This is the published version of a paper published in *Journal of Affective Disorders*.

Citation for the original published paper (version of record):

https://doi.org/10.1016/j.jad.2017.03.018

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:  
http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-136322
Research paper

Behavioral activation versus physical activity via the internet: A randomized controlled trial

Markus B.T. Nyström, Andreas Stenling, Emma Sjöström, Gregory Neely, Philip Lindström, Peter Hassmén, Gerhard Andersson, Christopher Martell, Per Carlbring

Abstract

Background: A major problem today is that only about fifty percent of those affected by depression seeks help. One way to reach more sufferers would be by offering easily accessible internet based treatments. The purpose of this study was to compare/evaluate four therapist supported internet administered treatments.

Method/results: Two hundred eighty six participants were included. The treatment period lasted twelve weeks, consisting of the following treatments: 1) physical activity without treatment rational, 2) physical activity with treatment rational, 3) behavioral activation without treatment rational and 4) behavioral activation with treatment rational. All groups (including a control-group) showed a significant decrease in depressive symptoms. When the treatment groups were pooled and compared to the control group, there were significant differences from pretest to posttest (Hedges \( g_{av} \) treatment =1.01, control group =0.47). This held true also when each of the four treatment groups was compared to the control group, with one exception: Physical activity without treatment rationale.

Limitations: The differences between how many modules the participants completed could indicate that there are other factors than the treatments that caused the symptom reduction, however, the dose-response analysis did not detect any significant differences on account of modules completed.

Conclusions: The results support the positive effects of internet administered treatments for depression, and highlights the importance of psychoeducation, which tends to affect both the treatment outcome and the probability of remaining in treatment. These aspects need to be considered when developing and conducting new treatments for depression, since they would increase the likelihood of positive treatment outcomes.

Keywords:
Depression
Physical activity
Behavioral activation
Treat
RCT
Growth curve modeling

1. Introduction

Major depressive disorder (MDD) is one of the most common disorders and the second most common cause of disability in the world (Lopez et al., 2006). Although depression can “self-heal” (Whiteford et al., 2013), those who recover without support have an increased risk of relapse, with the time between relapses decreasing over time (Angst and Preisig, 1995). In contrast, those who have received psychological treatment run a significantly lower risk for relapses (Cuijpers et al., 2013; Steinert et al., 2014).

Depression can occur in all stages of life, but a first onset is most common in young adult years (Hengartner et al., 2016) and is often associated with a diminished quality of life (Angst, 1995). Prolonged depression can also lead to exclusion in the labor market and weakening personal finances (Simon et al., 2002). Furthermore, depression significantly increases the risk for suicide (Angst, 1995; Moussavi et al., 2007). There is also a high comorbidity with other psychological conditions, such as anxiety disorders (Strine et al., 2015).

As an example, approximately 85% of patients diagnosed with depression also suffer from anxiety, and 90% of those diagnosed with anxiety also suffered from symptoms of depression (Tiller, 2013). A future major challenge in society will therefore be to meet the escalating need for effective mental health treatments. This need becomes even clearer when considering that only about 50% of those who suffer from...
depression seek help (Angst et al., 2002). There are probably a variety of reasons why not all who suffer from depression seek help, but likely contributors are stigma and the risk for side effects from antidepressant drugs (Johansson et al., 2013). Another possible reason may be the lack of appropriately trained clinicians (e.g., Clinical Psychologists and Psychiatrists).

One way of reaching more individuals with mild to moderate symptoms of depression is to offer internet-administered psychological treatments, which in recent years have become increasingly common (Andersson, 2016). A number of meta-analyses provide strong support for the effects of various types of internet-administered treatments for depressive disorders (Andersson et al., 2014; Cuijpers et al., 2015; Jakobsen et al., 2017). Internet-administered treatments for depression also have positive side effects in terms of increased self-esteem and quality of life, similar to face-to-face interventions (Crisp et al., 2014).

