First, the original validation study of the PHQ-9 recommends that scores of 0–4 are in the minimal depression range and scores of 5–9 in the mild depression range. We therefore defined remission on the PHQ-9 as a score of <5. Second, the cut-off value most widely used to identify a positive case for depressive disorder is a total score of 10 or higher. We therefore also used a more lenient, remission definition: a PHQ-9 total score <10. Response was assessed only at 12 weeks, remission both at 12 weeks and 52 weeks.

Blinding
Before the HDRS-17 interview, the researcher provided the interviewers only with the telephone number and name of the patient to be interviewed. In this way group assignment was kept blind. Before the interview patients were asked “not to reveal their treatment condition”. No data was collected on the success of the blinding, however, we asked the interviewers in the monthly consensus meetings in how many cases the blinding was broken. They concluded that in 85% to 95% of the cases patients had been able to retain the concealment.

Treatment
In order to optimise the GPC treatment we paid each participating GP a visit before the start of the study. During this visit we educated the GP in the treatment protocol and discussed the content of each of the contacts. After the visit we provided a ring binder with the treatment protocol (containing the content of the contacts) and informed the GP about the possibility of consulting an independent physician for questions about the treatment protocol during the intervention period.

To monitor the actual content of the provided GPC and potential supplementary treatment next to GPC and brief CBT, each GP received a short questionnaire at the end of the 12 week treatment period in which we asked how many appointments they had had with the patient, whether they prescribed an antidepressant (and the dosage) and whether they combined their treatment with a treatment by another health care professional or referred their patient.

Statistical methods
Adequacy of the randomisation was assessed by comparing the two groups at baseline on socio-demographic and clinical characteristics. For continuous variables we used the t-test for independent samples and for categorical variables we used the chi-square test. When expected cells counts were too small for the Chi Square Test, Fishers Exact Test was used.
We did not assess the data for Normality since the sample sizes of both groups were considered sufficient to make the t-test robust to deviation from normality.36

Multiple Imputation (MI) was used in case of missing data. In situations where missing data are missing completely at random (MCAR) or missing at random (MAR), multiple imputation leads to unbiased results with correct standard errors.37 The results of the multiple imputation analyses were combined using Rubin’s rules (1987).38 F-test values and degrees of freedom were calculated by a method proposed by Marshall et al.39 Since MI based pooled estimates are considered less biased, all analyses are based on the MI based estimates. In the tables, however, the imputed data as well as the actual observed data are shown. Furthermore, all analyses were intention to treat and involved all patients who were randomly assigned. Analyses were carried out with SPSS Statistics 18.0. All statistical tests were 2-sided with an alpha of 0.05.

Treatment effect
Continuous outcomes were analysed with a linear mixed model regression analysis with change between baseline and 12 weeks and change between baseline and 52 weeks as dependent variables; treatment group, time and the treatment group by time interaction as independent variables; and baseline severity as covariate. In this model the main effect of treatment group indicates treatment effect and the treatment by time interaction indicates whether this effect sustains. When the treatment by time interaction was significant, we performed planned contrasts to assess whether changes from baseline in HDRS-17 and PHQ-9 scores differed between both groups at 12 weeks and at 52 weeks.

Percentage response was analysed with logistic regression analysis with response in terms of the HDRS-17 or the PHQ-9 as dependent variable and treatment group as independent variable, and baseline severity as covariate. Percentage remission was analyzed with longitudinal logistic regression analysis, using the SPSS Generalized Estimating Equations programme (GEE) with a logit link function and a binomial error distribution. Dependent variable was remission in terms of the HDRS-17 or the PHQ-9. Independent variables were treatment group, time, the treatment group by time interaction and baseline severity as covariate.
Results

Patient flow and demographic characteristics

Figure 1 shows the patient flow through the study. During the recruitment phase, 175 patients were referred of whom 170 could be assessed for eligibility. Five patients (2.9%) improved during the recruitment phase and therefore declined participation. Of the 170 patients, 121 (71.2%) met the entry criteria and agreed to participate. Thirty four (20%) patients did not meet the inclusion criteria, and eight (4.7%) patients refused randomisation due to treatment preference. Six (3.5%) patients said they were improved and therefore declined randomisation, one (0.6%) patient did not accept the diagnosis of depression.

