Characteristics of effective psychological treatments of depression: a metaregression analysis

Pim Cuijpers; Annemieke Van Straten; Lisanne Warmerdam; Niels Smits

* Department of Clinical Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands

First published on: 06 November 2007
Characteristics of effective psychological treatments of depression: A metaregression analysis

PIM CUIJPERS, ANNEMIEKE VAN STRATEN, LISANNE WARMERDAM, & NIELS SMITS

Department of Clinical Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands

(Received 9 March 2007; revised 23 April 2007; accepted 7 May 2007)

Abstract

Although many meta-analyses have shown that psychological therapies are effective in the treatment of depression, no comprehensive metaregression analysis has been conducted to examine which characteristics of the intervention, target population, and study design are related to the effects. The authors conducted such a metaregression analysis with 83 studies (135 comparisons) in which a psychological treatment was compared with a control condition. The mean effect size of all comparisons was 0.69 (95% confidence interval = 0.60–0.79). In multivariate analyses, several variables were significant: Studies using problem-solving interventions and those aimed at women with postpartum depression or specific populations had higher effect sizes, whereas studies with students as therapists, those in which participants were recruited from clinical populations and through systematic screening, and those using care-as-usual or placebo control groups had lower effect sizes.

It is well established that psychological interventions are effective in the treatment of depression. In the past three decades, at least 160 controlled and comparative studies have examined the effects of psychological treatments compared with control conditions and other treatments (Cuijpers & Dekker, 2005). This large number of studies has clearly shown that most psychological treatments studied in a trial have large effects on depression.

Meta-analyses have played a major role in showing not only that psychological treatments are effective but also what types of psychological treatments are effective and for which types of patients. More than 40 meta-analyses have been conducted in this area (Cuijpers & Dekker, 2005).

First, these meta-analyses have examined whether psychological treatments are also effective in specific populations. These studies have shown that psychological treatments are effective not only in adults with depression but also in older adults (Cuijpers, van Straten, & Smit, 2006), women with postpartum depression (Lumley, Austin, & Mitchell, 2004), and patients with both depression and general medical disorders, including multiple sclerosis (Mohr & Goodkin, 1999), stroke (Hackett, Anderson, & House, 2004), and cancer (Sheard & McGuire, 1999) patients. A large number of studies have also shown that psychological treatments of depression in children and adolescents are effective but with smaller effect sizes than in adults (Weisz, McCarty, & Valeri, 2006).

A second group of meta-analyses has focused on the relative effects of specific types of treatments. Several meta-analyses found that cognitive-behavioral therapy (CBT) is more effective than other treatments (Churchill et al., 2001; Gloaguen, Cottraux, Cucherat, & Blackburn, 1998), but another meta-analysis suggested that there is no difference compared with other treatments (Wampold, Minami, Baskin, & Tierney, 2002).

Both individual and group treatments are effective in the treatment of depression (Churchill et al., 2001; McDermut, Miller, & Brown, 2001), but individual therapy may be somewhat more effective. Guided self-help and psychoeducational treatment of depression is also effective and has effect sizes comparable to those of other treatments (Cuijpers, 1997, 1998). In addition, in older adults, reminiscence and life review have effects comparable to those of other treatments (Bohlmeijer, Smit, & Cuijpers, 2003).

A third group of meta-analyses has focused on the comparative effects of psychological and pharmacological treatments. Overall, the effects of both types of treatment are comparable (Robinson, Berman, & Neimeyer, 1990), and the combination...
is somewhat more effective than treatment with
pharmacotherapy alone (Friedman et al., 2004; Panpanolla, Bollini, Tibaldi, Kupelnick, & Munizza, 2004). Although several studies suggest that combined treatments are more effective than psychological treatment alone, this has not yet been examined in a meta-analysis.

Most meta-analyses have focused on one specific subgroup of studies in order to examine the effects of one type of intervention or one target population. The last comprehensive meta-analysis of all studies on psychological treatments for depression was conducted by Robinson et al. in 1990. No comprehensive meta-analysis has attempted to examine whether characteristics of the target populations, interventions, and studies are related to their outcome using meta-regression techniques. These techniques can only be used when sufficient studies are available. The last comprehensive meta-analysis included “only” 38 controlled studies, but since then dozens of new studies have been conducted.

