Results from a trial of an unsupported internet intervention for depressive symptoms

Yan Leykin[1], **Ricardo F. Muñoz**[1,2], **Omar Contreras**[1], and **Melissa D. Latham**[1]

[1]University of California, San Francisco

[2]Palo Alto University

Abstract

Internet interventions provide an option for those who either cannot or choose not to engage with traditional treatments. Most research on internet interventions involves guided or supported interventions. However, unsupported interventions offer considerably more scalability and cost-effectiveness, which makes them attractive for large-scale implementation. In this study, 309 participants recruited via Google AdWords entered an unsupported cognitive-behavioral internet intervention for depressive symptoms. To maximize the ecological validity of the study, participants received no incentives or live contact with study personnel. Furthermore, the study was open to individuals at any level of depressive symptoms, and all participants received the active intervention. The main outcome measures were depressive symptom level and self-efficacy in managing depressive symptoms. At follow-up, depression scores were significantly lower than baseline scores at each follow-up point (1, 2, 4, and 7 months), with pre-post effect sizes ranging from medium to large. Follow-up depression self-efficacy scores were significantly higher than baseline scores at each follow-up point, with pre-post effect sizes in the medium range. The results remained significant when analyzing only participants with depression scores indicative of a presence of a major depressive episode; results likewise remained significant when employing the conservative last observation carried forward convention, even in the presence of high attrition observed in this study. The results illustrate the potential of unsupported internet intervention to address the health needs of the global community.

Keywords

self-help; major depression; unguided intervention; unsupported interventions; internet intervention; worldwide; global health

*Corresponding author: Yan Leykin, PhD, University of California, San Francisco, Department of Psychiatry, 3333 California St., Suite 465, San Francisco, CA 94143-0848, Phone: (415) 476-8799, Fax: (415) 476-7744, Yan.Leykin@ucsf.edu..

1. Introduction

Most individuals experiencing symptoms of depression lack access to effective interventions (World Health Organization, 2010). Many of them access the internet both to understand their symptoms and to look for resources that could address them. We have previously reported that, in a large worldwide sample of users visiting a depression screening website (Leykin et al., 2012b), 67% of participants screened positive for a current major depressive episode, yet only 25% of those screening positive for depression reported currently receiving depression treatment. These data suggest that innovative depression treatments must be developed and disseminated to individuals for whom traditional therapies fail to reach. Given that the internet already attracts individuals seeking alternative resources, developing and distributing such resources on the internet is a good strategy.

The key advantages of internet interventions, including unparalleled breadth of reach, scalability, and cost-effectiveness, can primarily be realized with interventions that are unsupported, that is, fully automated self-help interventions that do not rely on a provider, a coach, or on any other human contact for provision of services or for intervention effectiveness. Though human contact may improve engagement and outcomes of internet interventions (Fridrici et al., 2009; Leykin et al., 2012a; Muñoz et al., 2009), it also introduces considerable costs, which can substantially reduce or even negate the aforementioned benefits of internet interventions. In contrast, fully-automated unsupported interventions can be scaled reliably without increasing costs and without the need to engage the complex network of local health systems. Indeed, the largest trials of internet interventions for depression were unsupported, both in terms of the interventions themselves as well as in terms of trial administration (Christensen et al., 2002; Christensen et al., 2004b). The scalability and cost-effectiveness of unsupported interventions allow them to be both evaluated in and distributed to populations that do not usually participate in randomized trials, such as individuals with sub-syndromal symptoms (Powell et al., 2013). Thus, although interventions that are supported/guided by a clinician or a coach may yield somewhat greater improvement as compared to unguided interventions (Andersson and Cuijpers, 2009; Johansson and Andersson, 2012; Newman et al., 2011), these additional benefits are limited in scope given logistical challenges and costs of scaling such guidance to a larger population (Johansson and Andersson, 2012). To tackle the enormous challenge of reducing the global disease burden of depression (World Health Organization, 2010), systematic efforts should be undertaken to study and enhance the unsupported delivery model.

The significant benefit of internet interventions is the ability to disseminate them exactly as they were evaluated. This is in sharp contrast to traditional face-to-face treatment, which must be disseminated by training individual providers, each of whom can introduce "drift", that is, administering their own versions of the intervention, rather than the one that was manualized and tested (Shafran et al., 2009; Waller, 2009). An internet intervention, once tested and found successful, can be offered to the public without any changes and alterations, greatly increasing the likelihood that its effectiveness in the community will be very similar to the one observed during the trial, even if it is offered worldwide to thousands of users (which is possible with unsupported interventions). However, this consistency also

strongly suggests the need to evaluate an intervention in the same way as it is intended to be disseminated. Thus, common methods used by clinical trials to improve participant retention and engagement may compromise the generalizability of the original trials to their intended dissemination inasmuch as they introduce differences from the manner in which interventions will eventually be offered. For instance, financial incentives and phone-based follow-ups can improve retention in internet intervention trials (Fridrici et al., 2009; Leykin et al., 2012a; Muñoz et al., 2009). However, paying participants to visit the site or contacting them by phone to provide data exposes them to motivators that will not be present when the intervention is widely deployed. Thus, with financial incentives the ecological validity is reduced and engagement (and effectiveness) of an unincentivized intervention remains unknown. Similarly, phone follow-ups introduce variables that will not be present beyond the trial.

