REVIEW ARTICLE

Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis

VIOLA SPEK1,2*, PIM CUIJPERS2, IVAN NYKLÍČEK1, HELEEN RIPER4, JULES KEYZER2 AND VICTOR POP1,2

1 Department of Psychology and Health, Tilburg University, The Netherlands; 2 Diagnostic Centre Eindhoven, The Netherlands; 3 Department of Clinical Psychology, Vrije Universiteit Amsterdam, The Netherlands; 4 Trimbos-instituut, Netherlands Institute of Mental Health and Addiction, The Netherlands

ABSTRACT

Background. We studied to what extent internet-based cognitive behaviour therapy (CBT) programs for symptoms of depression and anxiety are effective.

Method. A meta-analysis of 12 randomized controlled trials.

Results. The effects of internet-based CBT were compared to control conditions in 13 contrast groups with a total number of 2334 participants. A meta-analysis on treatment contrasts resulted in a moderate to large mean effect size [fixed effects analysis (FEA) \(d = 0.40\), mixed effects analysis (MEA) \(d = 0.60\)] and significant heterogeneity. Therefore, two sets of post hoc subgroup analyses were carried out. Analyses on the type of symptoms revealed that interventions for symptoms of depression had a small mean effect size (FEA \(d = 0.27\), MEA \(d = 0.32\)) and significant heterogeneity. Further analyses showed that one study could be regarded as an outlier. Analyses without this study showed a small mean effect size and moderate, non-significant heterogeneity. Interventions for anxiety had a large mean effect size (FEA and MEA \(d = 0.96\)) and very low heterogeneity. When examining the second set of subgroups, based on therapist assistance, no significant heterogeneity was found. Interventions with therapist support (\(n = 5\)) had a large mean effect size, while interventions without therapist support (\(n = 6\)) had a small mean effect size (FEA \(d = 0.24\), MEA \(d = 0.26\)).

Conclusions. In general, effect sizes of internet-based interventions for symptoms of anxiety were larger than effect sizes for depressive symptoms; however, this might be explained by differences in the amount of therapist support.

INTRODUCTION

Cognitive behaviour therapy (CBT) is a widely used and effective form of therapy for a wide range of psychological disorders, including depression and anxiety disorders (Hollon et al. 2006). In the industrialized societies, the internet has become integrated into the daily lives of a large part of the population. The number of people using the internet is still rising. Internet use has even spread among the groups that are not usually the first to use a new technology, namely women, elderly people and minority groups (Lamerichs, 2003). The expansion of the internet offers new treatment opportunities. CBT is very suitable for adaptation to a computer format. It is a structured treatment approach with the aim of developing new types of behaviour and cognition.

Internet-based CBT has advantages over traditional CBT for both clients and health care.
The anonymity and accessibility of the internet make it very suitable for offering and receiving help with psychological problems. Clients who are treated on the internet can avoid the stigma incurred by seeing a therapist (Gega et al. 2004). They can obtain treatment at any time and place, work at their own pace, and review the material as often as desired. In internet-based treatment, clients are guided by programs to work on their problems. The level of therapist involvement can vary from no assistance, or minimal therapist contact by email or telephone, to the amount of involvement as seen in classic individual therapy. Thus, it may be possible to reduce the therapist time while maintaining efficacy (Wright et al. 2005). Furthermore, it may be possible to reach people through the internet who might otherwise not receive treatment for their problems.

Because internet-based interventions seem to form a very promising line of treatment, it is important to acquire more knowledge about the effectiveness of such interventions. In the past few years, the number of randomized studies examining the effects of internet interventions on mood and anxiety disorders has grown rapidly. This study aimed to integrate the results of these studies in a meta-analysis of randomized controlled trails examining the effects of internet-based cognitive behavioural programs, with or without minimal therapist assistance, for mood and anxiety disorders.

METHOD
Criteria for considering studies for this review
Types of studies
Only randomized controlled trials were included in this review. Both published and unpublished studies were included. We included only studies that compared internet-based CBT with control groups such as waiting-lists, treatment as usual, and placebos. Studies that compared internet-based CBT with active treatments were excluded.

Types of participants
As we also included prevention studies, there were no limitations in (minimal) significance of symptoms. Only studies with participants above 18 years old were included. Studies with children or adolescents were excluded. Both clinical patients and subjects recruited from the community were included.

