

Author's Accepted Manuscript

A systematic review and meta-analysis on the efficacy of Internet-delivered behavioral activation

Anna Huguet, Alyssa Miller, Steve Kisely, Sanjay Rao, Nelda Saadat, Patrick J McGrath



PII: S0165-0327(17)32154-7
DOI: <https://doi.org/10.1016/j.jad.2018.02.073>
Reference: JAD9619

To appear in: *Journal of Affective Disorders*

Received date: 19 October 2017
Revised date: 30 January 2018
Accepted date: 25 February 2018

Cite this article as: Anna Huguet, Alyssa Miller, Steve Kisely, Sanjay Rao, Nelda Saadat and Patrick J McGrath, A systematic review and meta-analysis on the efficacy of Internet-delivered behavioral activation, *Journal of Affective Disorders*, <https://doi.org/10.1016/j.jad.2018.02.073>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

A systematic review and meta-analysis on the efficacy of Internet-delivered behavioral activation

Anna Huguet^{1,2*}, Alyssa Miller¹, Steve Kisely³, Sanjay Rao^{4,5}, Nelda Saadat⁴, Patrick J McGrath^{1,2,6,7,8}

¹Center for Research in Family Health, IWK Health Centre, Halifax, Nova Scotia, Canada

²Department of Community Health & Epidemiology, Dalhousie University, Halifax, Nova Scotia, Canada

³School of Medicine, The University of Queensland, Brisbane, QLD, Australia

⁴Royal Ottawa Mental Health Centre, Ottawa, Ontario, Canada

⁵Department of Psychiatry, University of Ottawa, Ottawa, Ontario, Canada

⁶Department of Psychiatry, Dalhousie University, Halifax, Nova Scotia, Canada

⁷Faculty of Science, Dalhousie University, Halifax, Nova Scotia, Canada.

⁸Department of Pediatrics, Dalhousie University, Halifax, Nova Scotia, Canada;

*Corresponding author: Centre for Research in Family Health, Room K8531, 8th Floor Children's Building, IWK Health Centre, 5850/5980 University Avenue, P.O. Box 9700, Halifax, NS B3K 6R8. Phone: +1 902 470 7521; Fax: +1 902 470 6534. anna.huguet@iwk.nshealth.ca

Abstract:

Background:

Behavioral activation (BA) is an evidence-based treatment for depression which has attracted interest and started to accumulate evidence for other conditions when delivered face-to-face. Due to its parsimoniousness, it is suitable to be delivered via the Internet. The goal of this systematic review and meta-analysis was to examine evidence from randomized controlled trials (RCTs) to determine the efficacy of Internet-based BA and assess the quality of this evidence.

Methods:

Studies were identified from electronic databases (EMBASE, ISI Web of Knowledge, Medline, CINAHL, PsychINFO, Cochrane) and reference lists of included studies. Two reviewers independently screened articles for inclusion and extracted data. They assessed the quality of evidence for each outcome using The Grading of Recommendations Assessment, Development and Evaluation framework.

Results:

Nine RCTs on different forms of depression were included with 2,157 adult participants.

Random effects meta-analyses showed that in non-clinical settings, guided Internet-based BA was non-inferior to other forms of behavioral therapy and mindfulness (mainly very low to low quality evidence) and superior to physical activity (very low quality evidence), psychoeducation/treatment as usual (moderate quality evidence) and waitlist (low quality evidence) at reducing depression and anxiety outcomes at post-treatment and short follow-up.

Limitations:

The poor quality of some of the findings means that results should be cautiously interpreted.

Conclusions:

Evidence for the efficacy of Internet-based BA as a treatment for depression is promising.

However, high quality studies with longer follow-ups are needed to increase confidence in findings and determine its efficacy in clinical settings and other conditions.

Key words: Internet-based behavioral activation, systematic review, meta-analysis, efficacy, depression, other conditions

1. Introduction

Behavioral activation (BA) is a first-line treatment of depression (Parikh et al., 2016) which is also starting to be used as a low-intensity treatment for other psychological conditions (Kanter et al., 2010; Ekers et al., 2014). BA is a goal-oriented approach that aims to re-engage individuals back into their regular routines. BA focuses on the use of pleasurable activities to increase the chance of reconnecting the individual with sources of positive reinforcement, while diminishing escape or avoidance behaviors to decrease potential sources of negative reinforcement (Jacobson et al., 1996; Martell et al., 2010). BA was initially developed to treat

depression and has been shown to be effective when delivered face-to-face and compared against cognitive therapy, anti-depressant medication or control (Ekers et al., 2014; Richards et al., 2016). There is also early evidence, although less convincing, of the efficacy of BA when using traditional face-to-face delivery methods for the treatment of other conditions such as anxiety and posttraumatic stress disorder (Soleimani et al., 2015; Hopko et al., 2006; Hopko et al., 2016; Mulick et al., 2011).

Despite the increasing evidence for the efficacy of BA as a stand-alone intervention for depression and as a result, being recently emerged as a first-line treatment for depression (Parikh et al, 2016), BA is not often used. BA is oftentimes delivered as a component of cognitive behavioral therapy (CBT), which is the most established evidence-based first line treatment for many psychological conditions, including depression (Jacobson et al., 1996; Butler et al., 2006; Dragioti et al., 2017, Parikh et al, 2016). New evidence suggests that BA can become more widely chosen due to its potential to be more available than CBT. In fact, BA alone has recently been shown to be non-inferior and more cost-effective than CBT when offered as a stand-alone treatment for depression (Richards et al., 2016). BA is a more parsimonious form of treatment than CBT, making it potentially more suitable to be delivered through cost-effective means such as via the Internet (Ekers et al., 2011). This has important implications for practice as there are insufficient CBT-trained clinicians to meet demand, resulting in long wait times to see a specialist (Goldner et al., 2011; Hazell et al., 2017). However, choosing what type of psychological treatment should be considered as a first-line of treatment depends mainly on the efficacy and quality of its evidence. It is necessary to establish the efficacy of BA when delivered via the internet.

Because of BA's potential as an evidence-based low-intensity intervention that could be easily expanded to the treatment of depression and other conditions, there have been attempts to deliver it in an Internet-based format. There have been empirical studies of the efficacy of these interventions, but, to our knowledge, no systematic reviews of the area. Prior reviews have been restricted to evaluating the efficacy of BA in individuals with depression, regardless of delivery method, and only reported a summary effect, not separate effects by mode of delivery (Ekers et al., 2014; Shinohara et al., 2013; Soucy Chartier and Provencher, 2012). The relative effectiveness of different modes of delivery are thus unclear.

Therefore, the goals of this systematic review are to: (1) identify the conditions for which the efficacy of Internet-based BA has been evaluated; (2) quantitatively assess the efficacy of Internet-based BA when compared against control conditions (e.g., waiting list) or an active treatment for each of these conditions, and; (3) evaluate the overall quality of evidence of Internet-BA on the outcomes for each of these conditions.

2. Methods

A systematic review protocol was developed (available on request). We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement for reporting title, abstract, introduction, methods, results, and discussion (Moher et al., 2009). The review protocol was not registered but is available from the authors.

2.1. Eligibility criteria:

Studies were included if they: (1) were randomized controlled trials (RCT) that evaluated the efficacy of an Internet-based BA intervention, as the standalone intervention or as the main component, on any health outcome, (2) Studies published in English. Studies were excluded if

they offered BA via videoconferencing. Although this form of delivery is Internet-based, we felt this form was not sufficiently different from in-person therapy. No restrictions were imposed on publication date, publication status, type of population, population age, type of control group, assessment time points, or health outcomes considered. It is important to acknowledge that when the protocol was initially written, we were planning to include uncontrolled studies and non-randomized controlled studies that provided health outcome data in at least five individuals who had been treated with an Internet-based BA intervention as well. However, in light of our findings (we identified 9 RCTs and 3 uncontrolled studies), we decided that uncontrolled studies would not be included in our review as these studies are at high risk of bias and therefore may not yield clinically robust results about the effects of the Internet-based BA intervention.

2.2. Information sources and search:

The following electronic databases were searched from review inception until June 2017: Cochrane Database of Systematic Reviews and Health Technology Assessment Database (HTAD) via the Cochrane Database, EMBASE, ISI Web of Knowledge, Medline (via PubMed), CINAHL, and PsychINFO. The search strategy included key terms related to the type of the intervention (e.g., behavior therapy, behavioral activation) and its delivery method (e.g., Internet, Smartphone, app, web-based, online). Search strategies for each database were built and run by an experienced reference librarian (Supplement 1 presents the search strategies for each database). Reference lists of included articles and relevant reviews were manually searched.

2.3. Study selection:

Once the list of titles/abstracts was retrieved, two reviewers (AH, AM) independently screened a random set of 20% of titles/abstracts. Articles that clearly did not meet the eligibility

criteria were immediately excluded. With almost perfect agreement ($\kappa=0.895$) (Landis and Koch, 1977; Vierra and Garrett, 2005), remaining titles/abstracts were screened by only one reviewer (AM). Once the full list of titles/abstracts was screened, two independent reviewers (AH, AM) read a random set of 35% of the remaining articles in full to judge their inclusion. Reviewers recorded reasons for exclusion. With almost perfect agreement ($\kappa=0.900$) (Landis and Koch, 1977; Vierra and Garrett, 2005), remaining full articles were screened by one reviewer (AM). Discrepancies between reviewers at any level were discussed between the two reviewers and resolved. A third reviewer (SR) was consulted when reviewers were still unsure of inclusion after discussion; this occurred in one instance.