Though avoiding trial components that depart from ecological validity likely increases the generalizability of its findings, doing so also introduces problems with a traditionally important component of trials – the control group. Without financial incentives, participants allocated to a control condition in an internet-based trial would have few reasons to return to the site to provide follow-up data; those who would remain in the trial will likely be unrepresentative of all randomized to this arm. The promise of future participation (waitlist) is also unlikely to improve the follow-up rate, given the expectation of immediacy on the internet and likely availability of other internet resources. A possible solution is to conduct a trial as a single-condition study, without employing the control condition. Though a randomized controlled design may be preferable for understanding efficacy, such designs may also systematically exclude individuals who may be reluctant to participate in a study where they do not have control over treatment assignment or may risk being assigned to a non-treatment group; removing randomization and the associated control group may actually increase the ecological validity of the intervention and the representativeness of trial participants.

A number of studies and meta-analyses of these studies have confirmed the efficacy and the usefulness of internet interventions (Andersson and Cuijpers, 2009; Andrews et al., 2010; Griffiths et al., 2010; Van't Hof et al., 2009), yet few ecologically valid studies, conducted in a manner closely resembling eventual dissemination, exist. Indeed, studies that are described as self-help or unguided have used financial incentives (Clarke et al., 2005; Clarke et al., 2002) or phone interviews (Berger et al., 2011; Christensen et al., 2004a; Spek et al., 2007; Vernmark et al., 2010), or were administered in a structured setting such as a school (O'Kearney et al., 2006; O'Kearney et al., 2009). The few truly unsupported studies (Christensen et al., 2002; Christensen et al., 2004b; Donker et al., 2013; Meyer et al., 2009) used randomization, which may have turned away participants who were reluctant to receive a control condition or their non-preferred condition. Thus, the goal of this study was to understand the efficacy of an ecologically valid fully-automated, unsupported intervention for the reduction of depressive symptoms. As subthreshold depression carries considerable burden (Judd et al., 1994), individuals at any level of depressive symptoms were allowed to take part in the trial. To understand the effectiveness of the intervention as the user would experience it, and not under more idealized conditions that could produce unreplicable results, our "pragmatic" trial used the same methods that would be available once this

intervention is deployed outside of the research context. Thus, we offered no financial incentives or human support, and we employed a single-condition, unrandomized design.

2. Methods

2.1. Participants

A convenience sample of participants was recruited worldwide primarily via Google AdWords -- the placement of ads to the right of the search results in the Google search engine. The ads appeared when users searched for depression, depression treatment, and related keywords. Eligible participants were 18 years of age or older, proficient in the English language, with regular access to the Internet and email (at least 3 times per week). No exclusion was made on the basis of depressive symptom scores or geographical location. Of the 1,116 participants who provided enough data to evaluate eligibility, 521 (46.7%) signed the consent to participate in the study. Of these, 309 (59.3%) accessed the course. The rest (n=212) did not access the course, either because they failed to proceed to the end of the baseline assessment (access was granted at the end of baseline assessment; n=109), or because they did not enter the course even after completing baseline for undetermined reasons (n=103).

2.2. Measures

Demographics questionnaire asked general demographic information, i.e., age, gender, race, ethnicity, English language proficiency, as well as frequency of access to the Internet and email.

Quick Inventory of Depressive Symptomatology – Self-Report (QIDS; Rush et al., 2003) is a widely used 16-item self-report questionnaire measuring the severity of depressive symptoms. It assesses the presence and the severity of the nine symptoms that identify the presence of a major depressive episode according to the DSM-IV (APA, 2000). It has achieved good to excellent validity and reliability across numerous studies.

Depression Self-Efficacy questionnaire (DSEQ). The Self-Efficacy Questionnaire for Depression in Adolescents (Tonge et al., 2005) was adapted for adult use, and modified to conform to Bandura (Bandura, 2006) guidelines for creating efficacy questionnaires.

2.3. Procedures

Participants clicking on the Google AdWords ads arrived at the landing page that contained information about the study. Participants then completed the Demographics questionnaire, to determine eligibility. Participants who were ineligible were informed of their ineligibility, and were provided links to other depression resources (i.e., NIH and WHO depression sites). Participants interested in joining the study electronically signed the consent document; they were then asked to verify their understanding of the key points of the consent form by correctly answering a follow-up question about the nature of the study (not a replacement for a mental health professional; responses are not reviewed in real time by a clinician). Those refusing consent were asked to list reasons for their refusal, and provided links to other online depression resources. Consented participants were asked to provide their phone

number (this step could be skipped), though no participant has actually been contacted via phone. Participants then completed several baseline measures, including the QIDS and the DSEQ and were presented detailed feedback on their responses summarized on a single page. Those indicating acute suicidality were shown a statement indicating our concern and were provided a link to Befrienders.org – an international suicide telephone and email helpline database available in over 60 countries in over 20 languages. Participants were then provided with their unique username and password, as well as a link to the Depression Management Course. At 5, 12, and 21 days after consent, an email reminding participants to return to the website was sent.