Types of interventions
Internet-based CBT is defined as a standardized CBT treatment that the participant works through more or less independently on the internet. Studies are included if there is no therapist support, or if there is limited support, which is defined as contact that is supportive or facilitative regarding the course material. No traditional relationship between therapist and participant is developed; the therapist only supports the working through of the standardized treatment.

We selected only internet-based treatment and excluded computer-based treatment that did not involve the internet as the study designs are too different. In studies on computer-based treatment, participants usually have to go to a particular computer to receive treatment (e.g. Marks et al. 2003; Proudfoot et al. 2003). They have to make appointments and are expected to comply with these appointments. For internet-based treatment, there is no need to make an appointment. Participants can have treatment whenever they want. This seems to be an important advantage, but there is also a disadvantage. There is no social control on using the intervention and treatment sessions can be postponed infinitely. Furthermore, participants in internet-based treatment are really on their own. In computer-based treatments, there is often someone present to help the participant with technical problems, and the amount of personal attention, however little, that is given to the subject might keep the participant more involved in the study. Internet-based studies can seem quite impersonal to participants, as we sometimes heard from people who participated in internet-based trials. These differences may substantially affect the amount of treatment that people take.

We included studies with interventions aimed at treatment or prevention of symptoms of depression or anxiety. We followed the DSM-IV classification in mood and anxiety disorders; however, we applied no restrictions regarding the inclusion criteria applied by the authors of the studies. All symptoms were measured with validated questionnaires.
Types of outcome measures
As we were interested in the effects of internet-based CBT on symptoms of depression and anxiety, we used only those instruments that explicitly measure depression or anxiety. The following types of outcome measures were included: (1) self-rating scales measuring symptoms of depression or anxiety; and (2) clinician-rated scales. Other outcome measures, measuring intermediate outcomes, such as cognition, were not included. All outcome measures included, except two used in one study (Klein & Richards, 2001), are validated instruments.

Search strategy for identification of studies
Studies were retrieved through systematic literature searches in the databases of PubMed (1990 to February 2006), PsycINFO (1990 to February 2006), and the Social Science Citation Index. Searches were conducted with key words and text words, in which words indicative of internet treatment (computer, internet) were combined with words indicative of mood or anxiety disorders or problems or treatment (mood, depression, anxiety, treatment) and CBT (cognitive therapy, computer-based therapy). Literature dating from before 1990 was excluded because the rapid changes in computers and software packages mean that internet-based treatments dating from before 1990 cannot be compared with the current treatment programs. We also checked reference lists of retrieved papers, and of earlier reviews in the field (Ritterband et al. 2003; Andersson et al. 2004; Tate & Zabinski, 2004). We contacted the corresponding authors of all included papers to obtain information about any other published or unpublished studies they were aware of.

Study selection
The retrieved papers were assessed independently on inclusion criteria by two of the authors (H.R. and V.S.) to guarantee an error-free inclusion procedure (Fig. 1). When the two disagreed on inclusion of a paper, they discussed the differences until agreement was reached.

Methodological quality assessment
The methodological quality of the studies was assessed using three basic criteria: (1)
Table 1.  Selected characteristics of the studies