2.4. Data extraction:

A data extraction sheet was developed and tested with one randomly selected study and modified as needed. After this, two reviewers (AM, NS) independently extracted the information for all the retrieved articles to ensure reliability in use of the data extraction form. Discrepancies between reviewers were discussed between them and resolved. If an agreement could not be reached, a third reviewer (AH) was involved in the discussion.

The following data were extracted: authors, year of publication, country in which the trial was conducted, primary diagnosis, comorbid conditions if any, target participants' age, study setting (i.e., clinical setting, general population), number of arms, type of comparison group (i.e., treatment-as-usual (TAU), placebo, wait list, psychoeducation, and another active treatment), number of individuals at baseline, length of follow-up, total number of drop outs from baseline to last follow-up, duration of treatment, BA model used (i.e., Martell's model, (Martell et al., 2010), Lejuez, Hopko and Hopko's model (Lejuez et al., 2001), Lewinsohn's model (Lewinsohn

et al., 1976), and Kanter's model (Kanter et al., 2009)), guided vs self-guided BA intervention, facilitator type (i.e., professional, para-professional, student, and N/A-self-guided intervention), amount of support from the facilitator, type of application (i.e., web-based application, native app, email), treatment outcome variables, treatment outcome measures, and for each treatment outcome and each assessment time point: number of cases, mean and standard deviation for continuous variables, and number of cases and number of events for dichotomous variables. When data were not available, study authors were contacted and information was requested. Corresponding authors of 4 manuscripts were contacted, 2 of which provided the requested information.

2.5. Data analysis:

Meta-analyses were performed to quantify the pooled effect of Internet-based BA. It was planned, a priori, that inclusion would be restricted to controlled studies if there were at least three available since uncontrolled studies could introduce bias into results. We compared the effect of online-based BA against a comparator group (i.e., treatment as usual, psychoeducation or another type of treatment) on the same health outcomes at post-treatment, three-six months after treatment, six+nine months and nine+12 months after treatment. However, in light of the limited number of studies found, we performed meta-analysis if there were at least two RCTs or comparisons for any given outcome. When, within a manuscript, two different BA models were used, Martell's model was selected as the Internet-based BA group since this is the most commonly evaluated model in this review.

Review Manager version 5.2 for Windows, a statistical software package for analyzing Cochrane Collaboration systematic reviews, was used. We used standardized mean differences

(SMD) for continuous data even where studies used the same scale given findings that the SMD is more generalizable than the mean difference (Takeshima et al., 2014). The relative risk (RR) was planned to be reported for any dichotomous outcome. Where possible, intention-to-treat analyses were used.

Sensitivity analysis were performed to determine whether the following parameters could impact the results: (1) BA delivery method (i.e., email or a web-based application), (2) the use of alternative measures for the same outcome when multiple instruments were used (e.g. the Beck Depression Inventory (BDI), Hospital Anxiety-Depression Scale (HADS)), (3) the type of mood disorder (i.e., postpartum depression, complicated grief, depression with comorbid chronic medical condition), and (4) the use of different versions of the same comparison groups when assessed within the same study (i.e., physical activity, with and without rationale). All sensitivity analyses were defined a posteriori, once data were extracted.

Heterogeneity was assessed using the I^2 statistic, a measure that does not depend on the number of studies in the meta-analysis and hence has greater power to detect heterogeneity when the number of studies is small. It is calculated using the chi-squared statistic (Q) and its degrees of freedom (Higgins et al., 2008). An estimate of 50% or greater indicates possible heterogeneity, and scores of 75–100% indicate considerable heterogeneity.

The random effects model was used for all the analyses as between-study variation could not be definitely excluded even in the absence of statistical heterogeneity given the range of interventions under review.

Where there were a sufficient number of studies ($n > 10$), we intended to assess publication bias using funnel plot asymmetry.

2.6. Quality assessment:

2.6.1. Evaluating the quality of the Internet-based BA interventions:

The quality of BA interventions was based on the level of fidelity to the core components of BA models. These were derived from reviewing the literature for BA models (Kanter et al., 2009; Lejuez et al., 2001; Lewinsohn et al., 1976; Martell et al., 2010). One reviewer (AM) evaluated whether the interventions included each of the core components. This was done by reading intervention descriptions found in manuscripts and study protocols (when available). Core components for each model are briefly defined in Supplement 2.

2.6.2. Evaluating the quality of the evidence:

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework (Guyatt et al., 2011) was used to assess the overall quality of evidence for each outcome which we could estimate a pooled effect for. The GRADE approach categorizes the levels of the quality as very low, low, moderate, and high by utilizing several domains. According to this framework, ratings of quality are reduced by study limitations, inconsistency, indirectness, imprecision, and publication/reporting bias. Two reviewers (AH, AM) graded the overall quality consensually using the GRADE framework. The study limitations domain was assessed independently by two reviewers (AM, NS). Following the recommendations derived from a systematic review conducted by Zeng et al (2015), the Cochrane Risk of Bias tool (Higgins et al., 2011) was planned to be used to assess the limitations within each included randomized study.

3. Results

3.1. Study selection:

The literature search yielded 17,727 results. After screening all articles, nine studies met inclusion criteria (total n=2,157 participants). When screening full-text articles, the most common reasons for exclusion were articles assessing interventions which were not BA, and interventions which were not delivered via the Internet or were delivered via video-conferencing. Figure 1 displays a flowchart of the screening process for inclusion in the review.

3.2. Study characteristics:

Study characteristics are presented in Table 1. The following conditions were treated using Internet-based BA: depression (Carlbring et al., 2013; Ly et al., 2014, 2015; Nystrom et al., 2017), subthreshold depression (Buntrock et al., 2016) complicated grief and rumination (Eisma et al., 2015), depression with comorbid diabetes (Ebert et al., 2017) and postpartum depression (O'Mahen et al., 2013, 2014). All studies recruited the participants from non-clinical settings and all treated adults (18 years or older). The comparison group varied, with the most common being another active treatment. We did not find any studies of the use of Internet-based BA in other psychiatric conditions such as anxiety or post-traumatic stress disorder.

The most common outcomes of these nine studies was depression as assessed by standardized instruments such as the Hospital Anxiety-Depression Scale (HADS), Beck Depression Inventory (BDI-II), Edinburgh Postnatal Depression Scale (EPDS), Inventory of Complicated Grief-revised (ICG-R) or 9-item Personal Health Questionnaire (PHQ-9) in all included studies (Buntrock et al., 2016; Carlbring et al., 2013; Ebert et al., 2017; Eisma et al., 2015; Ly et al., 2014, 2015; Nystrom et al., 2017; O'Mahen et al., 2013, 2014), and anxiety as assessed by HADS, Beck Inventory of Anxiety (BAI) and Generalized Anxiety Disorder 7-item

Scale (GAD-7) in eight studies (Buntrock et al., 2016; Carlbring et al., 2013; Ebert et al., 2017; Eisma et al., 2015; Ly et al., 2014, 2015; Nystrom et al., 2017; O'Mahen et al., 2014). Less common outcomes included quality of life in five studies (Buntrock et al., 2016; Carlbring et al., 2013; Ebert et al., 2017; Ly et al., 2014, 2015), overall physical health (Buntrock et al., 2016; Ebert et al., 2017) and psychological flexibility in two studies (Ly et al., 2014, 2015), and chronic disease self-management (Ebert et al., 2017), post-traumatic stress (Eisma et al., 2015), postnatal bonding (O'Mahen et al., 2014), social adjustment (O'Mahen et al., 2014), social problem solving (Buntrock et al., 2016), insomnia (Buntrock et al., 2016) and mastery (Buntrock et al., 2016) in one study.

Study limitation (risk of bias) assessments for each RCT are reported in Figure 2. High risk of bias was found in all studies for two domains: performance bias and detection bias. Performance bias is unavoidable in studies of this nature, as it is not possible to conceal treatment group allocation from participants. Additionally, all studies in this review used self-assessment measures to collect outcomes, which is why they were deemed to be high risk for detection bias. All studies were classified as low risk of bias for selection bias and attrition bias.

3.3. Description and quality of BA interventions:

Six out of the nine interventions were solely BA, and three presented BA with another therapy component, namely problem solving therapy (PST) (Buntrock et al., 2016), PST and diabetes-specific content (Ebert et al., 2017) and acceptance and commitment therapy (ACT) (Carlbring et al., 2013).

All online treatments were guided self-help interventions; however, level of guidance varied across interventions from as much as 60 minutes per week to as little as 15 minutes per week.

Guidance was provided most commonly via email or web-based messaging. In all interventions, the facilitator who provided guidance was either a student or mental health care professional. See Table 2 for further details.

