Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis

Per Carlbring, Gerhard Andersson, Pim Cuijpers, Heleen Riper & Erik Hedman-Lagerlöf

To cite this article: Per Carlbring, Gerhard Andersson, Pim Cuijpers, Heleen Riper & Erik Hedman-Lagerlöf (2018) Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis, Cognitive Behaviour Therapy, 47:1, 1-18, DOI: 10.1080/16506073.2017.1401115

To link to this article: https://doi.org/10.1080/16506073.2017.1401115

© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 07 Dec 2017.

Article views: 564

View related articles

View Crossmark data
Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis

Per Carlbring\textsuperscript{a}, Gerhard Andersson\textsuperscript{b,c}, Pim Cuijpers\textsuperscript{d}, Heleen Riper\textsuperscript{d,e,f,g} and Erik Hedman-Lagerlöf\textsuperscript{h}

\textsuperscript{a}Department of Psychology, Stockholm University, Stockholm, Sweden; \textsuperscript{b}Department of Behavioural Sciences and Learning, Swedish Institute for Disability Research, Linköping University, Linköping, Sweden; \textsuperscript{c}Department of Clinical Neuroscience, Division of Psychiatry, Karolinska Institutet, Stockholm, Sweden; \textsuperscript{d}Department of Clinical, Neuro and Developmental Psychology, Section of Clinical Psychology, Vrije Universiteit, Amsterdam, The Netherlands; \textsuperscript{e}VU University Medical Centre/GGZ inGeest, Amsterdam, The Netherlands; \textsuperscript{f}Amsterdam Public Health Research Institute, Amsterdam, The Netherlands; \textsuperscript{g}Faculty of Health Science, Telepsychiatric Unit, University of Southern Denmark, University Hospital, Odense, Denmark; \textsuperscript{h}Department of Clinical Neuroscience, Osher Center for Integrative Medicine and Division of Psychology, Karolinska Institutet, Stockholm, Sweden

\textbf{ABSTRACT}

During the last two decades, Internet-delivered cognitive behavior therapy (ICBT) has been tested in hundreds of randomized controlled trials, often with promising results. However, the control groups were often waitlisted, care-as-usual or attention control. Hence, little is known about the relative efficacy of ICBT as compared to face-to-face cognitive behavior therapy (CBT). In the present systematic review and meta-analysis, which included 1418 participants, guided ICBT for psychiatric and somatic conditions were directly compared to face-to-face CBT within the same trial. Out of the 2078 articles screened, a total of 20 studies met all inclusion criteria. Results showed a pooled effect size at post-treatment of Hedges $g = .05$ (95\% CI, $-.09$ to $.20$), indicating that ICBT and face-to-face treatment produced equivalent overall effects. Study quality did not affect outcomes. While the overall results indicate equivalence, there have been few studies of the individual psychiatric and somatic conditions so far, and for the majority, guided ICBT has not been compared against face-to-face treatment. Thus, more research, preferably with larger sample sizes, is needed to establish the general equivalence of the two treatment formats.

\section*{Introduction}

The first treatment studies on Internet-delivered cognitive behavior therapy (ICBT) were carried out in the late 1990s (Andersson, Carlbring, \& Lindefors, 2016). They were designed to mirror face-to-face treatments in terms of content and length. Since

\textbf{ARTICLE HISTORY}

Received 23 August 2017
Accepted 1 November 2017

\textbf{KEYWORDS}

guided internet-delivered cognitive behavior therapy; face-to-face therapy; anxiety and mood disorders; somatic disorders; meta-analysis

© 2017 the author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

\textbf{CONTACT} Per Carlbring \textsuperscript{a} per@carlbring.se Department of Psychology, Stockholm University, Stockholm SE-10691, Sweden.
then, more than 200 randomized controlled trials have been published, often with promising results indicating that ICBT is clinically effective when compared to controls (Andersson, Carlbring, & Hadjistavropoulos, 2017). Diagnoses targeted have typically included anxiety disorders and depression (Arnberg, Linton, Hultcrantz, Heintz, & Jonsson, 2014), but there are ample examples of subclinical problems, such as interventions for procrastination (Rozental, Forsell, Svensson, Andersson, & Carlbring, 2015) and perfectionism (Shafran et al., 2017). ICBT can be successfully adapted to an internet-delivered format for psychiatric and somatic conditions even in children and adolescents (Vigerland et al., 2016).