<table>
<thead>
<tr>
<th>First author and year of publication</th>
<th>Recruitment; main inclusion criterion</th>
<th>Intervention: number of modules; therapist involvement</th>
<th>N</th>
<th>Outcome measures</th>
<th>Analyses</th>
<th>Control group</th>
<th>TAU allowed</th>
<th>Follow-up</th>
<th>Attrition rate (%)</th>
<th>Post-treatment comparison</th>
<th>Aim</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarke (2002) Community recruitment and clinical patients; No</td>
<td>7; None</td>
<td>299</td>
<td>CES-D</td>
<td>ITT</td>
<td>TAU</td>
<td>Yes, in both groups</td>
<td>4, 8, 16, 32 weeks</td>
<td>34</td>
<td>Intervention v. CTR</td>
<td>T</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Clarke (2005) Community recruitment and clinical patients; No</td>
<td>7; None</td>
<td>255</td>
<td>CES-D</td>
<td>ITT</td>
<td>TAU</td>
<td>Yes, in all groups</td>
<td>5, 10, 16 weeks</td>
<td>34</td>
<td>Intervention + postcard reminders v. intervention + phone reminders v. TAU</td>
<td>T</td>
<td>0.3 (mail)</td>
<td>0.2 (phone)</td>
</tr>
<tr>
<td>Christensen (2004) Community recruitment; Cut-off on KPDS</td>
<td>5; None</td>
<td>525</td>
<td>CES-D</td>
<td>ITT</td>
<td>Attention placebo</td>
<td>No</td>
<td>6 weeks</td>
<td>17</td>
<td>Intervention v. psycho education v. placebo</td>
<td>T</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>Andersson (2005) Community recruitment; Cut-off on CIDI-SF</td>
<td>5; Monitoring and feedback</td>
<td>117</td>
<td>BDI, MADRS</td>
<td>ITT</td>
<td>Participation in online discussion group</td>
<td>Yes, stable medication allowed</td>
<td>Post-treatment and 6 months</td>
<td>27</td>
<td>Intervention with participation in online discussion group v. participation in online discussion group</td>
<td>T</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>Patten (2003) Community recruitment; No</td>
<td>4; None</td>
<td>786</td>
<td>CES-D</td>
<td>Unclear</td>
<td>Psycho-education</td>
<td>Unclear</td>
<td>Post-treatment and 3 months</td>
<td>3</td>
<td>Intervention v. psycho education</td>
<td>P</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Klein (2001) Community recruitment; Panic disorder Unclear; None</td>
<td>22</td>
<td>PARF, DRF</td>
<td>CO</td>
<td>Self-monitoring</td>
<td>Unclear</td>
<td>Post-treatment</td>
<td>4</td>
<td>Intervention + self-monitoring v. self-monitoring</td>
<td>T</td>
<td>0.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klein (2006) Community recruitment; Panic disorder</td>
<td>6; Monitoring and feedback</td>
<td>55</td>
<td>Clinician rating PD and AP, no. of PA, PDSS, DASS</td>
<td>ITT</td>
<td>Therapist-assisted CBT manual and information only</td>
<td>No</td>
<td>Post-treatment and 3 months</td>
<td>16</td>
<td>Intervention v. information</td>
<td>T</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>Carlbring (2001) Community recruitment; Panic disorder</td>
<td>6; Monitoring and feedback</td>
<td>41</td>
<td>BSQ, MI, BAI</td>
<td>ITT</td>
<td>Waiting-list</td>
<td>Yes, if stable and if not CBT</td>
<td>Post-treatment</td>
<td>12</td>
<td>Intervention v. waiting-list</td>
<td>T</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Carlbring (in press) Community recruitment; Panic disorder 10; Monitoring and feedback + short weekly phone calls</td>
<td>60</td>
<td>BSQ, MI, BAI</td>
<td>ITT</td>
<td>Waiting-list</td>
<td>Yes, if stable and if not CBT</td>
<td>Post-treatment and 9 months</td>
<td>5</td>
<td>Intervention v. waiting-list</td>
<td>T</td>
<td>1.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andersson (in press) Community recruitment; Social phobia 9; Monitoring and feedback + 6 hours of group sessions</td>
<td>64</td>
<td>BAI, SPSQ, LSAS-SR, SPS</td>
<td>ITT</td>
<td>Waiting-list</td>
<td>Yes, but only stable medication</td>
<td>Post-treatment and 1 year</td>
<td>3</td>
<td>Intervention v. waiting-list</td>
<td>T</td>
<td>0.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
foreknowledge of treatment assignment is prevented; (2) assessors of outcomes are blinded for treatment assignment; (3) completeness of follow-up data (Higgins & Green, 2005). In most studies it was impossible to conceal treatment conditions from participants because of the kind of control conditions used (i.e. waiting-list), so this was not assessed.

### Treatment comparisons

Internet-based treatments with or without minimal therapist support were compared with control groups.

### Meta-analysis

First, we examined the effects of internet-based interventions compared to control conditions. We calculated effect sizes (d) by subtracting (at post-test) 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 group ($s.d_{ec}$). 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, while effect sizes of 0.33 to 0.55 are moderate, and effect sizes of 0 to 0.32 are small (Lipsey & Wilson, 2001).

In the calculations of effect sizes we only used those instruments that explicitly measure depression or anxiety (Table 1). When means and standard deviations were not reported, we used other statistics ($F$ value, $p$ value) to calculate effect sizes. If more than one measure was used, the mean of the effect sizes was calculated, so that each study (or contrast group) only had one effect size. In some studies, more than one experimental condition was compared to a control condition. In these cases, the number of subjects in the control condition was divided equally over the experimental conditions so that each subject was used only once in the meta-analyses.