Some of the most prevalent features of MDD are fatigue, inertia, and decreased desire to engage in things previously thought of as enjoyable, which in turn leads to increased withdrawal and a more inactive lifestyle. One way to reduce depressive symptoms is to break the pattern by reintroducing enjoyable activities in life (Dimidjian et al., 2011). Two different types of treatments, which have been shown to be effective in reducing depressive symptoms and target the inactivity elements of depression, are behavioral activation (BA; Cuijpers et al., 2007; Dobson et al., 2008; Mazzucchelli et al., 2009) and physical activity (PA; Blumenthal et al., 2007; Cooney et al., 2014; Dunn et al., 2005; Joseffson et al., 2014; Rethorst et al., 2009). One major difference, however, between these two treatments is that expert knowledge about psychological treatment often is regarded as a necessity when administering BA treatments. This is not considered as vital when administering PA treatments.

PA is defined as any type of muscle movement that increases energy consumption; it includes everything from regular exercise to gardening and other domestic chores (Garber et al., 2011). Low PA increases the risk of developing depressive symptoms as well as a variety of physiological disorders such as cardiovascular diseases (Barengo et al., 2004) and diabetes (Carnethon et al., 2007). Low PA can also prolong a current depressive state (an het Rot et al., 2009). A growing body of research has emerged over the last decade showing positive effects of psychological interventions aimed at increasing PA (Barbour et al., 2007; Cooney et al., 2014; Joseffson et al., 2014; Rethorst et al., 2009). The effect of PA on depression is comparable with cognitive behavior therapy (CBT) and/or anti-depressants (Blumenthal et al., 2007; Hallgren et al., 2015; Kvan et al., 2016; Rethorst et al., 2009; Rimer et al., 2012). A recent meta-analysis has even suggested that previous studies and meta-analysis may have underestimated the effects of PA on account of publication bias (Schuch et al., 2016). The exact antidepressant mechanism of PA is still unclear. Research, however, indicates that PA is associated with a number of positive psychological effects such as increased self-esteem, sense of control, feeling of success, increased sense of independence, and also the sense of belonging (Knapen et al., 2015). Researchers have not, however, reached consensus regarding the duration, frequency, and intensity of PA that has the best effect in reducing depressive symptoms. Both low and high frequency PA have shown an antidepressant effect and there is growing support for the importance of individualized PA (Nyström et al., 2015).

Another way to activate a person, but without the physical component is BA, aiming to break the vicious circle and reintroduce positive reinforcers in daily life (Dimidjian et al., 2011). BA has over the years become a well-established treatment for depression and its effects have been found to be comparable to various types of CBT and anti-depressants (Cuijpers et al., 2007; Dobson et al., 2008; Mazzucchelli et al., 2009). An additional benefit seems to be that the effect of treatment with BA lasts longer than treatment with anti-depressants (Dobson et al., 2008).

BA is, however, not a unitary concept; two of the most acknowledged and used models are Lewinsohn's (1974) and Martell's (2010). In Lewinsohn's model, the focus is on identifying pleasant events and then increasing the frequency of these, resulting in a positive change in mood (Lewinsohn, 1974). In Martell’s model, a greater emphasis is placed on understanding one's behavioral patterns and strategies, in order to be able to test new and more effective strategies to enhance one's own mood (Martell et al., 2010). Differences in efficacy, in reducing depressive symptoms between these two versions of BA have not yet, to our knowledge, been investigated.

Both PA and BA are based on concrete and observable behavior and are therefore relatively easy to administer and require less experienced therapists than other methods (Richards et al., 2016). Research further indicates that these treatments are time-efficient (Cuijpers et al., 2007). These methods thereby have the potential to be administered over the internet (Dimidjian et al., 2011), which has also been confirmed in several studies (O’Mahen et al., 2013; Rosenbaum et al., 2015).