Meta-regression analyses may help in answering questions about psychological treatments that have not yet been addressed. For example, it has been suggested that psychological treatments are less effective in more severe depression (Elkin et al., 1995), and most clinical guidelines suggest that pharmacological treatments are more effective in severe cases of depression. Research, however, suggests that CBT is equally effective in more severe depression (DeRubeis et al., 2005). Another example is the relative efficacy of different types of treatment. As indicated previously, several meta-analyses found superior effects of CBT, but the differences between CBT and other treatments disappeared after controlling for researcher allegiance (Robinson et al., 1990) and when comparisons between CBT and placebo psychological treatments are removed (Wampold et al., 2002).

Other examples of issues that have not yet been definitively answered concern the relative efficacy of individual therapy, group therapy, and bibliotherapy; the question of whether treatments are equally effective in all (adult) age groups; and the relationship between treatment duration and outcome.

We decided, therefore, to conduct a new comprehensive meta-analysis of all studies on psychological treatment for depression, focusing on meta-regression analyses, to examine which characteristics of the interventions, patient population, and study design are associated with the effects of the studies.

Method

Identification and Selection of Studies

Studies were traced by means of several methods. First, we conducted a comprehensive literature search (from 1966 to March 2006) in which we examined 5,178 abstracts in PubMed (1,224 abstracts), PsycINFO (1,336), EMBASE (1,118) and the Cochrane Central Register of Controlled Trials (1,500). We identified these abstracts by combining terms indicative of psychological treatment (psychotherapy, psychological treatment, cognitive therapy, behavior therapy, interpersonal therapy, counseling, counselling) and depression (both medical subject headings terms and text words). For this database, we also collected the primary studies from 42 meta-analyses of psychological treatment of depression that were collected from Cuijpers and Dekker's (2005) systematic review of these meta-analyses. We retrieved the full text of 777 reports for further study.

We included studies in which (a) the effects of a psychological treatment (b) on adults (c) with a depressive disorder or an elevated level of depressive symptomatology (d) were compared with a control condition (e) in a randomized controlled trial.

Psychological treatments were defined as interventions in which verbal communication between a therapist and a client was the core element or in which a psychological treatment was written down in book format (bibliotherapy), while the client worked through it more or less independently but with some kind of personal support from a therapist (by telephone, e-mail, or otherwise). No language restrictions were applied.

We included only studies with a wait-list control group, a care-as-usual control group, or a pill-placebo control group. Studies with attention control groups were not included because some interventions that are considered attention-placebo interventions in some studies are active interventions in other studies. For example, reminiscence is used as a control condition in some studies (e.g., Arean et al., 1993) and as an active intervention in others (e.g., Serrano, Latorre, Gatz, & Montanes, 2004).

We also excluded studies on children and adolescents (<18 years of age). Studies in which the psychological intervention could not be discerned from other elements of the intervention were excluded (managed-care interventions and disease management programs), as were studies in which a standardized effect size could not be calculated (mostly because no test was performed in which the difference between experimental and control
group was examined). We also excluded studies aimed at relapse prevention.

The abstracts from the electronic bibliographic databases were screened by Pim Cuijpers, and the 777 retrieved reports were examined independently by two reviewers for possible inclusion. When the two reviewers disagreed, they discussed the differences with the third reviewer until agreement was reached.

**Data Extraction**

Studies were coded according to intervention characteristics, patient characteristics, and study characteristics.

