The efficacy and effectiveness of online CBT
Ruwaard, J.-J.

Citation for published version (APA):
Ruwaard, J-J. (2013). The efficacy and effectiveness of online CBT
Chapter 7
General Discussion

Previous controlled trials established that Interapy, a particular form of online Cognitive Behaviour Therapy (CBT), yielded promising results in the treatment of post-traumatic stress. This raises the question whether similar positive results may be achieved with other mental health disorders, and whether online CBT performs well in routine clinical practice. In the preceding chapters, we presented several studies that were designed to examine the efficacy and effectiveness of online CBT. In this final chapter, we reflect on the aggregated evidence. What are our key findings? What are the strengths and limitations of the present studies? How do these findings inform clinical practice and future research initiatives? And finally, what general conclusions can be drawn with respect to the efficacy and effectiveness of online CBT?
Chapter 7. General Discussion

7.1 Key Findings

7.1.1 The efficacy of online CBT

To explore the wider applicability of online CBT, we developed four new protocols for the online treatment of burnout, depression, panic disorder and bulimia nervosa, and we assessed the efficacy of each protocol in a randomised controlled trial (RCT; Ruwaard et al., 2007, 2009, 2010; Ruwaard, Lange, Broeksteeg, et al., 2012; Chapters 2-5). The characteristics of these trials are listed in Table 7.1. The RCT’s included a total of $N = 456$ participants, who were randomly allocated to online CBT or a waiting-list/delayed treatment condition (in the RCT of the Bulimia Nervosa trial, bibliotherapy served as an additional experimental control). Outcome was assessed through well-validated self-report measures of symptom severity, which were administered at baseline (pre-test), immediately after treatment (post-test), and one to three years after treatment (long-term follow-up). Outcomes were analysed, on intention-to-treat basis, in terms of mean change in symptom severity over time, and in terms of clinically significant change (i.e., reliable recovery; Jacobson & Truax, 1991).

Dropout rates were encouragingly low: an (unweighted) average of 82% of the patients completed every step of treatment. Reductions in primary symptom severity were significantly larger with Interapy, in comparison to the experimental controls, as illustrated by the left forest plot in Figure 7.1(a). The standardised mean difference ($d$) in improvement between Interapy and the experimental controls ranged from $d = .7$ to $d = 1.0$, with a pooled effect size of $d = .8$ (95% CI: .6 to 1.0). These are large effects, and roughly equivalent to those of face-to-face CBT. As illustrated in Figure 7.1(b), our studies also revealed higher recovery rates with Interapy compared to those obtained with experimental controls. Across the trials, the average rate of reliable recovery was 47% in the online treatment groups, compared to 20% in the control groups, which equates to a significant odds ratio (OR) of 4.1 (95% CI: 2.4 to 7.0). One to three years after treatment, we found that treatment gains persisted.
7.1.2 The effectiveness of online CBT

To assess the effectiveness of online cognitive behavioural treatment in routine clinical practice, we examined treatment outcome of $N = 1500$ patients of the Interapy clinic (Ruwaard, Lange, Schrieken, Dolan, & Emmelkamp, 2012; chapter 6). This was an uncontrolled, pre-test/post-test study, with two follow-ups. Data were collected from unselected, consecutive electronic patient records of the clinic, which contained scores of self-report questionnaires that had been routinely administered at pre-test, post-test, and at 6-weeks and 1-year follow-up. Patients were Dutch adults with a referral for psychotherapy, who started online CBT for depression ($n = 413$), post-traumatic stress ($n = 478$), panic disorder ($n = 139$), or burnout ($n = 470$). Outcome variables were treatment adherence, primary and secondary symptom severity, and recovery rates.