Treatment typically consists of a person first being screened or diagnosed in order to match a specific treatment program to the individual’s unique set of problems. For example, a patient presenting with panic disorder symptoms will receive a dedicated program specifically designed to target that particular problem. However, there are a few transdiagnostic programs in which this matching is less central (Păsărelu, Andersson, Bergman Nordgren, & Dobrean, 2017). Following the assessment phase, the patient is often allocated a therapist who, using encrypted asynchronous text messages, introduces the patient to the program and platform (cf. Vlaescu, Alasjö, Miloff, Carlbring, & Andersson, 2016). This is usually done by writing that the treatment will last for a predetermined number of weeks and that a new treatment module will be assigned each week. The modules mimic the treatment given in face-to-face therapy and consist of information and exercises (e.g. psychoeducation, thought records, and behavior experiments). Each module ends with essay questions that must be completed and sent to the therapist, who in turn will give personalized feedback on the patient’s progress. While there are variations in how much human interaction is included in the guidance (usually 1–15 min per week), research shows that the guidance of a human therapist is typically beneficial for a patient’s outcome (Baumeister, Reichler, Munzinger, & Lin, 2014). However, human guidance can possibly be, to some extent, replaced by smart computer-generated responses and automated personalized feedback (Titov et al., 2013).

In addition to ICBT’s short-term effect sizes indicating equivalence to therapist-administered therapy (Andersson & Cuijpers, 2009; Andrews, Cuijpers, Craske, McEvoy, & Titov, 2010; Cuijpers, van Straten, & Andersson, 2008), a few long-term follow-up studies have shown that the ICBT effects are maintained for as long as five years post treatment (Hedman, Furmark, et al., 2011).

In spite of the promising results in controlled trials, in which ICBT is often compared with waitlist control groups, an outstanding question has been how well guided ICBT compares against standard manualized face-to-face treatment. This was investigated in a meta-analysis by Andersson, Cuijpers, Carlbring, Riper, and Hedman (2014) that included a total of 13 studies (N = 1053) published up until June 2013. The results showed a pooled effect size at post-treatment of Hedges g = −.01 (95% CI, −.13 to .12), suggesting that ICBT and face-to-face treatment produce equivalent overall effects. However, since the field is moving forward rapidly and a considerable number of new studies have been published, an updated systematic review and meta-analysis is needed. The aim of this study was to reinvestigate the efficacy of ICBT compared to face-to-face cognitive behavior therapy (CBT) for psychiatric and somatic disorders, giving consideration to studies published in the past four years.
Methods

Design and selection of studies

This was an updated systematic review and meta-analysis of original articles investigating the effect of ICBT compared to face-to-face treatment. We used the same methods as in the previously published systematic review (Andersson et al., 2014), which meant that, in order to be included, the studies had to (a) compare therapist-guided ICBT to face-to-face treatment using a randomized controlled design; (b) use interventions aimed at the treatment of psychiatric or somatic disorders (and not, for example, prevention or mere psychoeducation); (c) compare treatments that were similar in content in both treatment conditions; (d) investigate a form of ICBT wherein the internet treatment was the main component and not a secondary complement to other therapies; (e) investigate a form of full-length face-to-face treatment; (f) report outcome data from an adult patient sample; (g) report outcomes in terms of assessment of symptoms of the target problem; and (h) be written in English. We included only studies in which there was some therapist contact during the trial.

We calculated effect sizes based on the primary outcome measure at post-treatment in each study. If no primary outcome measure was specified in the original study, a validated measure assessing target symptoms of the clinical problem was used, following the procedures generated by Thomson and Page (2007).

Search methods

To identify studies, systematic searches in PubMed (the MEDLINE database) were conducted using various search terms related to psychiatric and somatic disorders, including “depression,” “panic disorder,” “social phobia,” “social anxiety disorder,” “generalized anxiety disorder,” “obsessive-compulsive disorder,” “post-traumatic stress disorder,” “specific phobia,” “hypochondriasis,” “bulimia,” “tinnitus,” “erectile dysfunction,” “chronic pain,” and “fatigue.” Each of these search terms was combined with each of the following: “Internet,” “computer,” and “computerized.” The search filter “randomized controlled trial” was used. Reference lists in the included studies were checked for potential additional studies. We did not search for unpublished English studies. In addition, studies needed to target an adult population, excluding a study on clinically anxious adolescents aged 12–18 years and their parent (Spence et al., 2011) and on treating youth depression and anxiety (Sethi, 2013). Finally, since the Internet-based therapy should not be merely a pure self-help book albeit delivered via email, a study on perfectionism was excluded (Egan et al., 2014). In the original search (Andersson et al., 2014), there were no restrictions regarding publication date but data collection ended in June 2013. This updated search identified additional studies published between July 2013 and February 2017. Hence, this study includes results from both the previous and the new search.

Assessment of quality of the included studies

Each included study was assessed using the quality criteria proposed by the Cochrane Collaboration (Higgins & Green, 2011. Available from www.cochrane-handbook.org.).
Five dimensions were evaluated: (I) risk of selection bias due to the method for generating the randomization sequence, (II) risk of selection bias in terms of allocation concealment (i.e. due to foreknowledge of the forthcoming allocations), (III) detection bias in terms of blinding of outcome assessors, (IV) attrition bias due to incomplete outcome data, and (V) reporting bias due to selective reporting of results. The criterion for performance bias relating to the masking of participants was not used because that form of masking is not possible in the types of treatments investigated in this review. For each dimension, the statuses of the studies were rated using one of the following response options: “low risk,” “high risk,” or “unclear.” The alternative “unclear” was used when there were no data to assess the quality criterion in the original study. In studies that used self-reporting, criterion III was judged to be not applicable.