![Patient flow diagram]

*BCBT* = brief cognitive behavioural therapy; GPC = general practitioners’ care

Figure 1. Patient flow according to the CONSORT criteria

---

111
### Table 1. Baseline demographic and clinical characteristics n=121

<table>
<thead>
<tr>
<th>Demographic characteristics</th>
<th>GPC (n=61)</th>
<th>bCBT (n=60)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age, mean (s.d)</strong></td>
<td>45.3 (11.0)</td>
<td>44.9 (11.6)</td>
</tr>
<tr>
<td><strong>Female %</strong></td>
<td>72.1</td>
<td>61.7</td>
</tr>
<tr>
<td><strong>Living situation %</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Alone</td>
<td>25.0</td>
<td>21.4</td>
</tr>
<tr>
<td>- With partner (with or without child)</td>
<td>50.1</td>
<td>55.4</td>
</tr>
<tr>
<td>- With child (without partner)</td>
<td>18.3</td>
<td>10.7</td>
</tr>
<tr>
<td>- With parent(s)</td>
<td>3.3</td>
<td>7.1</td>
</tr>
<tr>
<td>- Other</td>
<td>3.3</td>
<td>5.4</td>
</tr>
<tr>
<td><strong>Caucasian %</strong></td>
<td>78.3</td>
<td>78.6</td>
</tr>
<tr>
<td><strong>Marital status %</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Not married</td>
<td>45.0</td>
<td>42.8</td>
</tr>
<tr>
<td>- Married</td>
<td>36.7</td>
<td>41.1</td>
</tr>
<tr>
<td>- Divorced</td>
<td>11.7</td>
<td>14.3</td>
</tr>
<tr>
<td>- Widowed</td>
<td>6.6</td>
<td>1.8</td>
</tr>
<tr>
<td><strong>Primary role %</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Employed</td>
<td>56.7</td>
<td>62.5</td>
</tr>
<tr>
<td>- Student</td>
<td>0.0</td>
<td>1.8</td>
</tr>
<tr>
<td>- Unemployed</td>
<td>26.7</td>
<td>23.2</td>
</tr>
<tr>
<td>- Retired</td>
<td>3.3</td>
<td>1.8</td>
</tr>
<tr>
<td>- Other</td>
<td>13.3</td>
<td>10.7</td>
</tr>
<tr>
<td><strong>Educational attainment a %</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Low</td>
<td>31.7</td>
<td>33.9</td>
</tr>
<tr>
<td>- Medium</td>
<td>48.3</td>
<td>41.1</td>
</tr>
<tr>
<td>- High</td>
<td>20.0</td>
<td>25.0</td>
</tr>
<tr>
<td><strong>Clinical characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recurrent depression %</td>
<td>39.6</td>
<td>51.9</td>
</tr>
<tr>
<td>Severity depression (HDRS) %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Mild (total score 8-13)</td>
<td>42.6</td>
<td>46.7</td>
</tr>
<tr>
<td>- Moderate (total score 14-18)</td>
<td>36.1</td>
<td>38.3</td>
</tr>
<tr>
<td>- Severe (total score 19-52)</td>
<td>21.3</td>
<td>15.0</td>
</tr>
<tr>
<td>HDRS-17, mean (s.d)</td>
<td>19.8 (6.4)</td>
<td>18.3 (5.8)</td>
</tr>
<tr>
<td>PHQ-9, mean (s.d)</td>
<td>18.3 (4.0)</td>
<td>17.4 (5.1)</td>
</tr>
<tr>
<td>Co-morbid axis-I disorders, current b %</td>
<td>32.8</td>
<td>30.0</td>
</tr>
</tbody>
</table>

GPC=general practitioners’ care; bCBT=brief cognitive behavioural therapy; a Educational attainment was divided in three categories with respect to the Dutch Educational System; b Overlap between the disorder groups (anxiety-, somatoform disorders and addiction).
Mean age of the total sample (n=121) was 45.1 years (s.d 11.2) and 66.9% of the patients were women. Mean HDRS-17 score was 19.1 (s.d 6.1) and mean PHQ-9 score was 17.8 (s.d 4.6), both on the border of moderate and severe MDD. Fifty five patients (45.7%) had a recurrent depressive disorder and thirty eight patients (31.4%) suffered from a co-morbid DSM-IV axis-I disorder (anxiety-, somatoform- and/ or substance related disorder). Both groups were comparable on all demographic and clinical characteristics, both in terms of statistical significance and effect size. Figures stratified by intervention are shown in table 1.

**Treatment effect**

Data on outcome were obtained for 94 (78%) patients at 12 weeks and for 78 (64%) patients at 52 weeks.