Follow-up emails were sent to participants at 1, 2, 4, and 7 months after the date of consent. If a participant failed to respond to emails in 4 days, up to 2 additional emails, 4 days apart, were sent. Follow-up emails contained an invitation to return to the site and to complete a follow-up assessment containing QIDS and the DSEQ, as well as to provide any feedback about the site. As in the baseline assessment, participants were offered automated personalized feedback on their responses to the follow-up questionnaires, including responses suggesting suicidality.

Participants were not paid for participation in this study or for completing follow-up assessments. No contact with participants was made aside from the automated emails.

2.4. Intervention structure and content

The intervention consisted of eight lessons based on the classic texts of cognitive therapy for depression (Beck et al., 1979; Beck, 1995), along with other texts and manuals for treatment of depression using Cognitive-Behavioral Therapy (CBT) (e.g., (Lewinsohn et al., 1978; Muñoz et al., 2000). Most lessons contained an educational component, an interactive component, and a homework component. Three of the lessons were based on the standard CBT model (one lesson on behavior activation and two lessons on thoughts and thought processing). Three other lessons introduced content that, while based on the CBT framework, went beyond the most basic CBT approaches; these lessons taught people to make better decisions, to overcome perfectionism while boosting self-esteem, and to create a less depressogenic environment. Additionally, participants had access to the Introduction lesson, the Course Review / Relapse Prevention lesson and the Resources section (several interactive tools, downloadable worksheets and information sheets). All lessons were available to all participants upon first login to the intervention, and participants were able to read them in any order, with the sole exception of the Course Review / Relapse Prevention lesson, which only became available after all other lessons were completed.

Though participants could read the lessons in any order they chose, the actual presentation of the order varied somewhat between participants, depending on their responses to a baseline questionnaire. For instance, if a participant identified frequent low moods as their key concern, the site automatically rearranged the order of lessons on the lesson list, displaying thoughts and though processing lessons higher on the lesson list for that participant. The Introduction and Course Review / Relapse Prevention lessons were always presented first and last, respectively. After the Introduction, the three core CBT lessons were presented, with either the behavior activation lesson first, or the two lessons on thoughts

presented first. After the core lessons, the lessons on decision-making, perfectionism, and environment were presented, in the order determined by the participants' responses.

2.5. Analytical considerations

Characteristics of participants who did or did not enter the intervention were compared using chi-square tests, Fisher's exact tests, and ANOVAs, as appropriate.

The outcome analyses were limited to participants who entered the intervention; those who never entered the intervention were excluded, as their outcomes, positive or otherwise, could not be indicative of the intervention effectiveness. To avoid artificially lowering the scores on the QIDS and the DSEQ, individuals who skipped more than 3 questions on a measure were excluded from a given analysis. Paired t-tests were conducted to test for significant departures from baseline scores on both QIDS and DSEQ for each follow-up point (1, 2, 4, and 7 months). A subsample of individuals with baseline QIDS score of 10 or higher (which is indicative of the presence of a major depressive episode according to previous literature, (Rush et al., 2005; Rush et al., 2003; Trivedi et al., 2004) was also examined to determine whether the intervention is effective for individuals with more severe symptomatology. Two sets of analyses were carried out. The first only examined observed data; individuals who did not provide follow-up data for the follow-up period of interest were excluded. The second employed the last observation carried forward (LOCF) convention, wherein "gaps" in data are filled with data from the most recent available observation. Cohen's d effect sizes were calculated, using the correction for pre- and post-test correlation for within-subjects designs, as described in Morris and DeShon (2002).

3. Results

3.1. Participants

Consenting participants (N = 521) were 34.5 (SD = 12.5) years of age on average. Women comprised two thirds of the sample (66.0%), which is consistent with depression prevalence in the general population; 43.0% reported being married or in a relationship and 56.9% reported being employed. The participants who entered the intervention were just as likely to report being partnered and employed as those who did not, but they also reported being somewhat older (35.6 (SD = 12.8) vs 32.99 (SD = 11.9), $F(1,516) = 5.49$, $p < 0.02$); a somewhat higher proportion of participants who entered the intervention were female compared to participants who did not enter (70.4% vs 59.7%, Fisher's exact test, $p < 0.02$).

Participants' baseline QIDS scores ranged from 2 to 27 (possible range: 0 to 27), with a mean of 14.9 (SD = 4.7), which is indicative of the higher range of moderate depression. Baseline QIDS scores of individuals who entered the intervention were significantly higher than of those who did not (15.3 (SD = 4.5) vs 14.1 (SD = 4.9), $F(1,456) = 7.15$, $p < .01$). Participants' baseline scores on the Depression Self-Efficacy questionnaire ranged from 2 to 96 (possible range: 0 to 100), with a mean of 45.2 (SD = 18.3). No differences were observed in baseline depression self-efficacy scores between those who did and did not enter the intervention. Participants who entered the interventions visited it 3.89 times, on average

(SD=5.39), and saw 1.53 (SD=1.71) separate lessons, for an average of 2.25 (SD=3.49) lesson views.