A description of the level of fidelity of each intervention to the BA model on which it was based is presented in Table 3. It was not always possible to evaluate whether the interventions actually included each of the core components due to insufficient information. A median of 73% (range 0-100%) of components could be assessed per intervention. The most common BA model used was that of Martell and colleagues (Martell et al., 2010). For interventions which used Martell's model, the core component which was most commonly incorporated into the interventions was psychoeducation.

3.4. The efficacy of Internet-based BA and the overall quality of evidence:

At post-treatment, it was possible to perform meta-analyses including data from all nine controlled studies to estimate effect sizes for depression, meta-analyses including data from eight studies to estimate effect sizes for anxiety as well as meta-analyses including data from five controlled studies to estimate effect sizes for quality of life. Studies were grouped according to whether the comparison group was: (1) a form of behavioral therapy (i.e., exposure, in-person BA, another form of Internet-based BA), (2) mindfulness, (3) physical activity) (4) psychoeducation/treatment as usual (TAU), and (5) wait-list. Psychoeducation and TAU were grouped as all studies which used TAU as a control offered some form of psychoeducation. All outcomes were continuous as opposed to dichotomous.

Figures 3 and 4 show the estimated effect sizes of Internet-based BA intervention when compared with other forms of behavioral therapy, mindfulness, physical activity, psychoeducation/TAU or wait-list for depression and anxiety at post-treatment. In both depression and anxiety, the less intensive the control intervention, the better the result for BA. For example, Internet-based BA was non-inferior to other forms of behavioral therapy but superior to physical activity, psychoeducation/TAU and wait-list. Quality of evidence was found to be very low when comparing Internet-based BA to other forms of behavior therapy and physical activity, low for mindfulness and waitlist, and moderate for psychoeducation/TAU (Table 4).

We undertook several sensitivity analyses. For Eisma et al. (2015) the effect of using alternative depressive outcomes (i.e., the Inventory of Complicated Grief-revised, the Hospital Anxiety and Depression Scale) when Internet-based BA was compared against other forms of behavioral therapy and waitlist was investigated, while for Ly (2014 and 2015) the Beck Depression Inventory 2 (BDI-II) and Patient Health Questionnaire-9 (PHQ-9) were swapped when evaluating the effect of Internet-based BA against other forms of behavioral therapy or mindfulness. For Buntrock et al (2016), we looked at the effect of using the results of the Penn State Worrying Questionnaire (PSWQ) instead of the HADS-A (HADS – Anxiety subscale) as an outcome measure for anxiety when comparing Internet-based BA against psychoeducation/TAU. The measure used did not seem to alter findings (Supplement 3). Nystrom et al. (2017) compared the effect of two different Internet-based BA interventions based on two different BA models (i.e., Martell’s model and Lewinsohn’s model) against two types of physical activity (i.e., with and without rationale) (PA) on depression and anxiety. Sensitivity analyses of using alternative combinations of the PA versus BA study arms were performed.

Additionally, the effect of using Lewinsohn's BA model instead of Martell's BA model compared to psychoeducation/TAU was explored. None of these greatly altered the results (Supplement 3). Eisma et al. (2015) delivered their Internet-based BA intervention by email rather than directly through the Internet. When this study was excluded from the meta-analyses conducted to estimate the effect of Internet-based BA against other forms of behavioral therapy and waitlist, the results generally favoured web-based BA more strongly (Figures 3 and 4). Lastly, there was little difference in depression and anxiety outcomes when considering studies which explored the effect of Internet-based BA in comparison with psychoeducation/TAU with people with postpartum depression (O'Mahen et al., 2013, 2014) or people with comorbid diabetes (Ebert et al., 2017). Sensitivity analyses of excluding these studies to leave only the study of Buntrock et al. (2016), which was conducted with people with symptoms of depression did not alter any of the results (Figures 3 and 4).

Six studies reported depression and anxiety outcomes at three to six-month follow-up (Buntrock et al., 2016; Ebert et al., 2017; Eisma et al., 2015; Ly et al., 2014, 2015; O'Mahen et al., 2014). In terms of depression, we found low level of evidence that Internet-based BA was non-inferior to other forms of behavioral therapy ($n=114$, $k=2$, $SMD= 0.09$ [95% CI -0.19, 0.36]) and mindfulness ($n=81$, $k=1$, $SMD=-0.27$ [95% CI -0.71, 0.17]) and moderate levels of evidence that Internet-based BA was superior to psychoeducation/TAU ($n=720$, $k=3$, $SMD=-0.68$ [95% CI -0.83, -0.53]). There was low level of evidence coming from one study which compared Internet-based BA with wait-list controls that found no difference in outcomes ($n=21$, $k=1$, $SMD=-0.58$ [95% CI -1.22, 0.07]). However, this was the paper by Eisma et al. (2015) who delivered their intervention by email rather than directly through the Internet.

Identical results were found for anxiety. For instance, there was a moderate level of evidence that Internet-based BA was as effective as other forms of behavioral therapy ($n=116$, $k=2$, $SMD=-0.48$ [95% CI [-0.79, -0.17]]) and achieved significantly better outcomes than psychoeducation/TAU ($n=694$, $k=3$, $SMD=-0.48$ [95% CI [-0.79, -0.17]]). In both depression and anxiety, sensitivity analyses of swapping different outcome measures for other instruments for the same studies at three to six-month follow-up as was undertaken for the outcomes at post-treatment; this did not greatly change the results (Supplement 3). There was one exception: Internet-based BA was no longer superior to psychoeducation/TAU in the case of anxiety (Supplement 3).

Only two studies explored the effect of Internet-based BA on depression at 12-month follow-up (Buntrock et al., 2016; O'Mahen et al., 2014). There was a high level of evidence that the Internet-based BA group had significantly greater reductions in depression scores than psychoeducation/TAU ($n=460$, $k=2$, $SMD=-0.25$ [95% CI [-0.44, -0.07]]).

Five studies assessed the effect of Internet-based BA on quality of life (Buntrock et al., 2016; Carlbring et al., 2013; Ebert et al., 2017; Ly et al., 2014, 2015). There was a low level of evidence that at post-treatment, Internet-based BA achieved similar results to both alternative forms of behavioral therapy ($n=93$, $k=1$, $SMD= 0.01$ [95% CI -0.40, 0.42]) and mindfulness ($n=81$, $k=1$, $SMD= 0.04$ [-0.39, 0.48]). There was a moderate level of evidence that Internet-based BA leads to higher quality of life scores than psychoeducation/TAU ($n=662$, $k=2$, $SMD=0.64$ [95% CI 0.47, 0.80]). However, there was a low level of evidence from one study that compared Internet-based BA with wait-list controls which found no difference between the groups at post-treatment (Carlbring et al, 2013) ($n=80$, $k=1$, $SMD=0.02$ [-0.42, 0.46]). Results

when comparing Internet-based BA against psychoeducation/TAU at three to six-month follow-up were similar ($n=662$, $k=2$, $SMD=0.50$ [0.35, 0.66]); the quality of evidence was moderate.

3.5. Heterogeneity and publication bias:

I^2 values were less than 50% for all but two outcomes; specifically, anxiety at post-treatment for other behavioral therapy comparator group ($I^2=61\%$) and anxiety at three to six-month follow-up for psychoeducation/TAU comparator group ($I^2=67\%$). We could not test for publication bias as there were insufficient studies for any of the outcomes.

4. Discussion

The purpose of this review was to determine for which conditions the efficacy of Internet-based BA has been evaluated, to determine its efficacy, and assess the overall quality of evidence. The efficacy of Internet-based BA was only examined in non-clinical populations for the treatment of different types of depression (e.g., mild forms of depression, postpartum depression, depression with a comorbid chronic medical condition such as diabetes) as evaluated in RCTs. In this context, guided Internet-based BA was non-inferior to other forms of behavioral therapy and mindfulness but superior to physical activity, wait-list, psychoeducation/TAU for the most commonly evaluated outcomes of depression and anxiety at post-treatment and at short-term follow-up. The same was found for quality of life when available, except in the comparison of Internet-based BA against waitlist. The overall quality of evidence was mainly very low to low when comparing Internet-based BA to other forms of behavioral therapy, mindfulness and waitlist, very low for physical activity, and mainly moderate for psychoeducation/TAU. The quality of evidence was most commonly downgraded due to imprecision. Concerns over the quality of the evidence meant we excluded three uncontrolled studies despite our protocol.

Although BA has the potential to benefit individuals with other conditions (e.g., anxiety, post-traumatic stress disorder, problem gambling) when delivered through traditional face-to-face methods (Hopko et al., 2006, 2016; Luquiens et al., 2016; Mulick et al., 2011; Soleimani et al., 2015), we did not find any studies on the use of Internet-BA in other psychiatric conditions such as anxiety or post-traumatic stress disorder. There is therefore a need to evaluate the efficacy of Internet-BA in other conditions due to advantages that BA can offer over the firmly established CBT treatment for many mental conditions (Richards et al., 2016).