Continuous outcomes: table 2 shows that both GPC and brief CBT resulted in a reduction of depressive symptoms from baseline to 52 weeks. Mean HDRS-17 scores declined from 19.8 to 10.1 in the GPC group and from 18.3 to 8.2 in the brief CBT group. Mean PHQ-9 scores declined from 18.2 to 9.4 in the GPC group and from 17.4 to 7.1 in the brief CBT group. No significant treatment by time interaction (neither for the HDRS-17 nor for the PHQ-9) was found. Therefore we restricted ourselves to the regression model with only the main effects. Patients in the brief CBT group improved 2.17 points more on the HDRS-17 total score than patients in the GPC group (F(1, 18.42) = 1.96; 95% CI -0.86- 4.37; p=.178). Patients in the brief CBT group improved 2.38 points more on the PHQ-9 total score compared to patients in the GPC group (F(1,319.69) = 6.71; 95% CI 0.59- 4.78; p=.010).

Dichotomous outcomes: table 3 shows that on the HDRS-17 a response was shown by 34.1% of the GPC group and by 48.7% of the brief CBT group. For the PHQ-9 scores responses were respectively 39.3% and 57.0%. Logistic regression analyses showed no significant differences in percentage response on the HDRS-17 and PHQ-9 (Responder: HDRS-17, odds ratio 1.90; F(1,1183) = 2.70; 95% CI 0.88- 4.09; p=.100. PHQ-9, odds ratio 2.22; F(1,83)=3.42; 95% CI 0.95- 5.14; p=.068).

Finally we looked at remission from baseline to 52 weeks. Based on HDRS-17 scores 44.6% of the patients in the GPC group and 55.7% in the brief CBT group could be considered remitter (odds ratio 1.97; F(1,34.39)= 2.49; 95% CI 0.87- 3.70; p=.124). For the PHQ-9 (total score <10) these figures were respectively 54.8% and 71.7% (odds ratio 2.14; F(1,268)= 3.52; 95% CI 0.95-4.83; p=0.062). For the PHQ-9 (total score <5) these figures were respectively 23.0% and 33.0% (odds ratio 1.87; F(1,67)=3.42; 95% CI 0.97-3.63; p=0.069).
Table 3. Dichotomous depression outcome: Response and remission\(^{a,b}\)

<table>
<thead>
<tr>
<th></th>
<th>GPC(^c) (n=61)</th>
<th>bCBT(^c) (n=60)</th>
<th>F (df1, df2)(^{d,e})</th>
<th>p(^{d,e})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response(^f) (end of treatment)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDRS-17</td>
<td>34.1%</td>
<td>48.7%</td>
<td>2.70 (1, 1183)</td>
<td>0.100</td>
</tr>
<tr>
<td>PHQ-9</td>
<td>39.3%</td>
<td>57.0%</td>
<td>3.42 (1, 83)</td>
<td>0.068</td>
</tr>
<tr>
<td>Remission(^g)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDRS-17</td>
<td></td>
<td></td>
<td>2.49 (1, 34)</td>
<td>0.124</td>
</tr>
<tr>
<td>End of treatment</td>
<td>32.5%</td>
<td>51.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 year follow-up</td>
<td>44.6%</td>
<td>55.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHQ-9 Cut-off &lt;10</td>
<td></td>
<td></td>
<td>3.52 (1, 268)</td>
<td>0.062</td>
</tr>
<tr>
<td>End of treatment</td>
<td>42.6%</td>
<td>57.3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 year follow-up</td>
<td>54.8%</td>
<td>71.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHQ-9 Cut-off &lt;5</td>
<td></td>
<td></td>
<td>3.42 (1, 67)</td>
<td>0.069</td>
</tr>
<tr>
<td>End of treatment</td>
<td>10.8%</td>
<td>30.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 year follow-up</td>
<td>23.0%</td>
<td>33.0%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) Presented are multiple imputation based pooled estimates; \(^b\) Observed percentages response HDRS-17: GPC 36.7%, bCBT 51.1%; PHQ-9: GPC 39.5%, bCBT 61.8%; Observed percentages remission at 12 weeks: HDRS-17: GPC 36.7%, bCBT 55.6%; PHQ-9 (<10) GPC 44.2%, bCBT 61.8%; PHQ-9 (<5): GPC 11.6%, bCBT 35.3%; at 52 weeks HDRS-17: GPC 51.2%, bCBT 56.8%; PHQ-9 (<10): GPC 59.5%, bCBT 75.0%; PHQ-9 (<5): GPC 27.3%, bCBT 37.5%; \(^c\) GPC=general practitioners’ care; bCBT=brief cognitive behavioural therapy. \(^d\) Response: logistic regression analysis with response (Y/N) as outcome variable and treatment as dependent variable; remission: GEE analysis with remission (Y/N) as dependent and treatment and time as independent variables (treatment by time interaction was not significant); \(^e\) Pooled multivariate test according to Marshall et al. \(^{49}\); \(^f\) Response defined as at least 50% decrease of the baseline score; \(^g\) Remission defined as: HDRS-17 score ≤7 or PHQ-9 score <10 and PHQ-9 score <5.
Adherence and antidepressant use