Due to routine outcome measurement, post-treatment data were available for 79% of the patients. Full treatment was completed by 71%, which is comparable to reported completion rates in Dutch mental healthcare (GGZ Nederland, 2010). Symptom reductions met selected benchmarks of naturalistic studies of face-to-face CBT, and were comparable to those observed in the RCT’s, as shown in Figure 7.2. On the short-term (at post-test and six weeks follow-up), clients reported significant ($P < .001$) reductions in symptom severity, which represented a large pooled (uncontrolled) effect size of $d = 1.4$ (cf. Figure 7.3). Among treatment completers, 71% reliably improved and 52% experienced a clinically significant change (i.e., recovery). Follow-up measurements were difficult to interpret given an attrition rate of 67%. Nonetheless, available data suggest that improvements sustained up to one year after treatment.

As in the clinical trials, routine practice clients expressed high satisfaction with the treatment. Patients gave high ratings to their therapists ($M = 8.5$ on a 1-10 scale; $SD = 1.5$). Although 30% of the patients indicated that they had missed face-to-face contact during therapy, 83% evaluated online therapy as effective, and 89% would recommend online treatment to others.
Table 7.1: Characteristics of the studies in this dissertation.

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Target problem</th>
<th>Participants</th>
<th>Type</th>
<th>Experimental conditions</th>
<th>N</th>
<th>Dropout</th>
<th>Key Outcome Measures</th>
<th>Longest Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ruwaard 2007</td>
<td>Work-related stress</td>
<td>Adult Community</td>
<td>RCT</td>
<td>Online CBT Waiting list</td>
<td>177</td>
<td>62</td>
<td>Stress (DASS); Exhaustion (MBI); Depression (DASS); Anxiety (DASS)</td>
<td>3 Yrs.</td>
</tr>
<tr>
<td>Ruwaard 2009</td>
<td>Depression</td>
<td>Adult Community</td>
<td>RCT</td>
<td>Online CBT Waiting list</td>
<td>36</td>
<td>18</td>
<td>Depression (BDI, SCL-90); General psychopathology (DASS)</td>
<td>18 Mos.</td>
</tr>
<tr>
<td>Ruwaard 2010</td>
<td>Panic symptoms</td>
<td>Adult Community</td>
<td>RCT</td>
<td>Online CBT Waiting list</td>
<td>27</td>
<td>31</td>
<td>Panic symptoms(PDSS-SR); Attack frequency &amp; intensity (Panic diary)</td>
<td>3 Yrs.</td>
</tr>
<tr>
<td>Ruwaard 2012a</td>
<td>Bulimic symptoms</td>
<td>Adult Community</td>
<td>RCT</td>
<td>Online CBT Waiting list</td>
<td>35</td>
<td>35</td>
<td>Binging and purging frequency &amp; global eating disorder severity (EDE-Q); Body attitude (BAT)</td>
<td>1 Yr.</td>
</tr>
<tr>
<td>Ruwaard 2012b</td>
<td>Burnout Depression</td>
<td>Adult, Referred</td>
<td>Routine practice study</td>
<td>Online CBT Waiting list</td>
<td>1500</td>
<td>29%</td>
<td>Specific psychopathology (OLBI/DASS Stress, BDI, PDSS-SR, IES); General psychopathology (DASS)</td>
<td>1 Yr.</td>
</tr>
</tbody>
</table>

BAT: Body Attitude Test; BDI: Beck Depression Inventory (Beck et al., 1979); DASS: Depression Anxiety Stress Scales (Lovibond & Lovibond, 1995); EDE-Q: Eating Disorder Examination-Questionnaire (Fairburn & Beglin, 1994); IES: Impact of Event Scale (Horowitz et al., 1979); MB: Maslach Burnout Inventory (Maslach et al., 1996); OLBI: Oldenburgh Burnout Inventory (Demerouti et al., 2001); PDSS-SR: Panic Disorder Severity Scale, Self-rate version (Houck et al., 2002); Ruwaard 2012a: Ruwaard, Lange, Broeksteeg, et al., 2012; Ruwaard 2012b: Ruwaard, Lange, Schrieken, Dolan & Emmelkamp, 2012
7.1. Key Findings