BA, especially when delivered via distance-based formats, including the Internet, has the potential to be a cost-effective treatment option. In fact, early evidence coming from the only RCT that has compared an Internet-based BA treatment against a face-to-face BA treatment (Ly et al., 2015), suggests that the use of Internet can possibly reduce the therapist time in a full face-to-face BA treatment while maintaining the same treatment quality. Additionally, due to BA's parsimoniousness, it has potential to be delivered by coaches who are graduate workers rather than clinical professionals. Many studies included in this review assessed a guided-Internet based BA program with students acting as supporters. Consistent with these observations, the employment of graduate workers to deliver low-intensity therapies, like BA, has been successful under the United Kingdom's Improving Access to Psychological Therapies (IAPT) program (NHS England, 2016). However, due to a paucity of data, a lack of certainty currently exists regarding the optimum support that should be provided when delivering an Internet-based BA intervention. It was not possible to assess whether the type of facilitator (i.e., student, paraprofessional, professional therapist) could impact depression and anxiety outcomes. It was also not possible to explore what to include when providing support, nor how much and how often support should be provided. Apart from this, all of the interventions in this review were guided

and delivered through a web-based platform, therefore little can also be concluded regarding the most effective form of Internet-based BA. For instance, we do not know whether Internet-based BA interventions can be more effective when guidance is added, nor do we know whether web-based or native BA application interventions can be more effective than BA interventions delivered via email. Although unguided CBT in the treatment of depression is not recommended (NICE guidance on adult depression, 2009) because it is inferior to guided CBT, future research is needed to assess the efficacy of unguided distance-based BA therapy against guided distance-based BA therapy, as this could be an even more cost-effective treatment option. Also, of note, none of the retrieved studies in this review included an economic analysis of Internet-based BA treatment compared to in-person BA treatment. The studies retrieved by Donker et al (2015) showed promising evidence of cost-effectiveness for guided Internet-based treatment for depression, but as per our knowledge there is no data specifically on Internet-based BA treatments.

As there were no limitations placed on control group, the comparison interventions varied widely. However, this can actually be positive as it is very helpful to compare Internet-based BA to an already established, evidence-supported treatment as opposed to a wait-list or no treatment group, especially taking into account best practices guidelines (e.g., NICE guidance on adult depression, 2009). This provides more convincing evidence on efficacy and can either confirm or deny whether this treatment should be offered as a stand-alone alternative.

In addition to variations in the comparator group, there were also differences in the condition being treated. All types of depression were initially analyzed together to see if the magnitude of the effect across these different forms of depression would be more or less the same. Sensitivity analysis confirmed that excluding specific subtypes did not change the effect of

Internet-based BA, with the exception of complicated grief and rumination ($n = 1$ study) (Eisma et al., 2015). One explanation might be that this condition is somewhat different from other depressive disorders even if there is an overlap in symptoms (e.g., deep sadness, intense avoidance and social withdrawal). For grief, even with low mood, there is very reactive avoidance which is best treated with exposure therapy, although BA can help if activity levels are low (Boelen et al., 2010; Shear, 2015). However, while complicated grief may have more overlap with depression than normal grief, it is important to note that not all grief presents with low activity levels (Horowitz et al., 1997). Activity levels at baseline were not collected, therefore it is difficult to draw any conclusions regarding this hypothesis. Moreover, the inconsistency observed for the effect of Internet-based BA on complicated grief could also be explained by other factors including the study design, the delivery method, the baseline characteristics of the participants, and/or the levels of attrition as opposed to the condition being treated. This study was the only one to deliver its intervention via email instead of a web-based platform. An intervention delivered via email may be less appealing to individuals since persuasive system design features cannot be used, and that could have led to poorer adherence and as a result, poorer treatment outcomes. Additionally, baseline levels of depression were found to be higher in the BA group than those in the comparator groups (i.e., the wait-list and exposure groups). Finally, the BA group had higher levels of attrition than the other comparator groups.

Several models of BA exist, with activity monitoring and activity scheduling being core components common across the variations. Activity monitoring forms the basis for planning subsequent behavior change efforts, while activity scheduling involves scheduling activities into the individual's everyday life to improve mood. Models of BA differ in how these common core

components are implemented as well as in additional strategies and techniques that may also be incorporated into the program (e.g., assessment of goals and values, social skills and problem-solving training, targeting avoidance behaviors). Martell's model (Martell et al., 2010) is most commonly used in practice, and has also been found to be the most commonly empirically evaluated model in the context of Internet-based BA. When assessing the quality of these tested interventions based on the extent to which these interventions incorporate the models core components there was a lack of adequate information about their content in many manuscripts. Following the CONSORT recommendations regarding the reporting of intervention descriptions (Schulz et al., 2010), it is important that authors present this information, especially in the context of Internet-delivered treatments. It is a challenge to translate the core components of BA into distance-delivery formats while still adhering to the treatment manuals due to the lack of therapist interaction. Therefore, it is necessary for authors to appropriately describe the intervention to ensure it is, in fact, still in line with BA.

This review also identified a few other additional gaps in the literature such as the efficacy of Internet-based BA in youth populations, especially given the evidence is starting to show that face-to-face BA has applicability in the treatment of adolescents (Gaynor and Harris, 2008; Pass et al., 2017; Ritschel et al., 2011; Ruggiero et al., 2007; Wallis et al., 2012), and adolescents are technically savvy and Internet is routinely incorporated into their everyday life. This is particularly relevant as psychological treatments are recommended as first line interventions for children and youth (Parikh et al., 2016). Another gap identified was the lack of studies in clinical populations. As such, it is necessary to perform research on the efficacy of Internet-based BA in clinical settings. Finally, it is important to note the short follow-up time observed in the included studies; only two studies had a 12-month follow-up (Buntrock et al., 2016; O'Mahen et al.,

2014). Therefore, it is difficult to assess the efficacy of Internet-based BA over a longer period of time, warranting the need for further research on its extended efficacy.

4.1. Limitations of this review:

First, our search was restricted to papers in English and therefore, relevant papers written in other languages could have been missed.

Second, the quality of the interventions was only based on the whether the interventions implemented core components of the BA models. However, a multidimensional approach considering other aspects regarding the content, such as the extent to which the core components are actually implemented as theoretically designed or usability, should ideally be used. Taking into account the quality of the interventions is very important since this can affect the results of the trials (Herbert and Bo, 2005).

Third, when evaluating the risk of bias within each included study we used the Cochrane risk of bias tool. Although this tool has been suggested as the best available tool for assessing RCTs (Zeng et al., 2015), this tool does not include all the potential risk of bias. For instance, the involvement of developers in evaluation studies is a potential risk of bias (Petrosino and Soydan, 2005), but was not assessed. Apart from this, the extent to which program evaluators were involved in developing and delivering the evaluated intervention was not clear in any of the included studies in this review. Consequently, we cannot rule out the possibility that the estimated effect sizes obtained and reported in this review may be inflated by a conflict-of-interest.

Fourth, due to the limited number of studies, when estimating the effect size for Internet-based BA, although not ideal, it was necessary to group different conditions as well as different comparators.

Fifth, some of our results showed moderate heterogeneity and we could not test for publication bias as there were insufficient studies for any of the outcomes.

Finally, a limitation of the included RCTs and, therefore, of this review is that participants excluded people with current suicidality, psychosis, and other comorbidities. Additionally, clinical populations were not investigated. This limits the generalizability of our findings.

4.2. Conclusion:

Due to the limited number of studies and the poor quality of some of the findings, our results must be viewed with caution. Nevertheless, the results of this review suggest that guided Internet-based BA seems to be an efficacious therapy option for the treatment of depression in adults recruited from the general population, resulting in reduced depression and anxiety symptoms. In this context, guided Internet-based BA was superior to less intensive forms of treatment including psychoeducation/TAU and waitlist. It also seems to be non-inferior to in-person BA and other forms of behavior therapy.

Our results emphasize the need for high quality studies with longer follow-ups which assess the effectiveness and cost-effectiveness of Internet-based BA as an option for treating depression (both in adults and adolescents) as well as other conditions, especially in the clinical setting. This is taking into account the lack of studies and the growing demand for the healthcare system to adopt scalable delivery technologies that still offer personalized and effective treatments.

The following authors have contributed to this submission entitled **A systematic review and meta-analysis on the efficacy of Internet-delivered behavioral activation**: Anna Huguet, Alyssa Miller, Steve Kisely, Sanjay Rao, Nelda Saadat and Patrick McGrath. Specifically, the individual roles of the authors are as follows:

Huguet A wrote the protocol, screened items retrieved from the literature search, evaluated the overall quality of evidence, interpreted the data and helped to write the first draft of the manuscript.

Miller A screened items retrieved from the literature search, extracted data, evaluated the risk of bias, helped to evaluate the overall quality of evidence and helped to write the first draft of the manuscript.

Kisely S contributed to write the protocol, undertook the statistical analysis, and helped to interpret data.

Rao SN contributed to write the protocol and helped to interpret data.

Saadat N extracted data, and evaluated the risk of bias.

All authors contributed to and have approved the final manuscript.

Acknowledgements:

The authors do not wish to acknowledge anyone at this time.

The authors have no conflicts of interest to declare.