3.2. Follow-up rate (Figure 1)

Of the 309 participants who entered the intervention, 109 (35.3%) completed the 1 month follow-up, 92 (29.8%) completed the 2 month follow-up, 79 (25.6%) completed the 4 month follow-up, and 72 (23.3%) completed the 7 month follow-up. Considering the entire follow-up period, just over half of the participants 162 (52.4%) completed at least one follow-up, with 65 (21.0%) completing only one follow-up, 37 (12.0%) – two follow-ups, 27 (8.7%) – three follow-ups, and 33 (10.7%) – all four follow-ups.

3.3. Outcomes – full sample

Observed data—As can be seen from Figure 2, there was a gradual decline in observed QIDS scores over the length of the study, and a gradual increase of observed DESQ scores. Across the study period, first to last known QIDS scores for 107 participants have decreased (range of decrease: 1 to 20 points), and for 31 participants the scores have increased (range of increase: 1 to 7 points). QIDS scores for 24 participants remained the same. Comparing the follow-up scores with baseline ratings via paired t-tests, significant differences were found for all comparisons, and the Cohen's d effect sizes ranged from medium to large for QIDS, and from small to medium for DSEQ. Thus, for QIDS, baseline and 1 month follow-up comparison yielded a 0.59 effect size (paired $t(98) = 5.78$, $p < 0.0001$); baseline and 2 month follow-up comparison – 0.82 ES (paired $t(84) = 7.50$, $p < 0.0001$); baseline and 4 month follow-up comparison – 0.84 ES (paired $t(73) = 7.09$, $p < 0.0001$); and baseline and 7 month follow-up comparison – 1.06 ES (paired $t(63) = 8.38$, $p < 0.0001$). With DESQ, baseline and 1 month follow-up comparison yielded a −0.44 effect size (paired $t(68) = -3.53$, $p < 0.001$); baseline and 2 month follow-up comparison – −0.39 ES (paired $t(60) = -3.03$, $p < 0.004$); baseline and 4 month follow-up comparison – −0.46 ES (paired $t(50) = -3.21$, $p < 0.002$); and baseline and 7 month follow-up comparison – −0.67 ES (paired $t(44) = -4.47$, $p < 0.0001$). Note that the sample size for the DESQ is lower because fewer people completed this measure.

LOCF—The observed data analyses may introduce a bias if individuals who return for follow-ups are more likely to have benefited from the intervention. Thus, we repeated the analyses using the LOCF convention. In trials with significant rate of attrition from follow-up, LOCF, being a conservative convention, tends to shrink the size of the effects as it assumes that all individuals lost to follow-up remain in their earlier states vis-à-vis symptom level, thus group follow-up values are usually considerably worse than that of observed data. Nonetheless, follow-up QIDS scores were still significantly lower than baseline scores at each follow-up point, with effect sizes in the small-to-medium range (1 month follow-up: ES = 0.30, paired $t(306) = 5.23$, $p < 0.0001$; 2 month follow-up: ES = 0.30, paired $t(306) = 6.91$, $p < 0.0001$; 4 month follow-up: ES = 0.41, paired $t(306) = 7.17$, $p < 0.0001$; 7 month follow-up: ES = 0.48, paired $t(306) = 8.10$, $p < 0.0001$). Similarly, follow-up DSEQ scores were significantly higher than the baseline scores for all follow-up periods, though the effect size was only in the small range (1 month follow-up: ES = −0.20, paired $t(271) = -3.33$, $p < 0.0001$; 2 month follow-up: ES = −0.23, paired $t(271) = -3.84$, $p < 0.0001$; 4 month follow-

up: ES = −0.24, paired t(271) = −3.98, p < 0.0001; 7 month follow-up: ES = −0.28, paired t(271) = −4.50, p < 0.0001). Please see Figure 2 for a graphical illustration of the scores.

3.4. Outcomes – depressed sample

Observed data—To determine whether the intervention is effective for individuals with higher depressive symptoms, we excluded individuals whose baseline scores on the QIDS were lower than 10 (Rush et al., 2005; Rush et al., 2003; Trivedi et al., 2004), resulting in a sample of 271 participants. For this sub-sample, using observed data, there were still significant and substantial differences between baseline and follow-up scores on the QIDS, with the effect size being in the "large" range (1 month follow-up: ES = 0.73, paired t(84) = 6.36, p < 0.0001; 2 month follow-up: ES = 0.89, paired t(75) = 7.46, p < 0.0001; 4 month follow-up: ES = 1.06, paired t(63) = 7.71, p < 0.0001; 7 month follow-up: ES = 1.30, paired t(56) = 9.32, p < 0.0001). Similarly, DSEQ scores were significantly higher at follow-ups than they were for baseline, with effect sizes mostly in the medium range (1 month follow-up: ES = −0.44, paired t(58) = −3.26, p < 0.002; 2 month follow-up: ES = −0.45, paired t(53) = −3.31, p < 0.002; 4 month follow-up: ES = −0.38, paired t(43) = −2.48, p < 0.02; 7 month follow-up: ES = −0.79, paired t(38) = −4.82, p < 0.002).