Figure 7.1: Forest plots of between-group effect sizes in the controlled trials of Interapy online CBT for work-related stress, depression, panic symptoms and bulimic symptoms, with regard to a) the standardised difference (Cohen’s d) between the experimental groups in primary symptom severity improvements, and b) the relative odds (odds ratio) of reliable recovery. Cohen’s d is considered small when $d = .2$, medium when $d = .5$ and large when $d = .8$. Likewise, the odds ratio is small when OR = 1.5, medium when OR = 3.5 and large when OR = 9.0.
Figure 7.2: A comparison of RCT and routine practice data of pre- to post-treatment changes in primary symptom severity observed in clients, who followed online CBT. Solid lines represent mean scores of routine practice clients (Ruwaard, Lange, Schrieken, et al., 2012). Dashed lines represent mean scores of RCT participants, who completed online CBT (Ruwaard et al., 2007, 2009, 2010; Ruwaard, Lange, Broeksteeg, et al., 2012; Lange, Rietdijk, et al., 2003). Overall, treatment effects in routine practice were somewhat stronger than those in the controlled trials, because routine practice clients had more to gain from treatment. At pre-test, symptom severity levels of routine practice clients were higher in comparison to those of RCT participants.

Figure 7.3: Standardised changes in primary symptom severity in the routine practice dataset, from pre-test to one year follow-up.
7.2 Strengths and limitations

7.2.1 Strengths

**Therapist involvement.** Our studies complement a field of research that is largely focused on the reduction of therapist involvement. The vast majority of studies of internet-delivered CBT have evaluated self-help or guided self-help formats, in which therapist support is either not provided or substantially reduced. The focus on self-management is understandable, given concerns that available resources are inadequate to meet the rising demands for mental healthcare. However, reduced therapist involvement may also limit the effects of an intervention (Sharp et al., 2000; Palmqvist et al., 2007; Cuijpers et al., 2009). We examined a form of online treatment that matches manualised face-to-face CBT, in which therapist involvement is considered crucial and therefore more extensive. By doing so, we addressed a neglected area in internet intervention research.

**Large routine practice sample.** Until recently, little was known about the performance of online therapist-assisted CBT in routine clinical practice (Andersson et al., 2009). By demonstrating that the results of controlled trials of online therapist-supported CBT generalise well to clinical practice, we substantially added to the existing evidence base. To our knowledge, our study resulted in one of largest routine clinical samples in therapist-assisted online CBT research. Encouragingly, similar findings with therapist-assisted online CBT are being reported from other online virtual clinics around the world (e.g., Hilvert-Bruce, Rossouw, Wong, Sunderland, & Andrews, 2012; Sunderland, Wong, Hilvert-Bruce, & Andrews, 2012; Bergström et al., 2010). Policymakers, regulatory bodies, and insurance companies may find an argument in these results to support further implementation of online CBT.

**Acceptable adherence.** Poor adherence has been identified as one of the major challenges of online interventions. Many participants do not complete full treatment. The problem is most pronounced for self-help interventions. For instance, dropout was huge in the Farvolden et al. (2005) study of a program without therapist involvement, with as few as 12 out of 1161 participants (1%) completing the program. But guided
self-help interventions are also subject to considerable dropout. In the typical trial of internet-based guided self-help, the majority of participants do not complete the full program. Eysenbach (2005) went as far to suggest that high attrition should be considered an inherent aspect of all internet-based interventions. Our data challenge this suggestion. Treatment adherence rate was 83% in the clinical trials, and 71% in routine practice. Both figures clearly demonstrate that acceptable adherence can be achieved in internet-based treatment, which is an encouraging finding. In the absence of a formal component analysis, no definite explanation for the better adherence can be given. The protocols that we studied comprise a strong mix of cognitive behavioural procedures and techniques that enhance client motivation and the therapeutic relationship (see, for example, Lange, 2006; van der Velden et al., 2010). Future studies should cast further light on the interaction of these three factors.

**Strong effects.** Effects were evaluated against conservative criteria of outcome. We corrected for multiple testing to ensure a family-wise error rate of $P < .05$, we corrected for missing values using the last-observation-carried-forward intention-to-treat technique (which is increasingly considered to be over-conservative), and we applied the robust criteria of clinically significant change. Yet, despite these conservative criteria, the effects of online CBT on primary symptoms and general psychopathology were significant and large across the trials.