Even though both PA and BA have been shown to be effective in both face-to-face and internet-administered treatments separately, they have not yet been directly compared. Our primary aim with this study was therefore to evaluate and compare four different internet-administered treatments for depression: PA without rationale, PA with rationale, BA without rationale, and BA with rationale (see Method and Procedure). We hypothesize that there will be a difference in antidepressant effect between participants randomized to any of the four treatment groups and those randomized to the control group. In addition we aim to investigate if there are any differences in effect between the various treatments. Since previous research has revealed high comorbidity between depression and anxiety (Strine et al., 2015; Tiller, 2013), a secondary aim was to investigate if the change in anxiety symptoms follow the same path as the proposed change in depressive symptoms, which is our hypothesis.

2. Method

2.1. Participants

Participants from Sweden were recruited between January 2013 and May 2014 through advertisements in newspapers, on various websites and through social media (e.g., Facebook and Twitter). All registration to participate was made on the study’s website. To be included in the study the participant had to meet the criteria for mild to moderate depression according to DSM-IV-TR (American Psychiatric Association, 2000), score between 15 and 35 on the Montgomery–Åsberg Depression Rating Scale – self-rated version (MADRS-S; Svanborg and Åsberg, 1994), be aged ≥ 18 years, have a computer with access to the internet, be a resident in Sweden, and be able to read and write in Swedish. Individuals were excluded if they were regarded as suicidal or severely depressed (according to MADRS-S), presently participating in any other psychological treatment, had made changes in their anti-depressant medications (or other medications that may affect mood) during the last three months, were active exercisers (exercised more than once a week) or met criteria for another primary psychiatric diagnosis. In total, 1179 people registered interest in the study and filled out a number of questionnaires on the study website. People who met the inclusion criteria were then contacted for a semi-structured diagnostic interview and further assessments. Of the 1179 who responded, 312 were considered suitable for participation and were then block randomized independently of the research team (by the project administrator), using a specially designed computer program, into one of the five groups (four treatment groups and one control group). Of these, 286 completed at least one weekly measure or posttreatment and hence could be included in the analyses. For a description of the participants see Table 1. Those who were randomized to the control group were after 12 weeks (equivalent to the treatment period) re-randomized to one of the four treatment groups, and thus everyone who was included in the study was offered treatment.
at some point. For a more detailed description of the inclusion procedure see Fig. 1. This study was pre-registered in the Clinicaltrials.gov registry (NCT01619930) and constitutes the primary analysis of the Actua trial (Carlbring et al., 2013b). The study was approved by the Regional Ethical Board in Umeå, Sweden; written informed consent was acquired from all participants.

2.2. Procedure

One of the PA-groups was provided with a rationale, the other was not. The two BA-groups differed in that one of the conditions was based on Lewinshon’s (1974) model, the other on Martell’s model (Martell et al., 2010). Henceforth, we are referring to Lewinshon’s model as BAL and to Martell’s model as BAM. This resulted in four active treatment groups: PA without rationale, PA with rationale, BAL, and BAM. All four treatment groups followed a 12-week program consisting of eight separate modules. Each week, all participants (including the controls) were asked to fill out the Patient Health Questionnaire (PHQ-9), assessing their depression symptoms. After the 12-weeks, participants were encouraged to fill in a post-treatment form. On a weekly basis, the participants communicated with, and received feedback from, their therapist by email (for a more extended description of the procedure see, Carlbring et al., 2013b). The therapists were students from the Clinical Psychologist programs final four semesters (ten in total, over five years; equivalent to Clinical Psychology Master’s students). During the project, therapists received project-specific clinical supervision on a weekly basis from the principal investigator (PC). To minimize the
effect that the individual therapist had on treatment outcome the therapists were instructed and reminded at the weekly supervision, to devote no more than 15 min per week on correspondence with each participant. Furthermore, to balance out any selective treatment-effects, all therapists were assigned to treat participants from all treatment groups.