References:

Andersson, G., Titov, N., 2014. Advantages and limitations of Internet-based interventions for common mental disorders. *World Psychiatry*. 13, 4-11.

Boelen, P.A., de Keijser, J., van den Hout, M.A., van den Bout, J., 2010. Factors associated with outcome of cognitive-behavioural therapy for complicated grief: A preliminary study. *Clin Psychol and Psychot*. 18, 284-291.

- Buntrock, C., Ebert, D.D., Lehr, D., Smit, F., Riper, H., Berking, M., Cuijpers, P., 2016. Effect of a web-based guided self-help intervention for prevention of major depression in adults with subthreshold depression: A randomized clinical trial. *JAMA*. 315, 1854-1863.
- Butler, A.C., Chapman, J.E., Forman, E.M., Beck, A.T., 2006. The empirical status of cognitive-behavioral therapy: A review of meta-analyses. *Clinical Psychol Rev*. 26, 17-31.
- Carlbring, P., Hagglund, M., Luthstrom, A., Dahlin, M., Kadowaki, A., Vernmark, K., Andersson, G., 2013. Internet-based behavioral activation and acceptance-based treatment for depression: A randomized controlled trial. *J Affect Disorders*. 148, 331-337.
- Dimidjian, S., Hollon, S.D., Dobson, K.S., Schmalings, K.B., Kohlenberg, R.J., Addis, M.E., Gallop, R., McGlinchey, J.B., Markley, D.K., Gollan, J.K., Atkins, D.C., Dunner, D.L., Jacobson, N.S., 2006. Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *J Consult Clin Psych*. 74, 658-670.
- Donker, T., Blankers, M., Hedman, E., Ljótsson, B., Petrie, K., Christensen, H., 2015. Economic evaluations of Internet interventions for mental health: a systematic review. *Psychol Med*. 45, 3357-76.
- Dragioti, E., Karathanos, V., Gerdle, B., Evangelou, E., 2017. Does psychotherapy work? An umbrella review of meta-analyses of randomized controlled trials. *Acta Psychiat Scand*. 136, 236-246.
- Ebert, D.D., Nobis, S., Lehr, D., Baumeister, H., Riper, H., Auerbach, R.P., Snoek, F., Cuijpers, P., Berking, M., 2017. The 6-month effectiveness of Internet-based guided self-help for depression in adults with Type 1 and 2 diabetes mellitus. *Diabetic Med*. 34, 99-107.

- Eisma, M.C., Boelen, P.A., van den Bout, J., Stroebe, W., Schut, H.A.W., Lancee, J., Stroebe, M.S., 2015. Internet-Based Exposure and Behavioral Activation for Complicated Grief and Rumination: A Randomized Controlled Trial. *Behav Ther.* 46, 729-748.
- Ekers, D., Richards, D., McMillan, D., Bland, J.M., Gilbody, S., 2011. Behavioural activation delivered by the non-specialist: phase II randomized controlled trial. *Brit J of Psychiat.* 198, 66-72.
- Ekers, D., Webster, L., Van Straten, A., Cuijpers, P., Richards, D., Gilbody, S., 2014. Behavioural Activation for Depression; An Update of Meta-Analysis of Effectiveness and Sub Group Analysis David. *PLoS ONE.* 9, e100100.
- Gaynor, S. T., & Harris, A, 2008. Single-participant assessment of treatment mediators: Strategy description and examples from a behavioral activation intervention for depressed adolescents. *Behav Modif.* 32, 372–402.
- Goldner, E.M., Jones, W., Fang, M.L., 2011. Access to and waiting time for psychiatrist services in a Canadian urban area: A study in real time. *Can J of Psychiat.* 56, 474-480.
- Guyatt, G., Oxman, A.D., Akl, E.A., Kunz, R., Vist, G., Brozek, J., Norris, S., Falck-Ytter, Y., Glasziou, P., DeBeer, H., Jaeschke, R., Rind, D., Meerpohl, J., Dahm, P., Schunemann, H.J., 2011. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol.* 64, 383–394.
- Hazell, C.M., Strauss, C., Cavanagh, K., Hayward, M., 2017. Barriers to disseminating brief CBT for voices from a lived experience and clinician perspective. *PLoS ONE.* 12, e0178715.
- Herbert, R.D., Bo, K., 2005. Analysis of quality of interventions in systematic reviews. *BMJ.* 331,507.

- Higgins, J.P., Green, S., 2008. *Cochrane handbook for systematic reviews of interventions*: Wiley Online Library.
- Higgins, J.P., Altman, D.G., Gotzsche, P.C., Juni, P., Moher, D., Oxman, A.D., Savovic, J., Schulz, K.F., Weeks, L., Sterne, J.A.C., 2011. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 343, d5928.
- Hopko, D.R., Robertson, S.M.C., Lejuez, C.W., 2006. Behavioral activation for anxiety disorders. *Behav Anal Today*. 7, 212-231.
- Hopko, D.R., Lejuez, C.W., Ryba, M.M., Shorter, R.L., Bell, J.L., 2016. Support for the efficacy of behavioural activation in treating anxiety in breast cancer patients. *Clinical Psychol-UK*. 20, 17-26.
- Horowitz, M.J., Siegel, B., Holen, A., Bonanno, G.A., Milbrath, C., Stinson, C.H., 1997. Diagnostic criteria for complicated grief disorder. *Am J Psychiat*. 154, 904-910.
- Jacobson, N.S., Dobson, K.S., Truax, P.A., Addis, M.E., Koerner, K., Gollan, J.K., Gortner, E., Prince, S.E., 1996. A component analysis of cognitive-behavioral treatment for depression. *J Consult Clin Psych*. 64, 295-304.
- Kanter, J.W., Busch, A.M., Rusch, L.C., 2009. *Behavioral Activation: Distinctive Features*. Routledge, East Sussex.
- Kanter, J.W., Manos, R.C., Bower, W.M., Baruch, D.E., Busch, A.M., Rusch, L.C. 2010. What is behavioral activation? A review of the empirical literature. *Clin Psychol Rev*. 30, 608–620.
- Luquiens, A., Tanguy, M.L., Lagadec, M., Benyamina, A., Aubin, H.J., Reynaud, M., 2016. The efficacy of three modalities of Internet-based psychotherapy for non-treatment-seeking online problem gamblers: a randomized controlled trial. *J Med Internet Res*. 18, e36.

Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. *Biometrics*. 33, 159-174.

Lejuez, C.W., Hopko, D.R., Hopko, S.D., 2001. A brief behavioral activation treatment for depression. *Behav Modif*. 25, 255-286.

Lewinsohn, P.M., Biglan, A., Zeiss, A.M., 1976. Behavioral treatment for depression, in: Davidson, P.O. (Ed.), *Behavioral management of anxiety, depression and pain*. Brunner/Mazel, New York, pp 91-146.

Ly, K.H., Truschel, A., Jarl, L., Magnusson, S., Windahl, T., Johansson, R., Carlbring, P., Andersson, G., 2014. Behavioural activation versus mindfulness-based guided self-help treatment administered through a smartphone application: a randomised controlled trial. *BMJ Open*. 4, e003440.

Ly, K.H., Topooco, N., Cederlund, H., Wallin, A., Bergstrom, J., Molander, O., Carlbring, P., Andersson, G., 2015. Smartphone-supported versus full behavioural activation for depression: a randomised controlled trial. *PLoS ONE*. 10, e0126559.

Martell, C.R., Dimidjian, S., & Herman-Dunn, R., 2010. *Behavioral Activation for Depression: A Clinician's Guide*. Guilford Press, New York.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol*. 62, 1006-1012.

Mulick, P.S., Landes, S.J., Kanter, J.W., 2011. Contextual Behavior Therapies in the Treatment of PTSD: A Review. *Int J of Behav Consult Ther*. 7, 23–31.

National Institute for Health and Clinical Excellence, 2009. *Depression in adults: Recognition and management*. NICE guidance (CG90).

NHS England, Health Education England, 2016. 2015 Adult IAPT Workforce Census Report.

NHS England.

Nystrom, M.B.T., Stenling, A., Sjostrom, E., Neely, G., Lindner, P., Hassmen, P., Andersson, G., Martell, C., Carlbring, P., 2017. Behavioral activation versus physical activity via the internet: A randomized controlled trial. *J Affect Disorders*. 2015, 85-93.

O'Mahen, H.A., Woodford, J., McGinley, J., Warren, F.C., Richards, D.A., Lynch, T.R., Taylor, R.S., 2013. Internet-based behavioral activation—Treatment for postnatal depression (Netmums): A randomized controlled trial. *J Affect Disorders*. 150, 814-822.

O'Mahen, H.A., Richards, D.A., Woodford, J., Wilkinson, E., McGinley, J., Taylor, R.S., Warren, F.C., 2014. Netmums: a phase II randomized controlled trial of a guided Internet behavioural activation treatment for postpartum depression. *Psychol Med*. 44, 1675-1689.