**Statistical analysis**

Data were analyzed using Cochrane Review Manager (RevMan) version 5.1.0 (Higgins & Green, 2011. Available from www.cochrane-handbook.org.). In the main meta-analyses, we assessed the effect of ICBT compared to face-to-face treatment using the standardized mean difference at post-treatment (Hedges’ $g$) as the outcome, meaning that the difference between treatments was divided by the pooled standard deviation. If both intention-to-treat and per-protocol data were presented, the former estimate was used in the meta-analysis. Estimates of treatment effects were conducted using all included studies as well as separately for each clinical disorder (e.g. depression). Potential differences in dropout rates between ICBT and face-to-face treatment were analyzed using meta-analytic logistic regression. All pooled analyses were carried out within a random effects model framework, assuming variation in true effects in the included studies and accounting for the hypothesized distribution of effects (Borenstein, Hedges, Higgins, & Rothstein, 2009). Studies were assessed for heterogeneity using $\chi^2$ and $I^2$ tests. In addition, forest plots were inspected to assess variation in effects across studies. Sensitivity analyses were conducted to assess whether study quality was related to outcome by comparing studies judged as having a low risk of bias on all five quality criteria dimensions with the studies that did not (i.e. those assessed as “unclear” or “low risk” on at least one quality criterion). Publication bias was investigated using funnel plots. Power calculations were conducted as suggested by Borenstein et al. (2009) and showed that, in order to have a power of 80% to detect a small effect size ($d = .3$), given an alpha-level of .05, 14 studies, with an average of 25 participants in each treatment arm, were needed.

**Results**

**Studies included in the review**

Of the 2078 screened studies, 20 (total $N = 1418$) met all review criteria and were included in the study (13 previous and 7 new studies). Figure 1 displays the study inclusion process. The 20 studies all investigated ICBT against some form of CBT (individual format, $n = 10$ or group format, $n = 10$). In terms of conditions studied, three of the studies targeted social anxiety disorder (Andrews, Davies, & Titov, 2011; Botella et al., 2010; Hedman, Andersson, et al., 2011), three studied panic disorder (Bergström et al., 2010; Carlbring et al., 2005;
Kiropoulos et al., 2008), four studied depressive symptoms (Andersson, Hesser, et al., 2013; Lappalainen et al., 2014; Spek et al., 2007; Wagner, Horn, & Maercker, 2014), two studied body dissatisfaction (Gollings & Paxton, 2006; Paxton, McLean, Gollings, Faulkner, & Wertheim, 2007), two studied insomnia (Blom et al., 2015; Lancee, van Straten, Morina, Kaldo, & Kamphuis, 2016), two studied tinnitus (Jasper et al., 2014; Kaldo et al., 2008), one studied male sexual dysfunction (Schover et al., 2012), one studied spider phobia (Andersson et al., 2009), one studied snake phobia (Andersson, Waara, et al., 2013), and one studied fibromyalgia (Vallejo, Ortega, Rivera, Comeche, & Vallejo-Slocker, 2015). The total number of participants was 731 in ICBT and 687 in face-to-face conditions. The studies were carried out in Australia, Finland, the Netherlands, Spain, Sweden, Switzerland, and the USA. The smallest study had 26 participants, and the largest had 201. Twelve studies recruited participants solely through self-referral while the remainder included participants from clinical samples or used a mix of self-referral and clinical recruitment. All studies were published between the years 2005 and 2016. The characteristics of each study are presented in Table 1.

**Assessment of quality**

The results of the quality assessment in terms of random sequence generation, allocation concealment, blinding of outcome assessment, data completion, and reporting bias are presented in Table 1. When blinding of outcome assessment was included, only three studies were judged as having low risk of bias in all five quality dimensions (Bergström et al., 2010;
Table 1. Characteristics of the included studies.