2.3. Instruments

2.3.1. Screening

Pre-treatment depression was assessed using the Montgomery-Åsberg Depression Rating Scale (MADRS-S; Montgomery and Åsberg, 1979), a self-assessment questionnaire designed to measure the severity of depressive symptoms (Svanborg and Åsberg, 1994), and the Patient Health Questionnaire (PHQ-9, also used as the primary outcome measure described below). Anxiety was assessed with the Generalized Anxiety Disorder Screener 7 (GAD-7, used as the secondary outcome measure described below). In addition, the following instruments were also used at pre-treatment, to further assess potential participants: International Physical Activity Questionnaire (IPAQ; Hagströmer et al., 2007), and Quality of Life Inventory (Qoli; Frisch et al., 1992). During the telephone interviews a further depression assessment was performed, using the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I; First et al., 1997), as well as an assessment of motivation to participate in the project. Post-treatment the following measures were used: PHQ-9, IPAQ, GAD-7, and Qoli. Weekly assessments consisted of the PHQ-9 and GAD-7.

2.3.2. Primary outcome measure

The primary outcome measure used in this study was PHQ-9, a self-assessment questionnaire aimed at screening and diagnosing depres-

Table 1

<table>
<thead>
<tr>
<th>Variables</th>
<th>All randomized participants N =286</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Woman</td>
<td>218 (76%)</td>
</tr>
<tr>
<td>Man</td>
<td>68 (24%)</td>
</tr>
<tr>
<td>Age</td>
<td>Mean age (SD) 42.0 (13.5)</td>
</tr>
<tr>
<td></td>
<td>Min-Max 20–80</td>
</tr>
<tr>
<td>Civil status</td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>76 (26%)</td>
</tr>
<tr>
<td>In a relationship</td>
<td>186 (65%)</td>
</tr>
<tr>
<td>Divorced/ Widow / Widower</td>
<td>23 (8%)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>Living together with someone</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>182 (64%)</td>
</tr>
<tr>
<td>No</td>
<td>104 (36%)</td>
</tr>
<tr>
<td>Education Level</td>
<td></td>
</tr>
<tr>
<td>Elementary</td>
<td>9 (3%)</td>
</tr>
<tr>
<td>High school</td>
<td>100 (35%)</td>
</tr>
<tr>
<td>Graduate school</td>
<td>169 (59%)</td>
</tr>
<tr>
<td>Postgraduate</td>
<td>8 (3%)</td>
</tr>
<tr>
<td>Occupation Status</td>
<td></td>
</tr>
<tr>
<td>Working full time</td>
<td>155 (54%)</td>
</tr>
<tr>
<td>Part time</td>
<td>57 (20%)</td>
</tr>
<tr>
<td>Student</td>
<td>31 (11%)</td>
</tr>
<tr>
<td>Retired</td>
<td>17 (6%)</td>
</tr>
<tr>
<td>Unemployed</td>
<td>26 (9%)</td>
</tr>
<tr>
<td>Full- or part time on sick leave</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>21 (7%)</td>
</tr>
<tr>
<td>No</td>
<td>265 (93%)</td>
</tr>
<tr>
<td>Previous experiences of psychological treatment</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>163 (57%)</td>
</tr>
<tr>
<td>No</td>
<td>123 (43%)</td>
</tr>
</tbody>
</table>

Fig. 1. (Flow diagram of the recruitment process and treatment.)
sion (Kroenke et al., 2001). The form consists of nine questions based on the DSM-IV-TR’s diagnostic criteria for depression. Participants estimate the frequency of the various symptoms over the last fortnight (Kroenke et al., 2001). The scale ranges from 0 to 3, where 0 is “not at all” and 3 correspond to “almost every day” making the maximum score 27. The PHQ-9 has been shown to have good psychometric properties with high internal consistency (Coefficient α between 0.86–0.89; Bian et al., 2011; Kroenke et al., 2001; Löwe et al., 2004) and test-retest reliability (r =0.84–0.95; Bian et al., 2011; Kroenke et al., 2001). The instrument has also demonstrated good content and construct validity, as well as good external validity (Manea et al., 2012).

2.3.3. Secondary outcome measure

GAD-7 is a validated 7-item self-assessment instrument for screening and diagnosing generalized anxiety disorder (Spitzer et al., 2006). A total score of 8 is to be considered as the threshold for a generalized anxiety disorder. Although developed to measure generalized anxiety disorder, the measure has satisfactory psychometric properties for detecting a range of anxiety disorders (Dear et al., 2011; Kroenke et al., 2007) and has good internal consistency both when applied in primary care (α=0.92; Spitzer et al., 2006) and in clinical research settings (α=0.92; Dear et al., 2011).