Parikh, S.V., Quilty, L.C., Ravitz, P., Rosenbluth, M., Pavlova, B., Grigoriadis, S., Velyvis, V., Kennedy, S.H., Lam, R.W., MacQueen, G.M., Milev, R.V., Ravindran, A.V., Uher, R., 2016. CANMAT Depression Work Group. Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 Clinical Guidelines for the Management of Adults with Major Depressive Disorder: Section 2. Psychological Treatments. *Can J Psychiat*. 61, 524-39.

Pass, L., Lejuez, C., & Reynolds, S., 2017. Brief Behavioural Activation (Brief BA) for Adolescent Depression: A Pilot Study. *Behav Cogn Psychother*. 1-13. doi: 10.1017/S1352465817000443. [Epub ahead of print].

Petrosino, A., Soydan, H, 2005. The impact of program developers as evaluators on criminal recidivism: results from meta-analyses of experimental and quasi-experimental research. *J Exp Criminol*. 1, 435-50.

- Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., Hensley, M., 2011. Outcomes for implementation Research: Conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Hlth.* 38, 65-76.
- Richards, D.A., Ekers, D., McMillan, D., Taylor, R.S., Byford, S., Warren, F.C., Barrett, B., Farrand, P.A., Gilbody, S., Kuyken, W., O'Mahen, H., Watkins, E.R., Wright, K.A., Hollon, S.D., Reed, N., Rhodes, S., Fletcher, E. Finning, K., 2016. Cost and Outcome of Behavioural Activation versus Cognitive Behavioural Therapy for Depression (COBRA): a randomised, controlled, non-inferiority trial. *The Lancet.* 388, 871-880.
- Ritschel, L.A., Ramirez, C.L., Jones, M., Craighead, W.E., 2011. Behavioral activation for depressed teens: a pilot study. *Aust J Rural Health.* 18, 281-299.
- Ruggiero, K. J., Morris, T. L., Hopko, D. R., Lejuez, C. W., 2007. Application of behavioral activation treatment for depression to an adolescent with a history of child maltreatment. *Clin Case Stud.* 6, 64-78.
- Schulz, K.F., Altman, D.G., Moher, D., 2010. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomized trials. *BMJ.* 340, c332.
- Shear, M.K., 2015. Complicated grief. *N Engl J Med.* 372, 153-160.
- Shinohara, K., Honyashiki, M., Imai, H., Hunot, V., Caldwell, D.M., Davies, P., Moore, T.H.M., Furukawa, T.A., Churchill, R., 2013. Behavioural therapies versus other psychological therapies for depression. *Cochrane Db Systc Rev.* 10, CD008696.
- Soleimani, M., Mohammadkhani, P., Dolatshahi, B., Alizadeh, H., Overmann, K.A., Coolidge, F.L., 2015. A Comparative Study of Group Behavioral Activation and Cognitive Therapy in Reducing Subsyndromal Anxiety and Depressive Symptoms. *Iran J Psychiat.* 10, 71-78.

Soucy Chartier, I., Provencher, M.D., 2012. Behavioural activation for depression: Efficacy, effectiveness and dissemination. *J Affect Disorders*. 145, 292-299.

Takeshima, N., Sozu, T., Tajika, A., Ogawa, Y., Hayasaka, Y., Furukawa, T.A., 2014. Which is more generalizable, powerful and interpretable in meta-analyses, mean difference or standardized mean difference? *BMC Med Res Methodol*. 14, 30.

Viera, A.J., Garrett, J.M., 2005. Understanding interobserver agreement: the kappa statistic. *Fam Med*. 37, 360-3.

Wallis, A., Roeger, L., Milan, S., Walmsley, C., Allison, S., 2012. Behavioral activation for the treatment of rural adolescents with depression. *Aust J Rural Health*. 20, 95-6.

Zeng, X., Zhang, Y., Kwong, J.S., Zhang, C., Li, S., Sun, F., Niu, Y., Du, L., 2015. The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: a systematic review. *J Evid Based Med*. 8, 2-10.

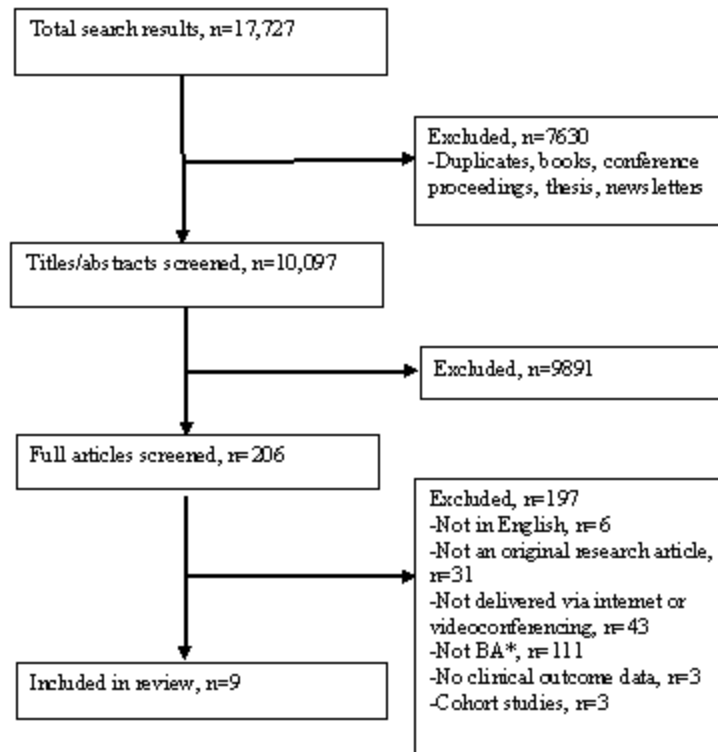


Fig 1. Flowchart of screening process *BA= Behavioral Activation

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Group similarity at baseline (other bias)	Co-interventions (other bias)
Buntrock 2016	+	+	-	-	+	+	+	+
Carlbring 2013	+	+	-	-	+	?	+	+
Ebert 2016	+	+	-	-	+	-	+	?
Eisma 2015	+	?	-	-	+	?	+	+
Ly 2014	+	+	-	-	+	?	+	+
Ly 2015	+	+	-	-	+	+	+	?
Nyström 2017	+	+	-	-	+	-	-	+
O'Mahen 2013	+	+	-	-	+	?	+	+
O'Mahen 2014	+	+	-	-	+	?	+	+

Fig 2. Cochrane risk of bias assessment for RCTs. Green represents low risk of bias, yellow represents unclear risk of bias and red represents high risk of bias.

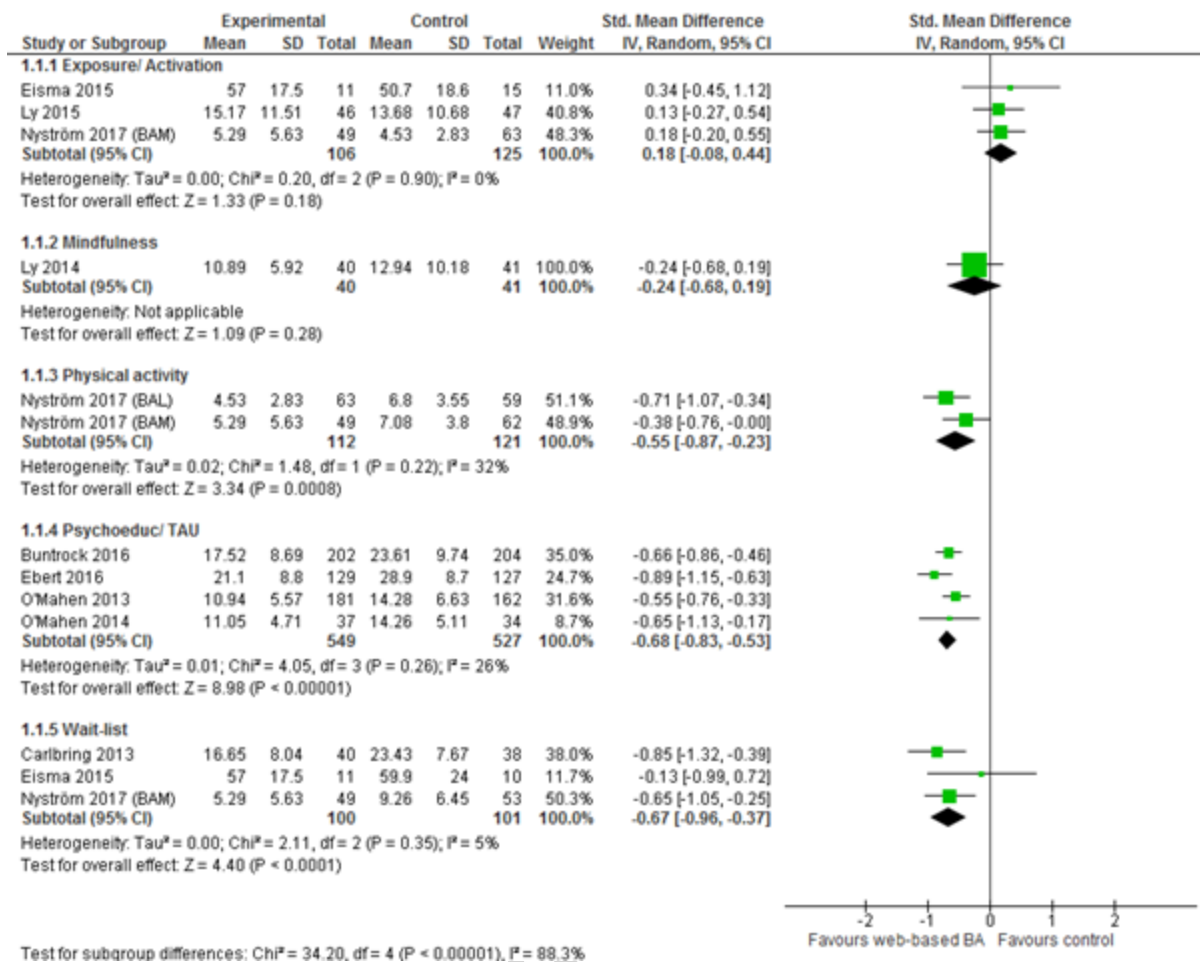


Fig 3. Forest plot displaying effects sizes of studies comparing Internet-based treatment with a comparison group on depression outcomes at post-treatment