Patients in the brief CBT group received a mean of 6.1 sessions (range 0-8; s.d 2.7). This is in line with other primary care studies on brief CBT in primary care.15 Three patients did not attend any treatment session. The most common reason for termination of therapy was lack of motivation to attend the sessions. Treatment integrity assessment of the brief CBT showed that the essential ingredients of the intervention (i.e. behavioural activation and challenging negative thoughts/ formulating rational thoughts) were present. No interference in the work procedures of the psychologists was necessary.

Patients in the GPC group received a mean of 4.9 (face to face or telephonic) appointments (range 1-12; s.d 2.2). This corresponds well to the recommended frequency of five contacts during the 12 week treatment period mentioned in the protocol. All patients attended at least one appointment. GPs prescribed antidepressants to 48% of the GPC patients and to 11% of brief CBT patients. In total one in seven patients received an outside referral for additional mental health treatment (11% in brief CBT and 18% in GPC).

Discussion

We investigated whether a protocolised brief CBT applied by psychologists was more effective than optimised GPC based on a clinical GP treatment guideline for primary care patients with MDD. The overall results showed that brief CBT did not lead to better outcomes than GPC. The continuous outcomes indicated that patients of both therapies experienced a substantial reduction of mean depression scores, both on the HDRS-17 and PHQ-9, from baseline to 1 year follow-up. Only the improvement in terms of mean PHQ-9 score differed significantly between both groups in favour of brief CBT. The therapies did not differ significantly in terms of the dichotomous outcomes; response (50% reduction of total scores) and remission (total score of 7 or less on the HDRS-17; total score of less than 10 and total score of less than 5 on the PHQ-9). The PHQ-9 remission rates based on the definition of a PHQ-9 score <10 were more comparable to the HDRS-17 remission rates than those based on the definition of a PHQ-9 score <5. The definition of the PHQ-9 score <5 proved to be too stringent.

Although we cannot conclude from our data that the outcome for both groups differs, the trend appears to consistently favour the CBT group. It cannot be ruled out that our study had insufficient power (due to the relatively small sample sizes) to detect effects of the observed magnitude.
Chapter 7

Considering process data, patients in the GPC group received on average 4.9 (face to face or telephonic) appointments, while this number was 6.1 sessions in the CBT group. This suggests a comparable number of professional contact in both groups. Furthermore the referral rate to other mental health treatments was comparable (1 in 7 patients).

ADs were one of the treatment options in the GPC and GPs were free to choose whether they prescribed an AD. This ultimately resulted in 48% of antidepressant description in the GPC group. This is low compared to other GP studies where 49-96% of the patients received AD.41,42,13

In summary, our hypothesis that brief CBT would outperform GPC was not supported. However, GPs prescribed ADs in approximately half of the GPC patients. GPs had received feedback of the diagnostic interview results (i.e. diagnosis and severity of the depression) and they were educated in the treatment protocol based on the primary care guidelines. So we believe they were well informed and were able to prescribe ADs in accordance with the guidelines.

Though psychologists were asked to refrain from referral to the GP for additional medication, 11% of the brief CBT patients was prescribed an additional AD. Since the majority of patients prefer psychological therapy to pharmacological therapy6 this about 40% AD reduction may be an advantage of brief CBT. However, since we made no measurements of patient experience or satisfaction, this advantage is not proven in this sample. Another possible advantage of CBT is its prophylactic effect on depression recurrence on the long term.43

Earlier studies on brief CBT in depressed primary care patients are scarce and showed brief CBT to be more effective than treatment as usual, especially in the short term, but the advantages were small.11-13, 44 Moreover, Conradi et al14 found that brief CBT outperformed treatment as usual only in patients with recurring depression with at least 4 episodes.