Our data are in line with the hypothesis that the effects of human-supported interventions are larger than those of unsupported interventions. The pooled controlled effect size that we found in our trials is considerably higher than the effect size of unsupported internet interventions as found in previous meta-analyses (Cuijpers et al., 2009; Andersson & Cuijpers, 2009; Spek et al., 2007).

In comparison to other human-supported online interventions (e.g., Andersson et al., 2007; Andrews et al., 2010), the present interventions appear to have similar effects. For instance, between-group effects in our depression trial were comparable to those found in a study of internet-based guided self-help by Andersson et al. (2005). This suggests that the effectiveness of internet-based interventions may be less dependent of the frequency of therapist support, as was previously shown in an Australian trial (Klein et al., 2009). This result is puzzling, since one would expect that increased adherence would result in better effects. It may not be enough to
simply compare standardized outcomes of separate studies. Differences in research methodology and sample characteristics may obscure subtle but important differences in outcome. It would be interesting to directly compare the current interventions with variants with less frequent support (e.g., Donker et al., 2009). Such comparisons should include the careful examination of the relationship between adherence and effect size.

### 7.2.2 Limitations

**No DSM-diagnoses.** Throughout our studies, we relied on symptom severity ratings obtained through validated self-report questionnaires rather than DSM-IV diagnoses. Formal diagnoses would have required in-person diagnostic interviews, which were incompatible with the goals of our research (i.e., to assess the effects of online CBT with no face-to-face contact). In effect, some included participants probably did not fully meet the diagnostic criteria for the targeted disorders. This might be considered a limitation. However, our dimensional approach also increased the external validity of the trial, since patients often present with subclinical symptoms in clinical practice. In addition, our data indicate that we included a rather severe group of participants, who suffered from chronic symptoms. A large percentage of the participants had been treated before. Face-to-face diagnostics interviews limit conclusions about the effects of online treatment per se, since it cannot be ruled out that the effects are confounded by the additional face-to-face contact. They also limit the generalisability of the results, since the required face-to-face contacts may deter those who tend to avoid treatment, because they are reticent about such contacts. In addition, these diagnostic interviews create a geographical limitation, which precludes the possibility to treat clients who live at distant locations. Telephonic diagnostic interviews may provide a future solution. There are indications that telephone interviews and face-to-face interviews yield comparable diagnostic conclusions (Cacciola et al., 1999; Aziz & Kenford, 2004; Crippa et al., 2008).

**Asymmetric post-test assessments.** The internal validity of the first two RCT’s suffered from asymmetric timing of symptom assessment between the experimental groups. We expected participants to complete treatment within the planned weeks,
but the average treatment took longer. Consequently, measurements in the experimental groups were taken some time after the measurements in the control condition. This resulted in interpretative difficulties, since the between-effects were possibly confounded by (uncontrolled) effects of the mere passage of time. In the study of online CBT for work-related stress (Chapter 2), post-hoc analyses failed to reveal a significant effect of treatment duration. This implied that the confounding effects were probably small. In the depression study (Chapter 3), corrections were necessary, since the participants in the control group reported strong reductions in depressive symptoms from pre- to post-test. To correct for this, we extrapolated the post-test scores of the control group, which resulted in more conservative estimates of the between-group effects. Although the post-hoc analyses and statistical corrections provided some resolution, it would of course had been better if measurements had been simultaneously taken in both groups. Fortunately, we were able to avoid asymmetric timing of assessments in the other RCT’s.

Uncontrolled follow-ups. The present evidence with regard to the long-term outcome of online CBT is encouraging, but needs to be corroborated in future research. Attrition at follow-up was 19% to 37% in the RCT’s, and 65% in the routine practice trial. Clients, who experienced a more favourable long-term outcome, may have been more inclined to respond to the follow-up invitations. An additional limitation of the follow-up is that it was uncontrolled, since participants in the control condition received treatment after post-test. During the period between the post-test and the follow-up, participants were subject to many influences which we could not control. Many clients reported chronic and recurrent depressive symptoms before treatment, and stable improvements up to three years after treatment. Apparently, treatment had made a lasting difference. However, these results should be interpreted with caution.