LOCF—Using observed data while imposing lower limits on severity can inadvertently inflate differences between baseline and follow-ups due to regression to the mean, that is, some individuals may naturally return to an improved state over the course of the study; these individuals may also be more likely to return for follow-up. However, in the presence of lower limit on severity, LOCF convention becomes even more conservative, as scores indicative of greater severity are retained in case of non-follow-up. Using the LOCF analyses, the differences between QIDS baseline and follow-up scores remained significant, and the effect sizes remained virtually the same as in the full sample of participants (1 month follow-up: ES = 0.34, paired t(270) = 5.53, p < 0.0001; 2 month follow-up: ES = 0.43, paired t(270) = 7.03, p < 0.0001; 4 month follow-up: ES = 0.50, paired t(270) = 7.58, p < 0.0001; 7 month follow-up: ES = 0.55, paired t(270) = 8.45, p < 0.0001). The same was true for DSEQ, where significance was retained and effect sizes remained largely unchanged (1 month follow-up: ES = −0.20, paired t(238) = −3.07, p < 0.002; 2 month follow-up: ES = −0.28, paired t(238) = −4.15, p < 0.0001; 4 month follow-up: ES = −0.27, paired t(238) = −4.02, p < 0.0001; 7 month follow-up: ES = −0.32, paired t(238) = −4.69, p < 0.0001).

4. Discussion

The purpose of this investigation was to understand the efficacy of an unsupported, fully automated internet intervention for depressive symptoms. One of the key goals of this investigation was to maximize ecological validity. Thus, we attempted to avoid those aspects of the trial that depart from the manner in which this intervention site will eventually be disseminated. Specifically, we did not provide financial incentives, either to encourage the use of the site or for providing follow-up data (Clarke et al., 2005; Clarke et al., 2002); we also did not require an interview with a live person to be accepted into the study (Berger et al., 2011; Spek et al., 2007; Vernmark et al., 2010) or to provide follow-up data (Christensen et al., 2004a). Additionally, as it is possible that requiring individuals to

consent to be randomized may exclude individuals with strong preferences or those who desire more control over their treatment decisions (Brewin and Bradley, 1989; Ward et al., 1999), we used a single-condition pre-post design with minimal exclusion criteria. We were able to consent 521 individuals for the trial, of whom 309 entered the intervention site; data from these 309 participants were used for outcome analyses. Importantly, individuals who entered the intervention site had higher baseline depression scores, suggesting that resources such as internet interventions attract those who need them most.

The results from this pragmatic trial were encouraging. We found that depressive symptoms as measured by the Quick Inventory of Depressive Symptomatology reduced significantly and substantially from baseline to the first follow-up, and remained lower for all subsequent follow-ups. Though our study design did not allow for a control group, our pre-post effect sizes are consistent with those previously reported in the literature (Hedman et al., 2014; Spek et al., 2007; Van Voorhees et al., 2005; Wagner et al., 2014), which offers more confidence that these observed declines in symptom levels are due to the beneficial effects of the intervention and not to other factors, such as spontaneous remission. Our finding of increases in depression self-efficacy from baseline to follow-up lends further support to the effectiveness of our intervention. The Depression Self-Efficacy Questionnaire measures the ability to manage and improve one's symptoms of depression. Whereas symptoms can improve with spontaneous remission, this is not necessarily true for the skills and abilities needed to manage symptoms; yet symptom management is the very skillset CBT aims to improve in individuals. The fact that self-efficacy ratings improved is therefore very encouraging.

Internet interventions tend to have relatively high attrition rates. The problem is exacerbated with unsupported interventions (Christensen et al., 2006; Donker et al., 2013), and indeed, in this study, while the follow-up rate is certainly less than ideal, it is in fact similar to that found with other unsupported interventions (Donker et al., 2013; Newman et al., 2011). High attrition rates raise concerns regarding the representativeness of those who do return for follow-up, specifically that these individuals might be the ones who benefited from the intervention the most, which increases effect sizes. To address these concerns, we conducted analyses using the "last observation carried forward" convention, which retains the last available data point in place of missing data. This is a conservative strategy that effectively "penalizes" attrition because it assumes that individuals who drop out remain unimproved. In trials with high attrition rates this "penalty" is obviously greater, as a greater proportion of participants are assumed to be unimproved. The fact that we have found significant pre-post differences in our study using LOCF analyses, especially given our high attrition rate, brings more confidence to our findings, which was further confirmed by the similarity of the effect sizes to other unsupported studies (Meyer et al., 2009). When a sample is further constrained to those with higher symptom levels, as we have done in our analysis of those scoring 10 or higher on the QIDS, the "penalty" imposed by LOCF is likely even higher (participants who are lost to follow-up cannot have low baseline symptoms). That we have found significant reduction in depression symptoms and significant increases in depression self-efficacy, even with LOCF, and even when excluding low-symptom participants is strongly indicative of the effectiveness of our intervention.