**Characteristics of the intervention.**

- Type of intervention (seven categories: CBT according to the manual by Beck, Rush, Shaw, & Emery, 1979; other interventions that were characterized as CBT [or in which cognitive restructuring was the core element]; interpersonal psychotherapy; problem-solving therapy; activity scheduling; counseling; and other interventions [which were clearly not one of the earlier interventions])
- Format (three categories: individual therapy, group therapy, bibliotherapy)
- Therapist (four categories: a professional therapist with at least a master’s degree in psychology or psychiatry; graduate student as therapist; other professional, such as nurse, health visitor, or social worker; mixed or unknown)
- Number of sessions (continuous variable)
- Total contact time between therapist and patient (in minutes; continuous variable)

**Characteristics of the patients.**

- Target group (four categories: adults in general; elderly in general; women with postpartum depression; other populations, such as patients with somatic illnesses, student populations, and women of low socioeconomic status [SES]).
- Recruitment (four categories: open or community recruitment [the possibility to participate was published in mass media]; clinical [patients were referred from specialized mental health care or primary care settings]; screening [patients were recruited through a systematic search of a predefined population]; other [all other recruitment strategies]).
- Definition of depression (two categories: diagnosed mood disorder [included studies in which patients had to have an established mood disorder according to diagnostic criteria]; other definitions of depression included all other studies, most of which used a high score on a self-report depression questionnaire).
- Pretest Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961; Beck & Steer, 1993) score of the study population (continuous measure; we chose the BDI because this instrument was used in the largest number of studies)
- Percentage of female participants (continuous measure)
- Mean age of the study participants (continuous measure)

**Study characteristics.**

- Country (three categories: United States, United Kingdom, other countries)
- Type of control group (three categories: wait list, care as usual, placebo)
- Analyses (two categories: intention-to-treat analyses, completers-only analyses)

We also coded other characteristics of the studies, such as percentage of married participants and mean number of years of education, but these were not used in the analyses because a considerable portion of the studies (typically 50% or more) did not present these data in their reports.

**Analyses**

We calculated effect sizes by subtracting (at posttest) the average score of the control group (\(M_c\)) from the average score of the experimental group (\(M_e\)) and dividing the result by the pooled standard deviations of the experimental and control groups (\(SD_{ec}\)). An effect size of 0.5 thus indicates that the mean of the experimental group is half a standard deviation larger than the mean of the control group. Effect sizes of 0.56 to 1.2 can be assumed to be large, whereas effect sizes of 0.33 to 0.55 are moderate and 0 to 0.32 are small (Lipsey & Wilson, 1993).

In the calculations of effect sizes, we used only those instruments that explicitly measure depression. If more than one depression measure was used, the mean of the effect sizes was calculated, so that each study (or contrast group) only had one effect size. To calculate pooled mean effect sizes, we used the computer program Comprehensive Meta-analysis (version 2.2.021), developed for support in meta-analysis. Because we expected to find considerable heterogeneity, we decided to calculate mean effect sizes with the random-effects model. As an indicator of heterogeneity, we calculated the \(Q\) statistic.
also calculated the $I^2$ statistic, which is an indicator of heterogeneity in percentages. A value of 0% indicates no observed heterogeneity, and larger values show increasing heterogeneity: 25%, low; 50%, moderate; 75%, high (Higgins, Thompson, Deeks, & Altman, 2003). Because we were interested in the relationship between effect size and characteristics of the study, we only used the effect sizes indicating the difference between the experimental and control conditions at posttest; no effect sizes indicating differences at follow-up were calculated.
We first examined whether specific characteristics of the studies were related to the effect sizes in univariate analyses. For continuous variables, we used the meta-regression analyses as implemented in Comprehensive Meta-analysis (version 2.2.021). In these analyses, it is tested whether there is a significant relation between the continuous variable and the effect size, as indicated with a $Z$ value and an associated $p$ value.

For all categorical variables, we conducted the subgroup analyses as implemented in Comprehensive Meta-analysis. In these analyses, we used the mixed-effects model in which the random-effects model is used to calculate the effect size within each subgroup, while the fixed-effects model is used to test whether the effect sizes for each subgroup differ significantly from each other.

Next, we conducted a multivariate metaregression with effect size as the dependent variable. To decide which variables should be entered as predictors in this regression model, we first defined a reference group within each category of variables described in Table I (e.g., among the treatment types, we defined CBT according to the manual from Beck et al., 1979, as the reference group; among the three formats, the individual format was defined as the reference group). The reference groups for each category are presented in Table III.