2.4. Treatments

2.4.1. PA without rationale

The treatment focus for this group was purely on PA; with participants encouraged to gradually increase their degree of PA over the weekly treatment modules. Each module contained a number of worksheets in which the participants could register their activity and thus get feedback from their counselor. In order to facilitate the participant to self-monitor their PA they were mailed a pedometer at the start of the treatment.

2.4.2. PA with rationale

The treatment was similar to the treatment offered to “PA without rationale”, with the difference that a seven-page long rationale was provided. The rationale aimed at helping the participants to view their depression in a wider context and increase the understanding of the disorder. This treatment group was also mailed a pedometer at the start of the treatment.

2.4.3. BAL

This group was provided treatment based on Lewinshon’s (1974) model of behavioral activation. Treatment modules contained worksheets where the participant were asked to identify pleasant events to find positive amplifiers and then increase the frequency of these activities to achieve positive changes in their mood.

2.4.4. BAM

This group received modules based on Martell’s model of behavioral activation (Martell et al., 2010). The focus for this group was to learn to understand their behavior patterns and strategies and to test new, more effective actions to enhance their mood. This group also received weekly homework assignments, with subsequent feedback by their therapists.

2.4.5. Control group

This group had to wait twelve weeks before being randomized into a treatment group. During the wait this group received regular monitoring, but only by filling out weekly self-assessment measures of their mood (PHQ9 and GAD7). When the first twelve weeks had passed, they filled out the same post-treatment assessment as the treatment groups.

2.5. Data analytic approach

2.5.1. Growth curve modeling

Mplus version 7.4 (Muthén and Muthén, 2005; Muthén and Lin, 2011) and the robust maximum likelihood estimator (MLR) were used for estimating the multilevel growth models (MGM; also known as random-coefficients models; Hoffman, 2015). The MGMs were based on data from the weekly PHQ-9 and GAD-7 scores over the 12 weeks of the intervention. MGM are particularly useful for handling nested data structures, such as repeated measures nested within persons (Mehta and Neale, 2005). The MGM consists of fixed effects (e.g., means) and random effects (e.g., variances and covariances). The fixed effects indicate the average level (which is dependent upon the placement of the intercept) and the average rate of change over the 12 weeks, whereas the random effects are individual differences around the group means (intercept and slope variances), the relationship between the intercept and slope, and residual variances. Fixed linear and quadratic slopes were estimated to examine the form of change and random effects were estimated to examine interindividual change differences. MGMs were estimated for the entire sample, the four treatment groups combined into a single group, and for each group (i.e., the four treatment groups and the control group) separately. Barr, Levy, Scheepers, and Tily (2013) recommended keeping the number of random effects in multilevel modeling at a maximum as long as there are no convergence problems. Following Barr et al. (2013), random linear and quadratic slope factors were estimated in all models.

The overall effect of treatment was estimated by combining the four treatment groups into a single treatment group; in a second step each of the four treatment groups were compared with the control group. The overall effect of treatment was first examined by including treatment or control group as a categorical predictor of the intercept and slope factors. MGMs were also estimated with each treatment group and the control group as a categorical predictor of the intercept and slope factors.

Effect size estimates of the mean differences between week 1 and week 12 of the PHQ-9 and GAD-7 scores were estimated within each group. Following the recommendations by Lakens (2013) we reported Hedges g_m, for dependent samples.

2.5.2. Missing data

The intention-to-treat principle was employed (Streiner, 2002; Streiner and Geddes, 2001), that is, all participants who provided data for at least one weekly measure or the post treatment evaluation were included in the analyses and the full information maximum likelihood (FIML) estimation was used to handle missing data based on the assumption that the data was missing at random (MAR; Enders, 2010). Pearson’s χ² tests were performed to check for any significant differences between the groups in dropout frequency. To investigate whether there was a significant difference between participants who announced dropout and those who completed treatment, separate independent t-tests for each group, with respect to depressive symptoms, gender and age, were performed.