This study has limitations that should be acknowledged. Most of these are a consequence of eschewing the methodology of traditional randomized controlled trials with carefully selected participants in favor of more ecologically valid "pragmatic" approaches. Thus, a control group was not included in this study, which only enabled us to conduct within-subject pre-post analyses. We have attempted to analyze these in ways that minimizes possible bias, and looked to other literature to determine whether our effect sizes are consistent with those found by other teams. However, a comparison with a control group, conducting active cohort maintenance strategies (e.g., personal follow-up, financial or other incentives), as well as establishing a more rigid selection criteria that could potentially minimize error variance would have yielded a more accurate estimate of the efficacy of our intervention (while also compromising ecological validity and the understanding of dissemination potential). Using Google AdWords as a recruitment method limited participants to those who use the Google search engine and click on their advertisements. The results may not generalize to individuals who are referred to this site via other means (e.g., social networks, public health websites, etc.). Finally, our participants were computer literate individuals looking for depression information online, which is the intended audience for unsupported internet interventions for depression, including ours; results may therefore not be generalizable to lower literacy populations, to people who do not use the Internet, or to those who are not at least somewhat aware of the nature of their symptoms.

Like any other treatments, unsupported internet interventions cannot be beneficial for all individuals with depressive symptoms. Indeed, we found that for 31 of our participants, depression scores have worsened by 1 to 7 points on the QIDS. Though this number is much smaller than the 107 participants for whom depression scores have decreased (by up to 20 points), this finding is nonetheless troubling. It is difficult to know whether the increase in symptom level is truly iatrogenic (that is, a direct consequence of using the intervention, and would not have occurred had this person not used the intervention), or due to lack of response to this intervention along with the natural worsening of depression. Nonetheless, this finding should be a reminder that internet interventions cannot serve all users and that we are likely to find that, as in face-to-face interventions (Coffman et al., 2007; Koenig et al., 2014; Lampropoulos, 2011) some people will worsen during treatment. There may be several reasons for lack of response. This intervention, like most other internet interventions, is based on CBT, and some individuals with depressive symptoms are less likely to respond to this treatment modality, though they might respond to another type of psychotherapy or to a medication. Internet interventions may be less likely to be meaningful to individuals with highly idiosyncratic symptom presentations or depression etiologies. This is likely especially true for unsupported interventions, where there is no therapist to offer guidance beyond intervention programming. Unlike supported treatments, unsupported interventions rely more heavily on the attractiveness of the program and user engagement, as there is no social commitment to a therapist of a coach to use the site, which points strongly to the need of developing interventions that would be appealing and user-friendly.

Unsupported internet interventions are scalable, cost-effective, and can be efficacious; such interventions can therefore eventually reduce health disparities (Muñoz, 2010) by making treatment options available for individuals who are unable or unwilling to pursue traditional treatment resources. With increasing internet penetration worldwide, individuals with

internet access, either personal or public, can gain access to information, resources, and interventions, and indeed, the majority of internet users is already searching online for health information (Fox, 2006; Powell et al., 2003). Internet penetration has been especially rapid in traditionally underserved markets, such as Africa, the Middle East, and Latin America, and given the paucity of existing services for these markets, internet interventions can be an invaluable tool for millions of individuals who may otherwise remain untreated. Unsupported internet interventions are perhaps the only treatment tool scalable enough to provide empirically valid treatment globally for individuals in need. Developing, evaluating, and making such interventions available has the potential to make a substantial difference in global public health.

Acknowledgments

References

Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cognitive behaviour therapy. 2009; 38:196–205. [PubMed: 20183695]

Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. PloS One. 2010; 5:e13196. [PubMed: 20967242]

APA. Diagnostic and statistical manual of mental disorders. Author; Washington DC: 2000.

Bandura, A. Guide for constructing self-efficacy scales.. In: Pajares, F.; Urdan, T., editors. Self-efficacy beliefs of adolescents. Information Age Publishing; 2006.

Beck, AT.; Rush, AJ.; Shaw, BF.; Emery, G. Cognitive therapy of depression. Guilford Press; New York: 1979.

Beck, J. Cognitive Therapy: Basics and Beyond. Guilford Press; New York, NY: 1995.

Berger T, Hammerli K, Gubser N, Andersson G, Caspar F. Internet-based treatment of depression: a randomized controlled trial comparing guided with unguided self-help. Cognitive behaviour therapy. 2011; 40:251–266. [PubMed: 22060248]

Brewin CR, Bradley C. Patient preferences and randomised clinical trials. British Medical Journal. 1989; 299:313–315. [PubMed: 2504416]

Christensen H, Griffiths K, Groves C, Korten A. Free range users and one hit wonders: community users of an Internet-based cognitive behaviour therapy program. Aust N Z J Psychiatry. 2006; 40:59–62. [PubMed: 16403040]

Christensen H, Griffiths KM, Jorm AF. Delivering interventions for depression by using the internet: randomised controlled trial. British Medical Journal. 2004a; 328:265. [PubMed: 14742346]

Christensen H, Griffiths KM, Korten A. Web-based cognitive behavior therapy: analysis of site usage and changes in depression and anxiety scores. Journal of Medical Internet Research. 2002; 4:e3. [PubMed: 11956035]