<table>
<thead>
<tr>
<th>Country (Reference)</th>
<th>Disorder</th>
<th>Quality criteria</th>
<th>N</th>
<th>N</th>
<th>Outcome</th>
<th>Mean (SD)</th>
<th>Mean (SD)</th>
<th>Mean (SD)</th>
<th>Mean (SD)</th>
<th>Type of therapy (number of modules/sessions; weeks of therapy)</th>
<th>N</th>
<th>N</th>
<th>INT</th>
<th>INT</th>
<th>FTF</th>
<th>FTF</th>
<th>Sample (clinical or self-ref.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAD</td>
<td>Social anxiety disorder</td>
<td>RSG (sel. bias)</td>
<td>64</td>
<td>62</td>
<td>LSAS</td>
<td>68.4 (21.0)</td>
<td>39.4 (19.9)</td>
<td>71.9 (22.9)</td>
<td>48.5 (25.0)</td>
<td>INT: CBT (15; 15 w) FTF: group CBT (15; 15 w)</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>12%</td>
<td>Y Mixed</td>
</tr>
<tr>
<td>Sweden Hedman, Andersson, et al. (2011)</td>
<td>Australia Social anxiety disorder</td>
<td>AC (sel. bias)</td>
<td>23</td>
<td>14</td>
<td>SIAAS</td>
<td>54.5 (12.4)</td>
<td>44.0 (15.9)</td>
<td>57.8 (43.9)</td>
<td>43.9 (18.7)</td>
<td>INT: CBT (6; 8 w) FTF: group CBT (7; 7 w)</td>
<td>Low risk</td>
<td>Unclear</td>
<td>NA-SR</td>
<td>Low risk</td>
<td>Low risk</td>
<td>32%</td>
<td>Y Clinical</td>
</tr>
<tr>
<td>Spain Botella et al. (2010)</td>
<td>Social anxiety disorder</td>
<td>Blinding (det. bias)</td>
<td>62</td>
<td>36</td>
<td>FPSQ</td>
<td>53.3 (14.3)</td>
<td>39.7 (15.5)</td>
<td>50.5 (11.9)</td>
<td>39.3 (13.0)</td>
<td>INT: CBT (-; 8 w) FTF: CBT(16;8 w)</td>
<td>Unclear</td>
<td>Unclear</td>
<td>NA-SR</td>
<td>High risk?</td>
<td>Low risk</td>
<td>55%</td>
<td>Y Mixed</td>
</tr>
<tr>
<td>Sweden Carlbring et al. (2005)</td>
<td>Panic disorder</td>
<td>Incomp. data (attr. bias)</td>
<td>25</td>
<td>24</td>
<td>BSQ</td>
<td>48.7 (11.7)</td>
<td>31.8 (11.6)</td>
<td>52.6 (10.8)</td>
<td>31.3 (9.1)</td>
<td>INT: CBT (10; 10 w) FTF: CBT(10;10 w)</td>
<td>Low risk</td>
<td>Low risk</td>
<td>NA-SR</td>
<td>Low risk</td>
<td>Low risk</td>
<td>12%</td>
<td>Y Self-Ref.</td>
</tr>
<tr>
<td>Sweden Bergström et al. (2010)</td>
<td>Panic disorder</td>
<td>Selectiv. report (report bias)</td>
<td>53</td>
<td>60</td>
<td>PDSS</td>
<td>14.1 (4.3)</td>
<td>6.3 (4.7)</td>
<td>14.2 (4.0)</td>
<td>6.3 (5.6)</td>
<td>INT: CBT (10;10 w) FTF: group CBT (10; 10 w)</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>18%</td>
<td>Y Mixed</td>
</tr>
<tr>
<td>Australia Kiropoulos et al. (2008)</td>
<td>Panic disorder</td>
<td>Drop-outs post</td>
<td>46</td>
<td>40</td>
<td>PDSS</td>
<td>14.9 (4.4)</td>
<td>9.9 (5.9)</td>
<td>14.8 (4.0)</td>
<td>9.2 (5.7)</td>
<td>INT: CBT (6; 12 w) FTF: CBT(12;12 w)</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>0%</td>
<td>Y Self-Ref.</td>
</tr>
<tr>
<td>Depression</td>
<td>Depressive symptoms in elderly</td>
<td>RSG (sel. bias)</td>
<td>102</td>
<td>99</td>
<td>BDI</td>
<td>19.2 (7.2)</td>
<td>12.0 (8.1)</td>
<td>17.9 (10.0)</td>
<td>11.4 (9.4)</td>
<td>INT: CBT (8;8 w) FTF: Group CBT (10; 10 w)</td>
<td>Low risk</td>
<td>Low risk</td>
<td>NA-SR</td>
<td>High risk?</td>
<td>Low risk</td>
<td>39%</td>
<td>Y Self-Ref.</td>
</tr>
<tr>
<td>The Netherlands Spek et al. (2007)</td>
<td>Depressive symptoms</td>
<td>AC (sel. bias)</td>
<td>32</td>
<td>30</td>
<td>BDI</td>
<td>23.0 (6.1)</td>
<td>12.4 (10.0)</td>
<td>23.4 (7.6)</td>
<td>12.3 (8.8)</td>
<td>INT: CBT (7;8 w) FTF: CBT (8; 8 w)</td>
<td>Low risk</td>
<td>Unclear</td>
<td>NA-SR</td>
<td>Low risk</td>
<td>Low risk</td>
<td>15%</td>
<td>Y Self-Ref.</td>
</tr>
<tr>
<td>Switzerland Wagner et al. (2014)</td>
<td>Depressive symptoms</td>
<td>Blinding (det. bias)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>Disorder</td>
<td>Sample</td>
<td>Outcome</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Type of therapy</td>