Our study differs from these earlier studies in several ways. First, we optimised treatment as usual by using a treatment protocol based on clinical GP treatment guidelines and we educated the GPs in this treatment protocol. We also asked GPs to refrain from referring the patient. Our study design results in a more valid comparison and consequently more valid estimates of the differences in treatment effects between brief CBT and GPC. To our knowledge only one other study also educated GPs in the depression guidelines. Conradi et al14 invited GPs to attend a 2-hour booster session about guidelines for the treatment of depression. However, they did not ask the GPs to refrain from referral.
Second, the number of offered sessions of the brief CBT differed. In two of the earlier studies brief CBT consisted of more than 10 sessions (range 10-12 sessions).\textsuperscript{11,14} Also, the content of the brief CBT differed. Scott et al\textsuperscript{13}, for instance, used mainly cognitive techniques instead of a mixture of cognitive and behavioural aspects, as we did.

Third, the strategy of treatment allocation differed in one of the five studies. This study allowed, in addition to randomisation, self selection of patients to their preferred allocation.\textsuperscript{12}

Finally, the way the severity of major depressive disorder was assessed differed. Only one other study used both a patient rated and a clinician rated instrument.\textsuperscript{13} The other studies used either a clinician rated instrument (HDRS-17)\textsuperscript{11} or a patient rated instrument (Beck Depression Inventory (BDI)).\textsuperscript{45} A recent meta-analysis showed that clinician-rated and patient-reported measures of improvement in depression studies are not equivalent and the authors recommend to use both.\textsuperscript{46}

\textit{Strengths and limitations}

Strengths of this study are the adequate randomisation procedure, the moderate attrition rate, the 52 week follow-up, and the detailed analyses of treatment effects. We have tried to improve generalisability by including a great number of general practices from two distinct geographical areas\textsuperscript{47,48} and by using lenient inclusion criteria, permitting co morbidity. Finally the optimisation and standardisation of the GP treatment and the refraining from referral allowed a fairer comparison.

Our study also has limitations. First, one could argue that a control or treatment as usual group should have been included. However we think that CBT had already been proven effective as compared to TAU\textsuperscript{15} and so we considered a placebo or TAU control group not necessary.

Second, we obtained no information about harms (adverse effects) of the GPC while, for example, side effects of medication can have a great influence on whether an intervention will be acceptable.

Third, although our follow-up time was sufficient, we could have add additional time points (e.g. at 6 and 9 months) to assess service use between 12 and 52 weeks. Service use between 12 and 52 weeks may have had consequences for what can be contributed to the randomised treatments after 52 weeks. However, this is often a difficult task in primary care.

Fourth, although we did not measure adherence to the protocol, we monitored the actual content of the provided GPC by sending each GP a short questionnaire at the end
Chapter 7

of the intervention period in which we asked how many appointments they had had with the patient, whether they prescribed an antidepressant (and the dosage) and whether they combined their treatment with a treatment by another health care professional or referred their patient. This provides an overall idea of adherence to the protocol.

Assessment of the degree to which the GP care was optimised in comparison to GP care in the wider literature is difficult since mostly the content of GP care varies widely and the content of the contacts is often not reported. Detailed data from meta-analyses of CBT versus usual care might help to put the outcomes in context. Three meta-analyses\textsuperscript{15,49,44} with forest plots showed effect sizes of CBT versus usual care in the range of 0.33 to 0.42. The effect size for the HDRS-17 of brief CBT versus GPC in our study was comparable (0.41).

Conclusion
Both therapy groups achieved a substantial reduction of symptoms from baseline to one year follow-up. For the majority of patients a primary care intervention seems sufficient, however one in seven patients was referred for additional therapy. We can conclude that brief CBT did not lead to better outcomes than GPC. However, to reach comparable results GPs relatively often prescribe antidepressants. For patients who prefer psychological therapy or have an aversion to antidepressants, brief CBT provides a good alternative. CBT may have a prophylactic effect on recurrence after treatment is stopped, while antidepressants have to be continued to reach comparable prophylaxis. Future research is needed to examine the relative cost-effectiveness of brief CBT and GPC. Furthermore, brief CBT is effective but can be offset by antidepressants, as these are also effective in a different way. Future studies should focus on identifying differential predictors for these treatments to reveal what works for whom.

Funding
This study was funded by a grant from the Netherlands Organisation for Health Research and Development (ZonMw), program Mental Health (# 100.003.005 and # 100.002.021) and the Academic Medical Center/University of Amsterdam.

Acknowledgements
We thank the participants, the participating GPs and the first line psychologists for their cooperation.
REFERENCES


Chapter 7