Christensen H, Griffiths KM, Korten AE, Brittliffe K, Groves C. A comparison of changes in anxiety and depression symptoms of spontaneous users and trial participants of a cognitive behavior therapy website. J Med Internet Res. 2004b; 6:e46. [PubMed: 15631970]

Clarke G, Eubanks D, Reid E, Kelleher C, O'Connor E, DeBar LL, Lynch F, Nunley S, Gullion C. Overcoming Depression on the Internet (ODIN) (2): a randomized trial of a self-help depression skills program with reminders. J Med Internet Res. 2005; 7:e16. [PubMed: 15998607]

Clarke G, Reid E, Eubanks D, O'Connor E, DeBar LL, Kelleher C, Lynch F, Nunley S. Overcoming depression on the Internet (ODIN): a randomized controlled trial of an Internet depression skills intervention program. Journal of Medical Internet Research. 2002; 4:e14. [PubMed: 12554545]

Coffman SJ, Martell CR, Dimidjian S, Gallop R, Hollon SD. Extreme nonresponse in cognitive therapy: can behavioral activation succeed where cognitive therapy fails? J Consult Clin Psychol. 2007; 75:531–541. [PubMed: 17663608]

Donker T, Bennett K, Bennett A, Mackinnon A, van Straten A, Cuijpers P, Christensen H, Griffiths KM. Internet-delivered interpersonal psychotherapy versus internet-delivered cognitive behavioral therapy for adults with depressive symptoms: randomized controlled noninferiority trial. J Med Internet Res. 2013; 15:e82. [PubMed: 23669884]

Fox S. Online health search 2006: Most Internet users start at a search engine when looking for health information online, very few check the source and date of the information they find. Pew Internet & American Life Project. 2006

Fridrici M, Lohaus A, Glass C. Effects of incentives in web-based prevention for adolescents: Results of an exploratory field study. Psychology & health. 2009; 24:663–675. [PubMed: 20205019]

Griffiths KM, Farrer L, Christensen H. The efficacy of internet interventions for depression and anxiety disorders: a review of randomised controlled trials. The Medical journal of Australia. 2010; 192:S4–11. [PubMed: 20528707]

Hedman E, Ljótsson B, Kaldo V, Hesser H, El Alaoui S, Kraeplin M, Andersson E, Ruck C, Svanborg C, Andrerssonm G, Lindefors N. Effectiveness of Internet-based cognitive behaviour therapy for depression in routine psychiatric care. Journal of affective disorders. 2014; 155:49–58. [PubMed: 24238951]

Johansson R, Andersson G. Internet-based psychological treatments for depression. Expert review of neurotherapeutics. 2012; 12:861–869. quiz 870. [PubMed: 22853793]

Judd LL, Rapaport MH, Paulus MP, Brown JL. Subsyndromal symptomatic depression: a new mood disorder? The Journal of clinical psychiatry. 1994; 55(Suppl):18–28. [PubMed: 8077164]

Koenig AM, Jarrett RB, Gallop R, Barrett MS, Thase ME. Extreme nonresponse to acute phase cognitive therapy for depression: an attempt to replicate and extend. Behavior therapy. 2014; 45:300–313. [PubMed: 24680227]

Lampropoulos GK. Failure in psychotherapy: an introduction. J Clin Psychol. 2011; 67:1093–1095. [PubMed: 21964920]

Lewinsohn, PM.; Muñoz, RF.; Youngren, MA.; Zeiss, AM. Control your depression. Prentice-Hall, Englewood Cliffs; NJ: 1978.

Leykin Y, Aguilera A, Torres LD, Perez-Stable EJ, Muñoz RF. Interpreting the outcomes of automated internet-based randomized trials: example of an International Smoking Cessation Study. J Med Internet Res. 2012a; 14:e5. [PubMed: 22314016]

Leykin Y, Muñoz RF, Contreras O. Are consumers of Internet health information "cyberchondriacs"? Characteristics of 24,965 users of a depression screening site. Depression and Anxiety. 2012b; 29:71–77. [PubMed: 21681872]

Meyer B, Berger T, Caspar F, Beevers CG, Andersson G, Weiss M. Effectiveness of a novel integrative online treatment for depression (Deprexis): randomized controlled trial. J Med Internet Res. 2009; 11:e15. [PubMed: 19632969]

Morris SB, DeShon RP. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. Psychological methods. 2002; 7:105–125. [PubMed: 11928886]

Muñoz RF. Using evidence-based internet interventions to reduce health disparities worldwide. J Med Internet Res. 2010; 12:e60. [PubMed: 21169162]

Muñoz RF, Barrera AZ, Delucchi K, Penilla C, Torres LD, Pérez-Stable EJ. International Spanish/ English Internet smoking cessation trial yields 20% abstinence rates at 1 year. Nicotine & tobacco research. 2009; 11:1025. [PubMed: 19640833]

Muñoz RF, Ippen C,G, Rao S, Le H-N, Dwyer EV. The manual for group cognitive behavioral therapy for major depression. 2000November 19, 2007from http://www.medschool.ucsf.edu/latino/manuals.aspx#GroupCognitiveBehavioralTherapyofMajorDepression University of California, San Francisco School of Medicine Website.