<td>RSG (sel. bias)</td>
<td>AC (sel. bias)</td>
<td>Blinding</td>
<td>Incomp. data (att. bias)</td>
<td>Selectiv. report</td>
<td>Drop-outs post ITT</td>
<td>Sample (clinical or self-ref.)</td>
<td>Sample (clinical or self-ref.)</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------</td>
<td>--------</td>
<td>---------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>------------------</td>
<td>----------------</td>
<td>----------------</td>
<td>-----------</td>
<td>--------------------------</td>
<td>-----------------</td>
<td>----------------</td>
<td>-----------------------------</td>
<td>-----------------------------</td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>Depressive symptoms</td>
<td>33</td>
<td>INT</td>
<td>23.6 (4.8)</td>
<td>13.6 (9.8)</td>
<td>24.1 (5.0)</td>
<td>17.1 (5.0)</td>
<td>INT:CBT (7; 7 w?)</td>
<td>Low risk</td>
<td>Low risk</td>
<td>NA-SR</td>
<td>Low risk</td>
<td>Low risk</td>
<td>6%</td>
<td>Y</td>
<td>Self-Ref.</td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>Depressive symptoms</td>
<td>19</td>
<td>FTF</td>
<td>20.8 (9.3)</td>
<td>10.3 (8.2)</td>
<td>23.1 (6.4)</td>
<td>9.2 (5.2)</td>
<td>INT:ACT (7; 6 w)</td>
<td>Unclear</td>
<td>Unclear</td>
<td>NA-SR</td>
<td>Low risk</td>
<td>Low risk</td>
<td>3%</td>
<td>N</td>
<td>Self-Ref.</td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>Body dissatisfaction</td>
<td>21</td>
<td>FTF</td>
<td>129.1 (27.3)</td>
<td>98.4 (35.6)</td>
<td>140.8 (37.2)</td>
<td>109.6 (47.7)</td>
<td>INT:CBT (8; 8 w)</td>
<td>Unclear</td>
<td>Unclear</td>
<td>NA-SR</td>
<td>Low risk</td>
<td>Low risk</td>
<td>17.5%</td>
<td>Y</td>
<td>Self-Ref.</td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>Body dissatisfaction, disordered eating</td>
<td>42</td>
<td>FTF</td>
<td>134.3 (22.5)</td>
<td>116.8 (35.9)</td>
<td>143.3 (28.9)</td>
<td>105.8 (34.0)</td>
<td>INT:CBT (8; 8 w)</td>
<td>Low risk</td>
<td>Low risk</td>
<td>NA-SR</td>
<td>Low risk</td>
<td>Low risk</td>
<td>26%</td>
<td>Y</td>
<td>Sel-Ref.</td>
<td></td>
</tr>
<tr>
<td>Tinnitus</td>
<td>Tinnitus</td>
<td>26</td>
<td>TRQ</td>
<td>26.4 (15.6)</td>
<td>18.0 (16.2)</td>
<td>30.0 (18.0)</td>
<td>18.6 (17.0)</td>
<td>INT:CBT (6; 6 w)</td>
<td>Low risk/unclear</td>
<td>Low risk/unclear</td>
<td>NA-SR</td>
<td>Low risk</td>
<td>Low risk</td>
<td>14%</td>
<td>Y</td>
<td>Mixed</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>Tinnitus</td>
<td>41</td>
<td>Mini-TQ</td>
<td>12.2 (4.6)</td>
<td>7.4 (5.3)</td>
<td>14.2 (4.5)</td>
<td>8.1 (4.9)</td>
<td>INT:(12–18; 10 w)</td>
<td>Low risk</td>
<td>Low risk</td>
<td>NA-SR</td>
<td>Low risk</td>
<td>Low risk</td>
<td>7%</td>
<td>Y</td>
<td>Mixed</td>
<td></td>
</tr>
<tr>
<td>Sexual problems</td>
<td>Male sexual dysfunction</td>
<td>41</td>
<td>IEF</td>
<td>27.4 (17.3)</td>
<td>31.3 (20.4)</td>
<td>26.4 (18.2)</td>
<td>34.4 (22.2)</td>
<td>INT:CBT (-, 12w)</td>
<td>Unclear</td>
<td>Unclear</td>
<td>NA-SR</td>
<td>Low risk</td>
<td>Low risk</td>
<td>20%</td>
<td>Y</td>
<td>Mixed</td>
<td></td>
</tr>
<tr>
<td>Chronic pain</td>
<td>Fibromyalgia</td>
<td>20</td>
<td>FIQ</td>
<td>56.6 (19.8)</td>
<td>57.0 (18.2)</td>
<td>68.4 (19.5)</td>
<td>58.2 (18.6)</td>
<td>INT:CBT (10;10w)</td>
<td>Low risk</td>
<td>Unclear</td>
<td>NA-SR</td>
<td>Low risk</td>
<td>Unclear</td>
<td>0%</td>
<td>Y</td>
<td>Clinical</td>
<td></td>
</tr>
<tr>
<td>Specific phobia</td>
<td>Spider phobia</td>
<td>15</td>
<td>BAT</td>
<td>10.5 (1.5)</td>
<td>11.1 (1.2)</td>
<td>11.1 (1.2)</td>
<td>11.1 (1.2)</td>
<td>INT:CBT (5;4 w)</td>
<td>Unclear</td>
<td>Unclear</td>
<td>NA-SR</td>
<td>Low risk</td>
<td>Low risk</td>
<td>10%</td>
<td>No</td>
<td>Self-Ref.</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
**Table 1. (Continued).**