To calculate pooled mean effect sizes, we used the computer program Comprehensive Meta-analysis, version 2.2.021 (Biostat, Englewood, NJ, USA).

Because it was not known before analyses whether we could expect heterogeneity among the studies, we used both the fixed effects model (FEM) and the random effects model (REM) to
calculate the pooled effect size. Heterogeneity was calculated with the \( Q \)-statistic and the \( I^2 \)-statistic. A significant \( Q \) rejects the null hypothesis of homogeneity and indicates that the variability among the effect sizes is greater than what is likely to have resulted from subject-level sampling error alone (Lipsey & Wilson, 2001). We also calculated \( I^2 \), which describes the percentage of total variation across studies that is due to heterogeneity rather than chance. An \( I^2 \) value of 25% is associated with low heterogeneity, 50% is associated with moderate heterogeneity, and 75% is associated with high heterogeneity (Higgins et al., 2003).

Post hoc subgroup analyses were conducted both with the fixed effects analysis (FEA) and the mixed effects analysis (MEA), as implemented in the Comprehensive Meta-analysis software. In the FEA, the FEM is used to calculate the effect sizes for each subgroup of studies, and also for the difference between the subgroups. In the MEA, the REM is used to calculate the effect size for each subgroup, while the FEM is used to test the difference between the subgroups of studies.

**Description of studies**

A total of 28 studies were retrieved. Of these, 16 studies did not meet the inclusion criteria and were excluded. A total of 12 trials with 2334 subjects were included. Five studies focused on depression (four on treatment and one on prevention). Seven studies were aimed at anxiety disorders (four on treatment of panic disorder, one on prevention of anxiety disorders, one on social phobia, and one on subclinical post-traumatic stress disorder). Control conditions varied from care-as-usual to an internet-based placebo condition. One of the five studies on interventions for depression aimed at prevention. The total number of subjects participating in the depression trials included was 1982. In none of the studies were subjects required to meet diagnostic criteria for a depressive disorder. In only one of the five treatment studies (Andersson et al. 2005) therapists monitored progress and gave feedback to participants; the other studies had no therapist involvement. Control conditions differed widely across studies: from care-as-usual (Clarke et al. 2002) to an attention placebo (Christensen et al. 2004). The four included studies on panic disorder had a total number of 178 participants. There was one study (Klein & Richards, 2001) in which the intervention was strictly self-help. Control conditions varied from waiting-lists to information about panic disorder (Klein et al. 2006). One study evaluated an intervention for social phobia: 64 participants were randomized to either an internet-based CBT for social phobia or to a waiting-list (Andersson et al. 2006). With two 3-hour group exposure sessions and individual feedback on homework, this is the most extensive intervention reviewed here. One trial was designed to investigate the efficacy of a preventive cognitive behavioural intervention for people at risk of developing anxiety disorders. Eighty-three participants with elevated anxiety sensitivity were randomized to either an intervention group or a waiting-list control group. One paper reported the comparison of an intervention for subclinical post-traumatic stress disorder to a waiting-list. In this study 33 participants were randomized. Selected characteristics of the included studies are summarized in Table 1.

**Methodological quality of included studies**

The quality of the included studies was reasonable to good. Foreknowledge of treatment assignment was prevented in all studies. In most studies all outcome measures were self-reported by participants. In two studies some outcome measures were not self-reported; in one study assessors of outcomes were blinded for treatment assignment (Patten, 2003), and in another it was unclear whether the assessors of outcomes were blinded for treatment condition (Klein et al. 2006). Drop-out rates varied between 3% and 34%.

**RESULTS**

A fixed effects meta-analysis on all contrasts was conducted (Fig. 2, Table 2), resulting in a mean effect size of 0.24 [95% confidence interval (CI) 0.16–0.33], while the REM resulted in a mean effect size of 0.51 (95% CI 0.28–0.74). The hypothesis of homogeneity was rejected because a significant \( Q \) value was found \((Q = 58.65, I^2 = 79.5\%)\). We examined possible sources of heterogeneity through post hoc subgroup analyses. A subgroup analysis based on the aim of the intervention (prevention or treatment) still
showed high heterogeneity among treatment studies \((n = 11, \; Q = 39.77, \; I^2 = 74.9\%)\) but not among prevention studies \((n = 2, \; Q = 1.43, \; I^2 = 30.2\%)\). Treatment studies were then further divided into two sets of subgroups: one set based on the symptoms that were treated and one set based on the inclusion of support in the interventions. These divisions are depicted in Fig. 3, for purposes of clarity prevention studies are not included in this figure.