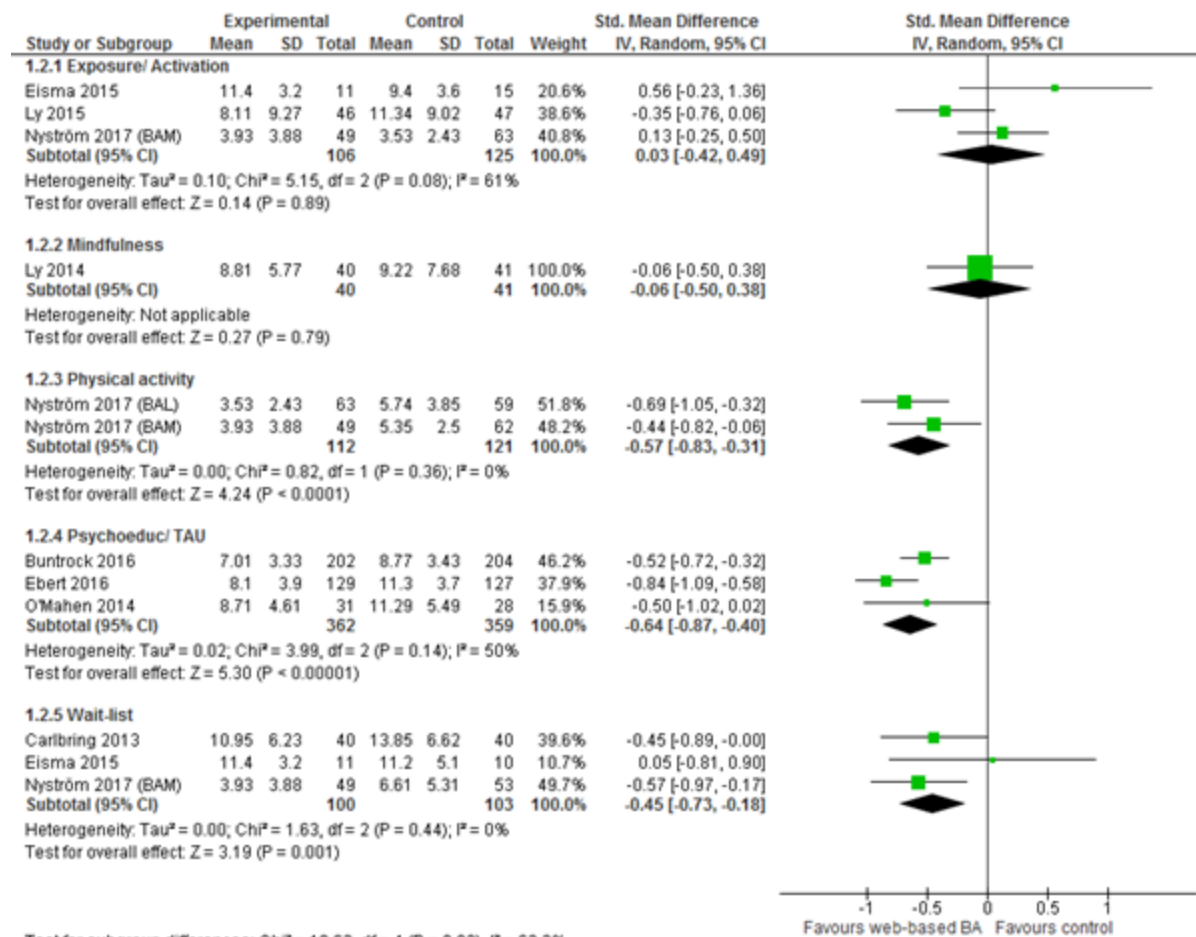


Fig 4. Forest plot displaying effects sizes of studies comparing Internet-based treatment with a comparison group on anxiety outcomes at post-treatment

Table 1. Description of studies included in the review

First author, Year	Number of arms	Country	Condition	Comparator (s)	n ¹	FU ²	Outcomes	Measures
Buntrock, 2016	2	Germany	Sub-threshold depression	Psychoeducation	40	6, 12	Depression Overall mental health Overall physical health Anxiety Behavioral activation Social problem solving	CES-D SF-12 MCS SF-12 PCS HADS-A BADSF SPSI-NPO, SPSI-

							Worrying Insomnia Mastery	PPO PSWQ ISI PSMS
Carlbring, 2013	2	Sweden	Depression	Wait list	80	6	Depression Quality of life Anxiety	BDI-II QOLI BAI
Ebert, 2017	2	Germany	Depression and diabetes	Psychoeducation	26	6	Depression Depression/anxiety Diabetes management Glycated haemoglobin level Diabetes self-management Acceptance of diabetes Overall physical health Overall mental health	CES-D HADS PAID Blood test DSMQ AADQ SF-12: PCS SF-12: MCS
Eisma, 2015	3	Netherlands	Complicated grief and rumination	Wait list, exposure	47	3	Complicated grief Posttraumatic stress Depression Anxiety Grief rumination Depressive brooding Depressive reflection	ICG-R PSS HADS HADS UGRS RRS RRS
Ly, 2014	2	Sweden	Depression	Mindfulness	81	6	Depression Anxiety Psychological flexibility Quality of life	BDI-II, PHQ-9 BAI AAQ-II QOLI
Ly, 2015	2	Sweden	Depression	In-person BA	93	6	Depression	BDI-II, PHQ-9

							Anxiety Psychological flexibility Quality of life	BAI AAQ-II QOLI
Nystrom, 2017	5	Sweden	Depression	Physical activity (with and without rationale), wait list	31 2	Non e	Depression Anxiety	PHQ-9 GAD-7
O'Mahen, 2013	2	UK	Postpartum depression	TAU	91 0	Non e	Depression	EPDS
O'Mahen, 2014	2	UK	Postpartum depression	TAU	83 6, 12		Depression Anxiety Work and social adjustment Postnatal bonding Social provision	EPDS GAD-7 WASAS PBQ SPS

BDI-II= Beck Depression Inventory, GAD-7= Generalized Anxiety Disorder 7-item scale, PHQ-9= Patient Health Questionnaire, SPS= Social Provision Scale, PBQ= Postnatal Bonding Questionnaire, WASAS= Work and Social Adjustment Scale, EPDS=Edinburgh Postnatal Depression Scale, AAQ-II= Acceptance and Action Questionnaire, BAI= Beck Inventory of Anxiety, RRS=Ruminative Response Scale, UGRS=Utrecht Grief Rumination Scale, HADS= Hospital Anxiety and Depression Scale, PSS=PTSD Symptom Scale, ICG-R=Inventory of Complicated Grief-Revised, SR-12 MCS= Short-Form Health Survey, Mental Health Summary Scale, SF-12 PCS=Short Form Health Survey - Physical Health Summary Scale, AADQ= acceptance and action diabetes questionnaire, DSMQ= Diabetes Self-Management Questionnaire, PAID= Problem Areas in Depression Scale, CESD= Centre for Epidemiological Studies Depression Scale, QOLI= Quality of Life Inventory, PSMS= Pearling and Schooler Mastery Scale, ISI= Insomnia Severity Inventory, PSWQ= Penn State Worrying Questionnaire, SPSI-NPO=Problem Solving Inventory-Revised, Negative Problem Orientation Subscale, SPSI-PPO=Social Problem Solving Inventory-Revised, Positive Problem Orientation subscale, BADS-SF= Behavioural Activation for Depression Scale Short Form, HADS-A= Hospital Anxiety and Depression Scale, Anxiety subscale