3. Results

3.1. Growth modeling

3.1.1. Primary outcome measure: PHQ-9

The Intraclass Correlation Coefficient (ICC) ranged from 46% to 76% across the groups indicating a high degree of between-person variability in the primary outcome variable PHQ-9.

The negative linear slope factor accompanied by a positive quadratic slope factor, a pattern observed in the entire sample as well as in all five groups, indicates that the participants decreased their depression scores over the 12 weeks but that the decrease was leveling off (i.e., decelerating) towards the end of the study period (see Fig. 2 and
The overall effect of the treatment compared to the control group was first examined by collapsing the four treatment groups into a single group and then used as a categorical predictor (treatment = 1, control group = 0); both the intercept and slope factors were included in the model. Group predicted the linear slope ($B = -0.669$, $SE = 0.198$, $p = 0.001$, 95% CI [$-1.058, -0.281$]), indicating a steeper decline in depression scores for the treatment group compared to the control group.

Secondly, we compared each of the four treatment groups with the control group. Three of the four treatment groups showed similar statistically significant effects on the linear slope factor when compared to the control group (PA with rationale versus control, $B = -0.603$, $SE = 0.260$, $p = 0.020$, 95% CI [$-1.113, -0.175$]; BAL versus control, $B = -0.853$, $SE = 0.261$, $p = 0.001$, 95% CI [$-1.365, -0.341$]; BAM versus control, $B = -0.883$, $SE = 0.266$, $p = 0.001$, 95% CI [$-1.404, -0.362$]). Being in the PA without rationale group did not, however, have a statistically significant effect on linear slope ($B = -0.375$, $SE = 0.288$, $p = 0.194$, 95% CI [$-0.940, 0.190$]).

The overall effect of the treatment compared to the control group was first examined by collapsing the four treatment groups into a single group and then used as a categorical predictor (treatment = 1, control group = 0); both the intercept and slope factors were included in the model. Group predicted the linear slope ($B = -0.669$, $SE = 0.198$, $p = 0.001$, 95% CI [$-1.058, -0.281$]), indicating a steeper decline in depression scores for the treatment group compared to the control group.

Secondly, we compared each of the four treatment groups with the control group. Three of the four treatment groups showed similar statistically significant effects on the linear slope factor when compared to the control group (PA with rationale versus control, $B = -0.603$, $SE = 0.260$, $p = 0.020$, 95% CI [$-1.113, -0.175$]; BAL versus control, $B = -0.853$, $SE = 0.261$, $p = 0.001$, 95% CI [$-1.365, -0.341$]; BAM versus control, $B = -0.883$, $SE = 0.266$, $p = 0.001$, 95% CI [$-1.404, -0.362$]). Being in the PA without rationale group did not, however, have a statistically significant effect on linear slope ($B = -0.375$, $SE = 0.288$, $p = 0.194$, 95% CI [$-0.940, 0.190$]).

Week 1 to week 12 effect sizes of the mean differences of PHQ-9 scores within each group are reported in Table 2. All four treatment groups had effect sizes of the mean difference larger than 1.00, indicating a substantive reduction in depression symptoms across the 12 weeks. The control group effect size PHQ-9 was 0.47, indicating that also the participants in the control group decreased their depression symptoms across the 12 weeks but to a lesser extent compared to the intervention groups.

3.1.2. Secondary outcome measure

**GAD-7.** The ICC ranged from 58% to 70% indicating a high degree of between-person variability in GAD-7. First, the overall effect of the treatment compared to the control group was estimated. All four treatment groups were collapsed into a single group and a categorical predictor (treatment = 1, control group = 0) of the intercept and slope factors were included in the model. Group predicted the linear slope ($B = -0.384$, $SE = 0.169$, $p = 0.023$, 95% CI [$-0.716, -0.052$]), indicating a steeper decline in anxiety scores for the treatment group compared to the control group.