7.3 Implications for clinical practice

Suitability. Our findings suggest that online CBT is a good alternative to face-to-face CBT for adults presenting with mild to moderate-severe forms of depression, panic disorder, bulimia nervosa, post-traumatic stress, or burnout. Online CBT might be offered to clients who prefer more anonymous forms of treatment, to clients with
busy or irregular life-styles, to expats, to clients who are physically unable to attend regular face-to-face CBT, and to clients who are reluctant to follow regular mental healthcare. Clients should have basic reading and writing skills, and need to be reasonably proficient in using the computer and the internet. Education level and age are not important. Suicidal ideation, risk of psychosis and dissociation, previous hospitalisation, and alcohol- and drug abuse are risk factors that need to be assessed at intake. Concurrent medication usage is an option, when usage is stable. Residual symptoms at post-test should be checked and addressed.

**Accessibility.** Online CBT is a highly accessible treatment. In theory, it has no geographical limitation. This is perhaps best illustrated by a recent trial, in which traumatised victims of war violence in Iraq received online CBT from German therapists, who operated from Berlin (Wagner et al., 2012). In practice, however, its availability is highly dependent on local codes of conduct. Online interventions do not fit traditional healthcare systems, and raise legal, ethical, and professional issues that are only partially resolved by current guidelines. In addition, financial barriers may limit the implementation of online treatment. In 2001, the Interapy initiative was well-received. Clients, referrers, and policymakers welcomed the new online treatment option. However, while costs of face-to-face treatment were fully reimbursed by Dutch public health insurance, costs of online treatment were not. Fortunately, this changed in 2005, when health regulatory bodies officially recognised the online services of the clinic as reimbursable healthcare.

**Efficiency.** There are several ways in which online CBT saves time. First, online CBT saves time related to travelling and the scheduling of therapy sessions. Second, online CBT saves time by freeing therapists of repetitive tasks such as the administration and scoring of outcome questionnaires. In a computerised environment, these tasks can be easily automated. Third, it is important to recognise that manualised CBT is still practiced by only a minority of therapists in clinical practice (Shafran et al., 2009). Although total time invested in online CBT may be comparable to the time that is invested in manualised face-to-face CBT, it may save time by providing an attractive alternative to the less structured (and presumably less efficient) treatments that are generally provided in routine practice. Fourth, online CBT saves time because it is
Chapter 7. General Discussion

associated with brief treatment duration. In the routine practice study, duration of
treatment ranged from 8 to 22 weeks, which is considerably less than the average
duration of treatment in Dutch specialised mental healthcare (1 year; Hilderink &
Van ‘t Land, 2009).

Practicability. Online CBT provides an excellent tool to maintain treatment fidelity,
since its treatment manual is computerised and highly standardised. It has been
argued that treatment manuals enable therapists to administer treatments more
effectively, reduce idiosyncrasy in therapeutic methods, increase the therapists’ focus
on specific treatment goals and techniques, and reduce differences in the quality
of treatment delivery (Crits-Christoph et al., 1991). Our data underscore these
suggestions. In our routine practice data, only 2% of the variance in outcome was
attributable to differences between therapists. As illustrated in Figure 7.4, therapists
generally achieved similar positive outcomes.

One particular advantage of online CBT is that therapists have time to reflect on
the best possible feedback, due to the asynchronous nature of the therapeutic dialogue.
Before sending their feedback, they can consult the help files that are provided by the
treatment manual to find motivating phrases, discuss a case with colleagues, or call
for supervision, if needed. As a result, the quality of feedback can be increased. In
our trials, the treatments were provided by therapists who were relatively young and
inexperienced. Yet, the outcomes of their treatments were on a par with the outcomes
of treatments of experienced therapists. In our view, the extra time provided by the
asynchronous therapeutic dialogue played a key part in this result.