To avoid collinearity among the predictors of the regression model, we first examined whether high correlations were found among the variables that could be entered into the model. We calculated the correlations between all variables described in Tables I and II (except for the variables that were used as reference groups). Two variables were found to have correlations higher than .60: The number of sessions correlated high with the total intervention time ($r = .69$), and the mean age correlated high with the dummy variable, indicating that older adults were the target population ($r = .80$). We decided to use number of sessions (not intervention time) and older adults as target population (and not mean age) as predictors in the model.

All remaining variables were entered as predictors into the regression model. Because Comprehensive Meta-analysis does not support multivariate

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>%</th>
<th>Point estimate</th>
<th>of slope*95% CI</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. sessions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2–7</td>
<td>50</td>
<td>37.0</td>
<td>0.01275</td>
<td>0.0013-0.0242</td>
<td>2.18*</td>
</tr>
<tr>
<td>8–12</td>
<td>60</td>
<td>44.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13–23</td>
<td>25</td>
<td>18.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contact time (min)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;240</td>
<td>23</td>
<td>17.0</td>
<td>0.00015</td>
<td>0.00002-0.0027</td>
<td>2.21*</td>
</tr>
<tr>
<td>240–599</td>
<td>29</td>
<td>21.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>600–839</td>
<td>23</td>
<td>17.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥840</td>
<td>32</td>
<td>23.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>28</td>
<td>20.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–49</td>
<td>6</td>
<td>4.4</td>
<td>-0.00113</td>
<td>-0.00380-0.00154</td>
<td>-0.08, ns</td>
</tr>
<tr>
<td>50–75</td>
<td>50</td>
<td>37.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>76–99</td>
<td>40</td>
<td>29.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>30</td>
<td>22.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>9</td>
<td>6.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean age (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–30</td>
<td>27</td>
<td>20.0</td>
<td>0.00490</td>
<td>0.0097-0.00882</td>
<td>2.45*</td>
</tr>
<tr>
<td>31–40</td>
<td>51</td>
<td>37.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41–60</td>
<td>24</td>
<td>17.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;60</td>
<td>16</td>
<td>11.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>17</td>
<td>12.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest BDI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20</td>
<td>13</td>
<td>9.6</td>
<td>-0.00189</td>
<td>0.02013-0.01634</td>
<td>-0.20, ns</td>
</tr>
<tr>
<td>21–25</td>
<td>31</td>
<td>23.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26–30</td>
<td>36</td>
<td>26.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;31</td>
<td>5</td>
<td>3.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>50</td>
<td>37.0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. The $Z$ value indicates whether the point estimate of the slope is significant. 95%CI = 95% confidence interval around the point estimate of the slope; BDI = Beck Depression Inventory.

*aThe point estimate of the slope indicates how much the effect size increases with one unit of the variable.

*p < .05.
regression analyses, these analyses were conducted in SPSS (version 12.0.2), and we used the instructions of Hedges and Olkin (1985) for metaregression analyses.

Results

Characteristics of the Included Studies

A total of 83 studies with 135 comparisons between a psychological intervention and a control condition were included. The total number of included participants was 5,733 (3,444 in treatment conditions and 2,289 in control conditions). A list of references of included studies is presented in the Appendix. The characteristics of the included studies and comparisons are summarized in Tables I and II. In brief, most studies examined CBT \((n/C3067)\), with considerably fewer studies examining other types of therapies, such as interpersonal psychotherapy \((n/C309)\), problem-solving therapy \((n/C309)\), activity scheduling \((n/C307)\), and counseling \((n/C308)\). In most studies, the therapy was delivered in individual format \((n/C3070)\), and most interventions were delivered by professionals \((n/C3025)\), students \((n/C3045)\), or other health professionals \((n/C3014)\). The majority of interventions had 12 or fewer sessions, with a total contact time of 14 hr.

Most studies were aimed at adults, who were recruited from the community. In about half of the studies, a diagnostic instrument was used to establish the presence of a mood disorder in participants. In virtually all studies, the majority of the participants were women, with a considerable number of studies in which only women were included. The mean age of the participants was between 31 and 60 years in most studies, and the average pretest score on the BDI was between 21 and 31 in most studies. Most studies were conducted in the United States, used a wait-list control group, and conducted completers-only analyses instead of intention-to-treat analyses.