<table>
<thead>
<tr>
<th>Country (Reference)</th>
<th>Disorder</th>
<th>N</th>
<th>N</th>
<th>Outcome</th>
<th>Type of therapy (number of modules/sessions; weeks of therapy)</th>
<th>Quality criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweden, Blom et al. (2015)</td>
<td>Insomnia</td>
<td>24</td>
<td>24</td>
<td>ISI</td>
<td>INT: CBT (8; 8 w) FTF: Group CBT (8; 8 w)</td>
<td>Low risk Low risk NA-SR Low risk Low risk 6% Y Self-Ref.</td>
</tr>
</tbody>
</table>

Abbreviations: INT, internet-treatment; FTF, face-to-face treatment; RSG, random sequence generation; AC, allocation concealment; Incomp., incomplete; Selectiv., selective; ITT, intention-to-treat; Sel-ref, self-referred; Sel., Selection; CBT, Cognitive Behavior Therapy; W, weeks; NA-SR, Not applicable due to self-report; LSAS, Liebowitz Social Anxiety Scale; SIAS, Social Interaction Anxiety Scale; FPSQ, Fear of Public Speaking Questionnaire; BSQ, Body Sensation Questionnaire; BSQ*, Body Shape Questionnaire; PDSS, Panic Disorder Severity Scale; BDI, Beck Depression Inventory; MADRS-S, Montgomery Åsberg Depression Rating Scale—Self-rated; TRQ, Tinnitus Reaction Questionnaire; Mini-TQ, Mini Tinnitus Questionnaire; IIEF, International Index of Erectile Function; Fiq, Fibromyalgia Impact Questionnaire; BAT, Behavioral approach test.
Hedman, Andersson, et al., 2011; Kiropoulos et al., 2008). If the criterion of blinding of outcome assessors was disregarded in the studies that assessed outcome only through self-reporting, 10 of the 20 studies were judged as having low risk of bias in all quality dimensions. Figure 2 displays the average risk of bias in the included studies. In terms of dropouts, meta-analytic logistic regression showed no significant difference between the two treatment formats (OR = .85; 95% CI, .63–1.15), indicating that dropouts did not systematically favor one treatment over the other.

**Main findings: ICBT vs. face-to-face treatment**

**All studies**

A forest plot presenting the effect sizes (g) of each study as well as the pooled between-group effect size of all studies is presented in Figure 3. Throughout the results, an effect size estimate (g) below 0 favors ICBT while an effect size above 0 represents larger effects for face-to-face CBT. The pooled between-group effect size at post-treatment across all 20 studies was $g = .05$ (95% CI, −.09 to .20), showing that ICBT and face-to-face treatment produced equivalent overall effects. Below, results are presented for each clinical disorder separately.

**Social anxiety disorder**

The pooled between-group effect size in the three studies targeting social anxiety disorder (Andrews et al., 2011; Botella et al., 2010; Hedman, Andersson, et al., 2011) was $g = -.16$ (95% CI, −.47 to .16) in non-significant favor of ICBT but indicating equivalent effects.
Panic disorder
The pooled between-group effect size in the three studies targeting panic disorder (Bergström et al., 2010; Carlbring et al., 2005; Kiropoulos et al., 2008) was $g = .05$ (95% CI, -.20 to .30), in line with the notion of equivalent effects.

Depressive symptoms
The pooled between-group effect size in the four studies targeting depressive symptoms (Andersson, Hesser, et al., 2013; Lappalainen et al., 2014; Spek et al., 2007; Wagner, Horn, & Maercker, 2014) was $g = -.02$ (95% CI, -.22 to .19), showing equivalent effects for this condition as well.

Body dissatisfaction
The pooled between-group effect size in the two studies targeting body dissatisfaction (Gollings & Paxton, 2006; Paxton et al., 2007) was $g = .07$ (95% CI, -.49 to .62), again showing largely equivalent effects.

Insomnia
The pooled between-group effect size in the two studies targeting body insomnia (Blom et al., 2015; Lancee et al., 2016) was $g = .71$ (95% CI, -.18 to 1.60), suggesting a non-significant but larger effect of face-to-face CBT compared to ICBT.