7.4 Suggestions for future research

Replication studies. Although the efficacy of the new interventions was demon-
strated, these findings should be confirmed in replication research. Each of these
interventions was tested in just one RCT. It might be argued that the aggregated
evidence of the five related studies strengthen the evidence of each individual inter-
vention. However, according to the criteria of empirically validated treatment (e.g.,
Chambless, 1993), the new interventions should be considered experimental until at
least one other group design study has been completed. Ideally, these replications will
7.4. Suggestions for future research

Comparative trials. In the clinical trials, the treatments have been evaluated exclusively by comparison with waiting-list comparison groups. Any reference to the equivalence of online CBT and face-to-face therapy is based on comparisons of effect sizes that were collected in different contexts. There is a clear need for direct comparisons between online CBT and alternative treatment options, preferably in the form of large-sample equivalence trials that are conducted in naturalistic settings.

Figure 7.4: Mean pre-test to six week follow-up changes in primary symptom severity (standardised scores) as observed for clients of the twenty most active therapists in the routine practice dataset (number of clients per therapist ranges from 21 to 151).
**Cost-effectiveness.** Modern healthcare is managed healthcare. The economic evaluation of an intervention is as important as the more traditional clinical evaluation. At present, cost-benefit comparisons between face-to-face treatment and online treatment are complicated by the limited availability of routine practice data and by unclear definitions of relevant health-economic variables. Nonetheless, scant research has shown that online treatment is cost-effective (e.g., Bergström et al., 2010; Tate et al., 2009), and it would be instructive to conduct a cost-effectiveness study for the present treatment as well.

**Pre-treatment withdrawal.** Pre-treatment withdrawal, voluntary drop-out that occurs during or even prior to the screening, demands attention. Throughout the studies, a considerable percentage of the applicants withdrew, during or even prior to the screening (40% in the routine practice study). While pre-treatment withdrawal rates are often not reported, research shows that high withdrawal rates are common, in both online and offline research (Issakidis & Andrews, 2004; Melville et al., 2010). The application of survival analysis methods, as suggested by Eysenbach (2005), might provide insight into the reasons for this withdrawal. Pre-treatment withdrawal might be caused by the ease with which one can apply for online therapy. This may result in impulsive applications (we found some indications that applicants, who started the screening without a referral, were less likely to start treatment). A second possibility is that patients withdraw from online treatment because they are unwilling to relinquish their anonymity (de Haas et al., 2009; Postel et al., 2009).

In the Netherlands, several organisations have started to provide free anonymous online CBT to clients who are unwilling to disclose their identity. Although this may be understandable given the alternative (no treatment), this raises questions about reimbursement, professional responsibility and parental consent (Lange & Ruwaard, 2010). Is it fair to provide treatment for free to some and not to others? Could providers of anonymous treatment be held responsible in case of crisis? Is it desirable that a person under 16 can start treatment when her parents are unaware of this? These questions require more discussion before wide-spread implementation of anonymous treatment should be considered.
Post-treatment booster sessions. When the effects of CBT are evaluated against conservative criteria of clinically significant change, as we did in this study, around 50% of the clients achieve high end-state functioning, i.e., recovery (Roth & Fonagy, 2005). Online CBT performed comparably in this respect. Although a large majority of clients experience a reliable change, a considerable percentage of client end treatment with residual symptoms that warrant further treatment. Currently, the options for continued online treatment are limited. The majority of the clients are referred back to their GP. For some clients with co-morbid symptoms, it makes sense to address symptoms by starting another treatment. In very rare cases, clients re-start the same treatment. However, the effects of these strategies are not clear. It may be fruitful to develop a generic booster session protocol in which clients with residual symptoms receive a series of follow-up sessions.

7.5 Conclusion

Our research strongly suggests that online therapist-assisted CBT is efficacious in the treatment of a variety of common adult mental health disorders. It is characterised by strong effects and clinically significant improvement, and by acceptable treatment adherence and high client satisfaction. Our findings further show that these effects are achieved both within the confines of the controlled trial and in the routine practice setting. Based on our results, we recommend further implementation of online CBT.