Overall Effects

The mean effect size of all comparisons together was 0.69 (95% confidence interval [CI]: 0.60–0.79). As expected, heterogeneity was high \((Q=434.06, p<.001; I^2=69.13)\).

Univariate Analyses: Categorical Variables

In the univariate analyses of categorical variables, we found several characteristics to be significantly related to effect size. First, we found type of...
treatment to be significantly related to effect size. Counseling had a significantly smaller effect size than Beck's CBT (which means that the 95% CIs of these two categories of interventions did not overlap). Therapist background was also related to effect size; studies in which health professionals delivered the intervention had smaller effect sizes than those in which the intervention was delivered by mixed or other therapists and those that did not report the background of the therapist. Type of recruitment was also related to effect size. Studies using community recruitment had significantly higher effect sizes than those using clinical recruitment or recruitment through screening. Studies conducted in the United Kingdom had significantly lower effect sizes than those conducted in the United States and other countries. Studies using wait-list control groups had significantly higher effect sizes than those using care-as-usual or placebo control groups, and studies using care-as-usual control groups had significantly higher effect sizes than those using placebo control groups.

No indications were found that format, target population, and the issue of whether or not all patients had a diagnosed mood disorder were significantly related to the effect size.

In most of the univariate analyses, heterogeneity was high. The only analyses in which the Q statistic was not significant were those that used CBT (according to Beck's procedures) as an intervention, those that used counseling as an intervention, studies of women with postpartum depression, studies in which the intervention was delivered by other health professionals, studies in which participants were recruited through clinical settings, and studies conducted in the United Kingdom.

Univariate Analyses: Continuous Variables

Three of the five continuous variables we examined were significantly related to the effect size (Table II). The point estimate of the slope in Table II indicates how much the effect size increases with one unit of the variable. The three significant variables are number of sessions (each extra session results in an increase in effect size of 0.01275; \( p < .05 \)), contact time between therapist and patient (each extra minute of contact time results in an increase in effect size of 0.00015; \( p < .05 \)), and mean age (1 extra year results in an increase in effect size of 0.00490; \( p < .05 \)). No indication was found that the percentage of women and the BDI score at pretest were significantly related to effect size.

Multivariate Analyses

Next, we conducted a multivariate metaregression with effect size as the dependent variable. The results of these analyses are presented in Table III. Several variables were significant in these analyses: studies in which problem solving was used as an intervention had significantly higher effect sizes; studies in which the intervention was delivered by students had lower effect sizes; studies aimed at women with postpartum depression and at specific populations had higher effect sizes; studies in which participants were recruited from clinical populations and through systematic screening had lower effect sizes than those using community recruitment; and studies using care-as-usual control groups or placebo control groups had lower effect sizes than those using wait-list control groups.

Discussion

Which characteristics of psychological treatments, target populations, and studies determine the size of the effects? We examined the characteristics that are reported in most studies and found several characteristics that were related to the effect size, whereas several others were not.

First, we found that problem-solving interventions are significantly more effective than other interventions. In a recent meta-analysis of problem-solving therapies for depression (Cuijpers, van Straten, & Warmerdam, 2007), we found that these interventions did indeed have a large effect size in comparison to control groups at posttest (0.83). In this meta-analysis, we also found, however, that heterogeneity was very high, and several subgroup analyses did not give any clear indications what caused this heterogeneity. So we did find indications that problem-solving therapies are more effective than other treatments, but this is difficult to interpret when high heterogeneity remains.

In addition, we found that interventions conducted by students had lower effect sizes than those conducted by psychologists and other health professionals. This is not in agreement with an earlier meta-analysis, in which no support was found for the hypothesis that interventions conducted by professional therapists were more effective than those conducted by paraprofessionals (Den Boer, Wiersma, Russo, & van den Bosch, 2005). In that study, the term paraprofessional referred to a broad category of mental health personnel who are not qualified as psychiatrists, psychologists, social workers, or nurses and who are below a master’s-degree
level of education. In that study, however, no difference was made between types of paraprofessionals, and our study suggests that interventions conducted by students may be less effective than those conducted by professionals, which is in agreement with clinical experiences.