Tinnitus
The pooled effect size in the two studies targeting tinnitus (Jasper et al., 2014; Kaldo et al., 2008) was $g = -.09$ (95% CI, -.43 to .25), suggesting no difference between the formats for this condition.

Male sexual dysfunction
The pooled between-group effect size for the study targeting male sexual dysfunction (Schover et al., 2012), which used a clinical sample of patients that had been treated for prostate cancer, was $g = -.14$ (95% CI, -.58 to .29). This represents a small, non-significant effect in favor of ICBT.
Spider phobia
The effect size in the study targeting spider phobia (Andersson et al., 2009) was $g = .43$ (95% CI, −.30 to 1.15) in favor of face-to-face treatment, but given the small size of the study, the finding was not significant.

Snake phobia
The effect size in the study targeting snake phobia (Andersson, Waara, et al., 2013) was $g = .67$ (95% CI, −.17 to 1.42) in favor of face-to-face treatment, but given the small size of the study, the finding was not significant.

Fibromyalgia
The effect size in the study targeting fibromyalgia (Vallejo et al., 2015) was $g = -.06$ (95% CI, −.068 to 0.56), indicating equivalent effects.

Sensitivity analysis
In order to estimate whether there was an association between study quality and treatment effects, subgroup analyses were conducted. In the three studies judged to have low risk of bias on all five quality criteria, the pooled effect size (ICBT vs. face-to-face treatment) was $g = -.11$ (95% CI, −.42 to .21) while the corresponding effect size for the 10 other studies was $g = .10$ (95% CI, −.06 to .26), suggesting that study quality did not affect the outcomes significantly.

Publication bias
Figure 4 presents a funnel plot relating effect sizes on the primary outcomes of the studies to the standard errors of the estimates. As shown in Figure 4, effect sizes were evenly distributed around the averaged effect. Of specific interest, the lower right section of the funnel plot includes studies suggesting that there is no major bias of the pooled effect estimate. This is due to small, unpublished studies with results favoring face-to-face treatment.

Test of heterogeneity
Tests showed significant heterogeneity ($\chi^2_{(19)} = 32.91; I^2 = 42%; p = .02$). This heterogeneity was largely driven by the study of ICBT vs. face-to-face CBT for insomnia by Lancee et al. (2016). If this study was removed from the analysis, $I^2$ dropped from 42 to 0%, and there was no longer significant heterogeneity ($\chi^2_{(18)} = 16.19; p = .58$). The pooled effect size across all studies changed marginally from $g = .05$ (95% CI, −.09 to .20) to $g = -.01$ (95% CI, −.12 to .10) if this study was removed from the analysis.

Discussion
This updated meta-analysis strengthens the previous finding by Andersson et al. (2014), which indicated equivalence between the two treatment formats. High dropout and non-compliance is sometimes put forward as a major challenge for internet interventions (Christensen, Griffiths, & Farrer, 2009). However, this meta-analysis found that dropouts
did not systematically favor one treatment format over the other. In fact, the average dropout rate across the studies was 15.7%, which is perfectly in line with a recent meta-analysis on dropouts from individual psychotherapy for major depression (Cooper & Conklin, 2015).

While there were no general or specific effects in favor of either ICBT or face-to-face administrated treatment, there were a number of interesting, albeit non-significant, findings in opposite directions, depending on diagnosis. Most evident was the non-significant advantage of face-to-face treatment over ICBT in the two animal phobia trials. A possible reason for this is that the gold-standard face-to-face treatment is highly effective (Ollendick & Davis III, 2013) and also solves the problem of acquiring the phobic stimuli, which is needed for successful treatment. In face-to-face sessions, the therapist is the supplier of the snake or spider while an internet intervention only uses text and pictures as stimuli. Perhaps this will change with the advent of virtual reality (Lindner et al., 2017). The other notable difference was that, in treatment of social anxiety disorder, the results were in slight favor of ICBT. Perhaps this non-significant result can be understood in light of the possibility that the therapist herself may be a phobic object. Hence, in face-to-face therapy, the patient’s self-focus will be heightened and, thus, his ability to fully concentrate on the therapy might be hampered.

The quality of the studies varied both in terms of rigor and sample size. However, study quality did not affect outcomes significantly, although the number of studies was too small to conclude that with certainty. In addition, there was no major bias of the pooled effect estimate. The present meta-analysis had some strengths, including a consistent outcome across studies regarding the efficacy of ICBT compared to face-to-face CBT and the relatively high quality of the trials included. However, the study also had limitations. First, there was a potential problem with heterogeneity. Since that was driven by a single outlier study on insomnia (Lancee et al., 2016) that clearly favored face-to-face therapy ($d = 1.16$), we decided to present the results both with and without that study included in the analysis.
However, the relative impact of that study on the pooled effect size across all studies was negligible, with $g = .05$ changing to $g = -.01$ if it was excluded.