Each of the four treatment groups was then compared to the control group. Being in the PA without rationale group was not associated with a statistically significant steeper linear change compared to the control group ($B = -0.106$, $SE = 0.239$, $p = 0.659$, 95% CI [$-0.575, 0.364$]). Being in the PA with rationale group was associated with a steeper linear change compared to the control group ($B = -0.368$, $SE = 0.265$, $p = 0.165$, 95% CI [$-0.888, 0.152$]), but the effect was not statistically significant. The BAL group had a steeper linear decline compared to the control group ($B = -0.536$, $SE = 0.206$, $p = 0.009$, 95% CI [$-0.939, -0.133$], as did the BAM ($B = -0.498$, $SE = 0.221$, $p = 0.024$, 95% CI [$-0.931, -0.066$]). To summarize, all treatment groups, except for the PA without rationale group, decreased their GAD-7 scores over the 12
weeks compared to the control group.

Week 1 to week 12 effect size estimates of the mean difference of GAD-7 scores ranged from 0.76 to 1.55 indicating a medium to large reduction in anxiety symptoms in the four interventions groups. Also the control group participants to some extent decreased their anxiety symptoms but the magnitude of the effect size was smaller (0.42) compared to the four treatment groups (Table 2).

### 3.2. Influence of treatment adherence

As a supplementary analysis, the association between the number of completed treatment modules and rates of change in depression over the intervention period was estimated. The linear and quadratic slope factors were regressed on the number of completed treatment modules, and the intervention period was estimated. The linear and quadratic slope factors were regressed on the number of completed treatment modules.

### 3.3. Missing data

Pearson’s χ²-test showed that there were no statistically significant differences between the five groups in number of actively announced dropouts. Neither were there any statistically significant differences between those who dropped out and those who completed the treatment, on account of gender, severity of symptoms (both for depression and anxiety), nor age.

### 4. Discussion

The primary aim of the present study was to evaluate and compare four different internet-administered treatments for depression. We hypothesized that there would be a statistically significant decrease in depressive symptoms for those randomized to a treatment compared to the control group. We also investigated whether there were any relative differences in anti-depressive effects between the respective treatment groups. A secondary aim of this study was to see if a change in anxiety symptoms follows the same path as the proposed change in depressive symptoms.

When the treatment groups were pooled and compared with the control group, the results showed that those who received treatment had a significantly larger reduction in depressive symptoms compared with the control group. This was also true when the respective treatment groups were compared to the control group, with one exception, PA without rationale, where the difference was not statistically significant. When we compared the anti-depressive effect of the four treatments to each other, there were no significant differences between the treatment groups.

The results from the corresponding, secondary outcome, analysis showed that the change in anxiety symptoms followed a similar pattern as the depressive symptoms. This was the case, both when the treatments were pooled and when the respective treatment groups were compared to the control group. The decrease in anxiety symptoms was, however, not of the same magnitude as it was for depressive symptoms, when the treatment groups were pooled and compared to the controls.

The results indicate that participants in three of the groups (PA with rationale, BAL and BAM) had a greater decline in depressive symptoms, even though not statistically significant, than PA without rationale. These results seem to suggest that some kind of psychoeducation (e.g., rationale) is important for reduction of depressive symptoms as it might provide the participants with a better understanding of their own mood. Participants in the groups with a more thorough psychoeducation (e.g., BAM) might have experienced that they received more support, since they received more material to work with, which have had an effect on treatment outcome in previous studies (Johansson and Andersson, 2012; Richards and Richardson, 2012). It is possible that this sense of support is particularly important in
internet-administered treatments and in a way compensates for the lack of physical contact and functions as a form of therapeutic alliance. Results from previous studies on the importance of therapeutic input (e.g., therapeutic alliance) in ICBT have however been mixed. Some studies have suggested that alliance may not be as important in ICBT as in face-to-face treatments of depression (Andersson et al., 2012), while some have implied that therapeutic input is an important determinant of treatment outcome (Nordgren et al., 2013; Palmqvist et al., 2007).