We also found that studies aimed at women with postpartum depression and at specific populations (e.g., patients with somatic illnesses, student populations, and women of low SES status) were more effective than those aimed at other target populations. This is not in agreement with a 2007 meta-analysis of psychological treatments of postpartum depression in which a somewhat smaller mean effect size ($d = 0.51$; Cuijpers, Brännmark, & Van Straten, 2007) was found than the overall mean effect size we found in the current study ($d = 0.69$). However, it may well be possible that controlling for characteristics of the interventions, target populations, and studies work in the opposite direction and result in higher effect sizes.

The relationship between effect size and type of recruitment we found in this metaregression study seems quite plausible. Individuals who are invited to a treatment after systematic screening are probably less motivated or their problems are considered to be less severe than in other depressed participants; otherwise, they probably would have sought help themselves at an earlier stage. Individuals who are responding to community recruitment are probably very motivated, which may improve their results compared with those who have to rely on care as usual. We found no indication that the effects of a treatment are related to the severity of the patients’ condition at pretest (which could also have explained the relationship between effect size and type of recruitment).

Equally, the relationship between effect size and type of control condition we found in the current study seems plausible. Individuals in wait-list control groups probably do not take active action to reduce their problems themselves during the waiting period because they are expecting professional help in the future. This explains why there is probably less spontaneous recovery in these individuals than in those who receive care as usual. In addition, it is known from studies on antidepressant medication that the effects of placebo medications can be considerable (Kirsch & Sapirstein, 1998).

Perhaps the most remarkable result of this study is that we found no other characteristic of the intervention, the target population, or the study itself that was significantly related to the effect size in the multivariate analyses. This has to be considered with caution and cannot be regarded as very hard evidence for the absence of a relationship between these characteristics and effect size. To show that there is no association between a characteristic and effect size, it is better to conduct a trial in which individuals are randomized to an intervention with the characteristic or an intervention without that characteristic. For example, using this method, several trials have examined the effects of individual versus group treatment (e.g., Brown & Lewinsohn, 1984; M. J. Scott & Stradling, 1990), longer versus shorter treatments (e.g., Shapiro, Barksam, Rees, Hardy, Reynolds, & Startup, 1994), and the comparative effects of different types of treatment (e.g., Elkin et al., 1989; Zeist et al., 1979).

However, despite this caution, it is remarkable that no significant association was found between the examined characteristics and several other characteristics that could have been hypothesized to be related to effect size. For example, we found no evidence that the severity of the depressive symptoms at the start of the treatment is related to the effects, although the studies varied considerably in the level of depressive symptomatology at pretest. We also found no differences between individual treatments and group treatments, and guided self-help is equally effective as other treatments. Studies in which the participants have been diagnosed with a major depressive disorder do not result in other effect sizes than studies in which participants score highly on a self-report depression measure. Treatments aimed at the elderly are equally effective as treatments aimed at younger adults, and whether or not intention-to-treat analyses have been conducted does not result in an overestimation of the effect size.

This study has several limitations. First, although we were able to include 83 studies with a total of 135 comparisons, this is still relatively small compared with the number of variables we examined. Second, several studies compared more than one psychological treatment with a control condition, and we analyzed them as though all these comparisons were independent from each other. However, the comparisons in studies with more than one comparison are clearly not independent from each other. This may have influenced the results. Third, as indicated earlier, to show that there is no relationship between a characteristic and effect size, it is better to conduct a trial in which participants are randomized to an intervention with the characteristic or an intervention without that characteristic.

Despite these limitations, however, we found strong indications that the effects of a psychological treatment are determined by only a few selected characteristics, whereas other characteristics of the intervention, the population, or the study do not seem to make much difference.
References


Appendix: Studies Included in the Meta-Analysis


Some studies are described in more than one report. In this overview, we present only the report with the main results, so that each study only has one report.