The studies on depression \((n = 5)\) had a mean effect size of \(0.27\) \((95\% \; CI \; 0.15–0.40)\) according to the FEA and \(0.32\) \((95\% \; CI \; 0.08–0.57)\) according to the MEA. The \(Q\) value was 13.37
and $I^2$ was 70-1%, indicating considerable heterogeneity. However, further analyses showed that one study (Andersson et al. 2005) could be regarded as an outlier. Analyses without this study showed a mean effect size of 0.22 for both the FEA and the MEA (95% CI 0.09–0.35 and 0.03–0.41 respectively) and moderate, non-significant heterogeneity ($Q=5.75$, $I^2=47.8\%$).

For anxiety studies ($n=6$), both the FEA and the MEA resulted in an effect size of 0.96 (95% CI 0.69–1.24), a $Q$ value of 5.10, and an $I^2$ of 2.0%. As heterogeneity in depression studies was caused by one outlier that was also the only depression treatment with therapist support, we conducted other subgroup analyses based on therapist support (Fig. 3). These showed low heterogeneity in both subgroups: $Q=8.02$, $I^2=37.6\%$ for studies without support ($n=6$) and $Q=3.24$, $I^2=0\%$ for studies with support ($n=5$). Interventions without support had a pooled mean effect size of 0.24 (95% CI 0.11–0.37) in the FEA and 0.26 (95% CI: 0.08–0.44) in the MEA, which is small. Interventions with support had a large pooled mean effect size: 1.00 (95% CI 0.75–1.24) in both the FEA and the MEA and no heterogeneity ($I^2$ was 0).

**DISCUSSION**

When looking at all studies in this meta-analysis of internet-based CBT for symptoms of depression and anxiety, we found a moderate overall mean effect size and significant heterogeneity. Subsequently, when looking at prevention and treatment studies separately, a small effect size and non-significant heterogeneity were found for prevention studies. Treatment studies showed a large mean effect size and significant heterogeneity. Therefore, treatment studies were divided into two sets of subgroups, one based on the symptoms that were addressed and another based on the inclusion of support in the interventions. The first set of subgroup analyses showed a large mean effect size with non-significant heterogeneity for anxiety treatment. The analyses on treatment for depression showed a small mean effect size with significant heterogeneity, which was mainly explained by one outlier. After the exclusion of this study, a small mean effect size with non-significant heterogeneity was demonstrated. In the second set of subgroup analyses, treatment with support showed a large mean effect size and no
heterogeneity. Treatment without support showed a small mean effect size and non-significant heterogeneity.

A large effect for treatment with support was also found in one of the studies by Carlbring et al. (2005), in which internet-based self-help with therapist support proved to be as effective as traditional individual CBT. In this meta-analysis, the only study with a high effect size in the depression treatment studies subgroup was shown to be an internet-based intervention with therapist support.

These results suggest that it is not so much the type of problem (symptoms of depression or anxiety) that differentiates between large and small effect sizes but rather the distinction between whether support is added or not. However, because of the substantial differences in the design of the studies that were included (differences in symptoms and differences in treatment), future studies are needed to support this hypothesis.

This meta-analysis has several limitations. Because internet-based CBT is a relatively new area of research, the number of studies that met the inclusion criteria was small. This first meta-analysis included studies on interventions for symptoms of depression and anxiety, which is a fairly broad range of symptoms. Therefore, heterogeneity was found and subgroup analyses had to be carried out. As a consequence, power declined.

A second limitation is the distribution of numbers of subjects across studies. The studies on depression all had large numbers of subjects; the studies on anxiety disorders all had small numbers of subjects. This means that power differed largely across studies. Finally, studies used different inclusion criteria for participants. In only five of the 11 studies included was the presence or absence of a disorder established. Three studies had a cut-off score on a questionnaire as the main inclusion criterion. Three studies had no such inclusion criteria.

Despite these limitations, our study indicates that internet-based interventions, especially those with therapist support, are effective. More research is needed to further evaluate the effectiveness of internet-based CBT. If it can be proved that internet-based treatment is effective, it could be a very promising line of treatment, reaching people who otherwise would not receive treatment.

ACKNOWLEDGEMENTS

This study was supported by a grant from ZON-MW, the Netherlands Organization for Health Research and Development.

DECLARATION OF INTEREST

None.

REFERENCES