It is notable that all groups, including the control group, showed a statistical significant reduction in depressive symptoms between pre- and post-treatment. This could possibly be an indication of the expectation effect, just knowing that you are in line for treatment or help/acknowledgment of your situation have been shown to predict treatment outcome (Curry et al., 2006). Another possible explanation, other than regression to the mean, for the control group’s decrease in depressive symptoms, could be the so-called spontaneous remission (Whiteford et al., 2013). It could further be argued that the control group should not be seen as completely untreated, since they on a weekly basis were asked to self-assess their mood, which could be seen as an intervention in itself and possibly explain the remission in symptoms in the control group (Sharpe and Gilbert, 1998). Also worth noting is that the reduction of depressive symptoms tended to decelerate towards the end of the treatment, which suggests that the same effect could be achieved with a shorter treatment. There are a number of possible reasons for this deceleration; one might be that participants with severe depression were excluded. Perhaps more symptom-ridden individuals need longer treatments to reach the same levels as people with fewer symptoms. To substantiate this assumption further, additional studies are required. The above is worth considering when designing and developing new treatments/studies for depressed populations.

4.2. Strengths and limitations

4.2.1. Strengths

First and foremost, the large numbers of participants, for a treatment study, which enables a comparison between the aforementioned treatments, and the relatively strict inclusion criteria, which excluded everyone who had another primary diagnosis than depression (e.g., an anxiety diagnosis). Another strength of this study was the effort that was placed on reducing therapist-related differences, which have been shown to be a source of variation in previous studies (Paxling et al., 2013). A further strength is the statistical approach used, which have several beneficial aspects (e.g., handling of missing data) and have been shown to be very well suited for treatment studies in general and studies conducted over the internet in particular (Hesser, 2015). Other strengths are the clinical implications of this study, such as the efficacy of internet administered BA and PA treatments for depression, the importance of psychoeducation and sense of support, and future development of treatments for depression.

4.2.2. Limitations

One limitation is the difference between how many modules the participants completed, which means that it cannot be ruled out that something other than the treatments can account for the change in symptoms. However, the dose-response analysis performed did not indicate any statistically significant differences on account of completed modules. A second potential limitation was that no account was taken to previous depressive episodes. The assumption is, however, that the randomization process will control for potential imbalances between the groups. A further limitation was the difficulty in capturing and operationalizing the relatively subtle differences between the two BA groups, although efforts have been made to clearly describe and implement the differences. Another limitation is the importance for treatment outcome that could be due to the different therapists. Previous studies have suggested that personal characteristics of the therapist can influence the treatment outcome (Baldwin et al., 2011; Paxling et al., 2013). In this study, however, this was taken into account because the majority of therapists treated participants from all groups, and they received clear instructions to minimize the importance of individual differences. A relating possible limitation is the risk that some therapist devoted more that the recommended 15 min per week in correspondence with each participant, which could influence treatment fidelity. However, this possible risk was acknowledged and the therapists, both when they were introduced to the project as well as during each of the weekly supervisions, were continually reminded of the importance of not spending more than 15 min per week on correspondence with each participant.

5. Conclusions

It would be easy to assume, by looking at the results and effect sizes from this study, that BA is more effective in reducing depressive symptoms than PA. One should keep in mind, however, that it could also be that it is the subjective sense/notion of receiving more support that lead to a greater reduction in symptoms. The results, however, indicate that three of the four treatments have a statistically significant decreasing effect on depressive symptoms. The added benefits of PA on physical working capacity and reduction of known health risks such as diabetes mellitus and cardiovascular diseases warrant interdisciplinary studies in the future to assess both mental and bodily health effects.

Disclosure statement

The authors have declared that no conflict of interest exists.

Acknowledgments

The Actua research study is funded by a grant from the Swedish Council for Working Life and Social Research (FORTE 2011-0477).

References

M.R.T. Nyström et al.


Archiv für psychiatrie und Psychiatrische Gesamtkrankheiten (Zürich, Switzerland : 1895) 146 (1), 5–16.