¹ n=Number of participants enrolled in study, all arms

² FU= Length of follow-up in months

Table 2: Description of interventions included in the review

First author, year	Delivery mode	Duration of treatment (weeks)	Intervention components	BA Model	Facilitator	Minutes per week/participant	Communication form
Buntrock, 2016	Web-based application	6	BA, PST ¹	Lewinsohn	Professional	20-40	Online
Carlbri ng, 2013	Web-based application	10	BA, ACT ²	Kanter	Student	15	Web-based messaging
Ebert, 2017	Web-based application	8	BA, PST	Unclear	Professional	30-60	Web-based messaging
Eisma, 2015	Email	6-8	BA	Lejuez, Hopko and Hopko	Professional	60	Unclear
Ly, 2014	Native application	8	BA	Combined	Student	20	SMS, email
Ly, 2015	Native application	9	BA	Martell	Student	Unclear	In-person, SMS
Nystrom, 2017	Web-based application	12	BA	Martell, Lewinsohn	Student	15	Email
O'Mahen, 2013	Web-based application	15	BA	Martell	Professional	Unclear	Web-based messaging
O'Mahen, 2014	Web-based application	17	BA	Martell	Professional	20-30	Phone

¹PST= Problem solving therapy²ACT= Acceptance and commitment therapy

Table 3. Fidelity of the interventions to core components of the BA models on which they are based

Martell's model						
Author, year	Psychoeducation	Setting goals	Functional analysis	Review/relapse prevention	Activation strategy used	
Ly, 2014 ¹	✓	?	?	?	Focused activation.	
Ly, 2015	✓	?	?	✓	Focused activation.	
O'Mahen, 2013	✓	✓	✓	✓	Focused activation, avoidance modification, rumination prevention, overcoming barriers.	
O'Mahen, 2014	✓	✓	✓	✓	Avoidance modification, overcoming barriers.	
Nystrom, 2017	?	?	?	?	Unclear.	
Lejuez, Hopko and Hopko's model						
Author, year	Psychoeducation/rationale	Establishing baseline	Values assessment	Activity goals/scheduling	Rewards/reinforcement	
Eisma, 2016	✓	✓	✓	✓	?	
Ly, 2014 ¹	✓	?	?	✓	✓	
Lewinsohn's model						
Author, year	Establishing baseline	Setting goals	Pleasant activity scheduling			
Buntrock, 2016	?	?	✓			
Nystrom, 2017	?	?	?			
Kanter's model						
Author, year	Establishing baseline/rationale	Values assessment/activity monitoring		Simple activation	Functional assessment	Relapse prevention
Carlbring, 2013	✓	?		✓	✓	✓

Note: ✓=Included. ?= Not evident to be included, ✖=Not included. Ebert, 2017 was not assessed as it was unclear which BA model they used.

¹Ly, 2014 based on Martell's model and Lejuez, Hopko and Hopko's model.

Table 4. GRADE evidence profile

		Quality assessment					Summary of findings				
		# of studies (Design)	Limitations ¹	Inconsistency	Indirectness	Imprecision ⁶	Publication bias ⁷	# of individuals		Quality	
							Internet-based BA	Comparator	SMD (95% CI)		
Internet-based BA against other forms of behavioral therapy	Depression at post-treatment	3 (RCT)	Serious limitations ²	No serious	No serious	Very serious	Undetected	106	125	0.18 (-0.08, 0.44)	Very low
	Anxiety at post-treatment	3 (RCT)	Serious limitations ²	No serious ⁴	No serious	Very serious	Undetected	106	125	0.03 (-0.42, 0.49)	Very low
	Depression at 3-6 month follow-up	2 (RCT)	No limitations	No serious	No serious	Very serious	Undetected	57	57	0.09 (-0.19, 0.36)	Low
	Anxiety at 3-6 month follow-up	2 (RCT)	No limitations	No serious	No serious	Serious	Undetected	57	59	-0.48 (-0.79, -0.17)	Moderate
	Quality of life at post-treatment	1 (RCT)	No limitations	No serious	No serious	Very serious	Undetected	46	47	0.01 (-0.40, 0.42)	Low
Internet-based BA	Depression at post-treatment	1 (RCT)	No limitations	No serious	No serious	Very serious	Undetected	40	41	-0.24 (-0.68, 0.19)	Low
	Anxiety at post-treatment	1 (RCT)	No limitations	No serious	No serious	Very serious	Undetected	40	41	-0.06 (-0.50, 0.38)	Low

Quality of life at post-treatment										
	1	No	No	No	Very	Undete	40	41	0.04	Lo
	(RC	serious	serious	serious	serious	cted			(-0.39,	w
	T)	limitatio							0.48)	
		ns								
Depression at 3-6 month follow-up										
	1	No	No	No	Very	Undete	40	41	-0.27	Lo
	(RC	serious	serious	serious	serious	cted			(-0.71,	w
	T)	limitatio							0.17)	
		ns								
Internet-based BA against physical	Depression at post-treatment									
	1	Very	No	No	Serious	Undete	112	121	-	Very
	(RCT)	serious	serious	serious		cted			0.55	low
		s							(-	
		limita							0.87	
		tions ²							,-	
								0.23)	
	Anxiety at post-treatment									
	1	Very	No	No	Serious	Undete	112	121	-0.57	Ve
	(RCT)	serious	serious	serious		cted			(-	ry
		s							0.83,-	lo
		limita							0.31)	w
		tions ²								
	Depression at post-treatment									
	4	Serio	No	No	No	Undete	549	527	-0.68	Mo
	(RCT)	us	serious	serious	serious	cted			(-	der
		limita							0.83,-	ate
		tions ³							0.53)	
	Anxiety at post-treatment									
	3	Serio	No	No	No	Undete	362	359	-0.64	Mo
	(RCT)	us	serious	serious	serious	cted			(-	der
		limita							0.87,-	ate
		tions ³							0.40)	
	Depression at 3-6 month follow-up									
	3	Serio	No	No	No	Undete	361	359	-0.68	Mo
	(RCT)	us	serious	serious	serious	cted			(-	der
		limita							0.83,-	ate
		tions ³							0.53)	
	Anxiety at 3-6 month follow-up									
	3	Serio	No	No	No	Undete	347	347	-0.48	Mo
	(RCT)	us	serious ⁵	serious	serious	cted			(-	der
		limita							0.79,-	ate
		tions ³							0.17)	

Depression at 12-month follow-up											
2	No	No	No	No	Undete	223	237	-0.25	Hi		
(RCT)	seriou	serious	serious	serious	cted			(-	gh		
	s							0.44,-			
	limita							0.07)			
	tions										
Quality of life at post-treatment											
2	Serio	No	No	No	Undete	331	331	0.64	Mo		
(RCT)	us	serious	serious	serious	cted			(0.47,	der		
	limita							0.80)	ate		
	tions ³										
Quality of life at 3-6 month follow-up											
2	Serio	No	No	No	Undete	330	331	0.50	Mo		
(RCT)	us	serious	serious	serious	cted			(0.35,	der		
	limita							0.66)	ate		
	tions ³										
Internet-based BA against waitlist	Depression at post-treatment										
	3	Serious	No	No	Serious	Undetected	100	101	-0.67	Low	
	(RCT)	limitations ²	serious	serious					(-		
									0.96,-		
									0.37)		
	Anxiety at post-treatment										
	3	Serious	No	No	Serious	Undetected	100	103	-0.45	Low	
	(RCT)	limitations ²	serious	serious					(-		
								0.73,-			
								0.18)			
Depression at 3-6 month follow-up											
1	No serious	No	No	Very	Undetected	11	10	-0.58	Low		
(RCT)	limitations	serious	serious	serious				(-			
								1.22,			
								0.07)			
Quality of life at post-treatment											
1	No serious	No	No	Very	Undetected	40	40	0.02	Low		
(RCT)	limitations	serious	serious	serious				(-			
								0.42,			
								0.46)			

¹Lack of blinding of participants and personnel and lack of blinding of outcome assessment has not been considered a potential risk of bias within studies when evaluating the limitations domains because blinding was not possible in any of the trials included in this trial. Blinding of participants and personnel is impossible when participants assigned to the comparator group are assigned to non-Internet comparators groups (e.g., wait list, mindfulness etc.). In these cases, study participants discover what condition they have been assigned to. Blinding of outcome assessment is not possible either, because explored outcomes, which are subjective in nature, were self-reported in all trials.

²Selective reporting and no group similarity at baseline.

³ Selective reporting.

⁴ Identified heterogeneity in results across studies ($I^2 > 50\%$) could be explained by differences across the comparator intervention.

⁵ Identified heterogeneity in results across studies ($I^2 > 50\%$) could be explained by a small trial (O'Mahen et al., 2014) combined in the meta-analysis.

⁶ If there is a wide confidence interval around the estimate of the effect (i.e., CI overlaps no or little effect **and** the upper or lower confidence limit crosses the effect size of SMD of 0.5 in either direction), or less than 400 participants were included.

⁷ Due to the small number of trials published, it is very difficult to assess likelihood of publication bias. The authors feel that the small number of trials at this stage is due to the still developing (early) evidence base available in the literature. For this reason, they have decided that the publication bias is still undetected at this stage.

Highlights:

- The efficacy of BA has been evaluated in adults with different forms of depression.
- Results supporting the efficacy of guided Internet BA for depression are promising.
- Findings are applicable only to the general population, not clinical population.
- High quality studies with long follow-up are needed to improve quality of evidence.
- Research is needed to evaluate the efficacy of BA for other conditions/populations.