Second, potential negative effects, such as increased or novel symptoms, of the treatments were not assessed at all. It has been suggested that internet interventions have great potential for alleviating emotional distress, promoting mental health, and enhancing well-being, but there could also be negative effects associated with treatment (Rozental et al., 2014). Indeed, a meta-analysis found that the conditions of 5.8% of participants involved in ICBT deteriorated (Rozental, Magnusson, Boettcher, Andersson, & Carlbring, 2017). That number corresponds well with the 5–10% found in face-to-face treatments (Hansen, Lambert, & Forman, 2002) and is much lower than the 17.4% of control group participants in internet-delivered trials (Rozental et al., 2017).

Third, the treatment reviewed is based only on CBT, which makes generalizations to other treatment approaches difficult. While the great majority of current treatments are based on cognitive behavioral principles, there are other forms of internet-delivered psychotherapy, such as psychodynamic psychotherapy (Johansson et al., 2012), physical exercise (Nyström et al., 2017), and radically different forms of ICBT, including attention bias modification (Carlbring et al., 2012), problem solving therapy (Warmerdam, van Straten, Twisk, Riper, & Cuijpers, 2008), and acceptance and commitment therapy (Ivanova et al., 2016). These internet-based intervention programs were not included in the analysis since they typically do not make direct comparisons with face-to-face psychotherapy (Johansson et al., in press).

Fourth, we have compared ICBT to face-to-face therapy regardless of whether it was delivered in an individual or group setting. It could be argued that group CBT is a suboptimal comparison (Morrison, 2001), at least when it comes to patient preferences. To untangle the relative efficacy of individual and group settings, head-to-head comparisons need to be carried out (cf. Sandell et al., 2015).

Fifth, we analyzed only the primary outcome measures in the trials; we did not include secondary outcomes. Indeed, the heterogeneity of the clinical conditions included can be viewed as a problem on its own. We cannot, at this stage and with the few studies available for each condition, conclude that ICBT and face-to-face therapy are equally effective on all conditions. This does not necessarily imply that face-to-face therapy is better. For example, there are very few studies on knowledge acquisition following CBT and even fewer on ICBT (Andersson, Carlbring, Furmark, & on behalf of the SOFIE Research Group, 2012), and the therapy formats may differ in this regard. In addition, patient characteristics, such as cognitive flexibility (Lindner et al., 2016), have not been taken into account. This is potentially important since there are studies suggesting that different predictors of outcome are relevant when comparing face-to-face treatment vs. Internet treatment (Ebert et al., 2016).

Sixth, a majority of the studies recruited participants solely through self-referral or using a mixture of self-referral and clinical recruitment. It has been suggested that recruiting through sources that imply more active treatment-seeking behaviors (e.g. Google searches, viewing postings on mental health websites) results in participants with more severe depression and anxiety than those recruited through more passive sources of information (Lindner, Nyström, Hassmén, Andersson, & Carlbring, 2015).

Finally, we did not analyze the long-term effects of the treatments due to large differences in follow-up times and since drop-out were handled in dissimilar ways. This is a possible area for future research; in the type of trials included, randomization can be maintained for long time periods (Andersson, Rozental, Shafran, & Carlbring, in press).
The results of our meta-analysis are thought-provoking both from theoretical and practical points of view. In terms of theories about change in psychotherapeutic interventions, the results suggest that the role of a face-to-face therapist may not be as crucial as suggested in the literature (Wampold, 2001) for generating large treatment effects. Even if factors such as therapeutic alliance are established in guided ICBT (Andersson, Paxling, et al., 2012), they are rarely important for its outcome. Indeed, understanding what makes ICBT work is a challenge for future research as only a few studies to date have investigated mediators of outcome (e.g. Hedman et al., 2013; Hesser, Westin, & Andersson, 2014; Karyotaki et al., 2015).

In conclusion, the aim of this systematic review and meta-analysis was to collect and analyze studies in which ICBT had been directly compared with face-to-face CBT. The findings are clear in that the overall effect for the main outcomes was close to zero, indicating that the two treatment formats are equally effective in treating social anxiety disorder, panic disorder, depressive symptoms, body dissatisfaction, insomnia, tinnitus, male sexual dysfunction, spider phobia, snake phobia, and fibromyalgia.

Acknowledgements

We would like to thank the following people for assisting in the assessment of abstracts: Martin Oscarsson, Axel D’Angelo, Ina Wood, Lars Taberman, Mikael Ersson, Sandra Frööjd, and Josefin Garberg.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Per Carlbring http://orcid.org/0000-0002-2172-8813
Gerhard Andersson http://orcid.org/0000-0003-4753-6745
Pim Cuijpers http://orcid.org/0000-0001-5497-2743
Heleen Riper http://orcid.org/0000-0002-8144-8901
Erik Hedman-Lagerlöf http://orcid.org/0000-0002-7939-9848

References