Newman MG, Szkodny LE, Llera SJ, Przeworski A. A review of technology-assisted self-help and minimal contact therapies for anxiety and depression: is human contact necessary for therapeutic efficacy? Clin Psychol Rev. 2011; 31:89–103. [PubMed: 21130939]

O'Kearney R, Gibson M, Christensen H, Griffiths KM. Effects of a cognitive-behavioural internet program on depression, vulnerability to depression and stigma in adolescent males: a school-based controlled trial. Cognitive behaviour therapy. 2006; 35:43–54. [PubMed: 16500776]

O'Kearney R, Kang K, Christensen H, Griffiths K. A controlled trial of a school-based Internet program for reducing depressive symptoms in adolescent girls. Depress Anxiety. 2009; 26:65–72. [PubMed: 18828141]

Powell J, Hamborg T, Stallard N, Burls A, McSorley J, Bennett K, Griffiths KM, Christensen H. Effectiveness of a web-based cognitive-behavioral tool to improve mental well-being in the general population: randomized controlled trial. J Med Internet Res. 2013; 15:e2. [PubMed: 23302475]

Powell JA, Darvell M, Gray JAM. The doctor, the patient and the world-wide web: how the internet is changing healthcare. Journal of the Royal Society of Medicine. 2003; 96:74. [PubMed: 12562977]

Rush AJ, Trivedi MH, Carmody TJ, Ibrahim HM, Markowitz JC, Keitner GI, Kornstein SG, Arnow B, Klein DN, Manber R, Dunner DL, Gelenberg AJ, Kocsis JH, Nemeroff CB, Fawcett J, Thase ME, Russell JM, Jody DN, Borian FE, Keller MB. Self-reported depressive symptom measures: sensitivity to detecting change in a randomized, controlled trial of chronically depressed, nonpsychotic outpatients. Neuropsychopharmacology. 2005; 30:405–416. [PubMed: 15578008]

Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, Markowitz JC, Ninan PT, Kornstein S, Manber R, Thase ME, Kocsis JH, Keller MB. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biological Psychiatry. 2003; 54:573–583. [PubMed: 12946886]

Shafran R, Clark DM, Fairburn CG, Arntz A, Barlow DH, Ehlers A, Freeston M, Garety PA, Hollon SD, Ost LG, Salkovskis PM, Williams JM, Wilson GT. Mind the gap: Improving the dissemination of CBT. Behav Res Ther. 2009; 47:902–909. [PubMed: 19664756]

Spek V, Nyklíŉek I, Smits N, Cuijpers P, Riper H, Keyzer J, Pop V. Internet-based cognitive behavioural therapy for subthreshold depression in people over 50 years old: a randomized controlled clinical trial. Psychological medicine. 2007; 37:1797–1806. [PubMed: 17466110]

Tonge B, King N, Klimkeit E, Melvin G, Heyne D, Gordon M. The Self-Efficacy Questionnaire for Depression in Adolescents (SEQ-DA). European child & adolescent psychiatry. 2005; 14:357–363. [PubMed: 16254764]

Trivedi MH, Rush AJ, Ibrahim HM, Carmody TJ, Biggs MM, Suppes T, Crismon ML, Shores-Wilson K, Toprac MG, Dennehy EB. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. Psychological medicine. 2004; 34:73–82. [PubMed: 14971628]

Van't Hof E, Cuijpers P, Stein DJ. Self-help and Internet-guided interventions in depression and anxiety disorders: a systematic review of meta-analyses. CNS spectrums. 2009; 14:34–40. [PubMed: 19238128]

Van Voorhees BW, Ellis J, Stuart S, Fogel J, Ford DE. Pilot study of a primary care internet-based depression prevention intervention for late adolescents. The Canadian child and adolescent psychiatry review = La revue canadienne de psychiatrie de l'enfant et de l'adolescent. 2005; 14:40–43.

Vernmark K, Lenndin J, Bjarehed J, Carlsson M, Karlsson J, Oberg J, Carlbring P, Eriksson T, Andersson G. Internet administered guided self-help versus individualized e-mail therapy: A randomized trial of two versions of CBT for major depression. Behav Res Ther. 2010; 48:368–376. [PubMed: 20152960]

Wagner B, Horn AB, Maercker A. Internet-based versus face-to-face cognitive-behavioral intervention for depression: a randomized controlled non-inferiority trial. Journal of affective disorders. 2014; 152-154:113–121. [PubMed: 23886401]

Waller G. Evidence-based treatment and therapist drift. Behav Res Ther. 2009; 47:119–127. [PubMed: 19036354]

Ward E, King M, Lloyd M, Bower P, Friedli K. Conducting randomized trials in general practice: methodological and practical issues. The British journal of general practice : the journal of the Royal College of General Practitioners. 1999; 49:919–922. [PubMed: 10818663]

World Health Organization. [September 12, 2010] Depression. 2010. from http://www.who.int/mental_health/management/depression/defintion/en/

Highlights

- 309 participants recruited worldwide entered an internet intervention for depression

- No financial incentives or human support was provided; 52.4% completed at least one follow-up

- Significant reduction of depression symptoms and increases in depression self-efficacy were observed

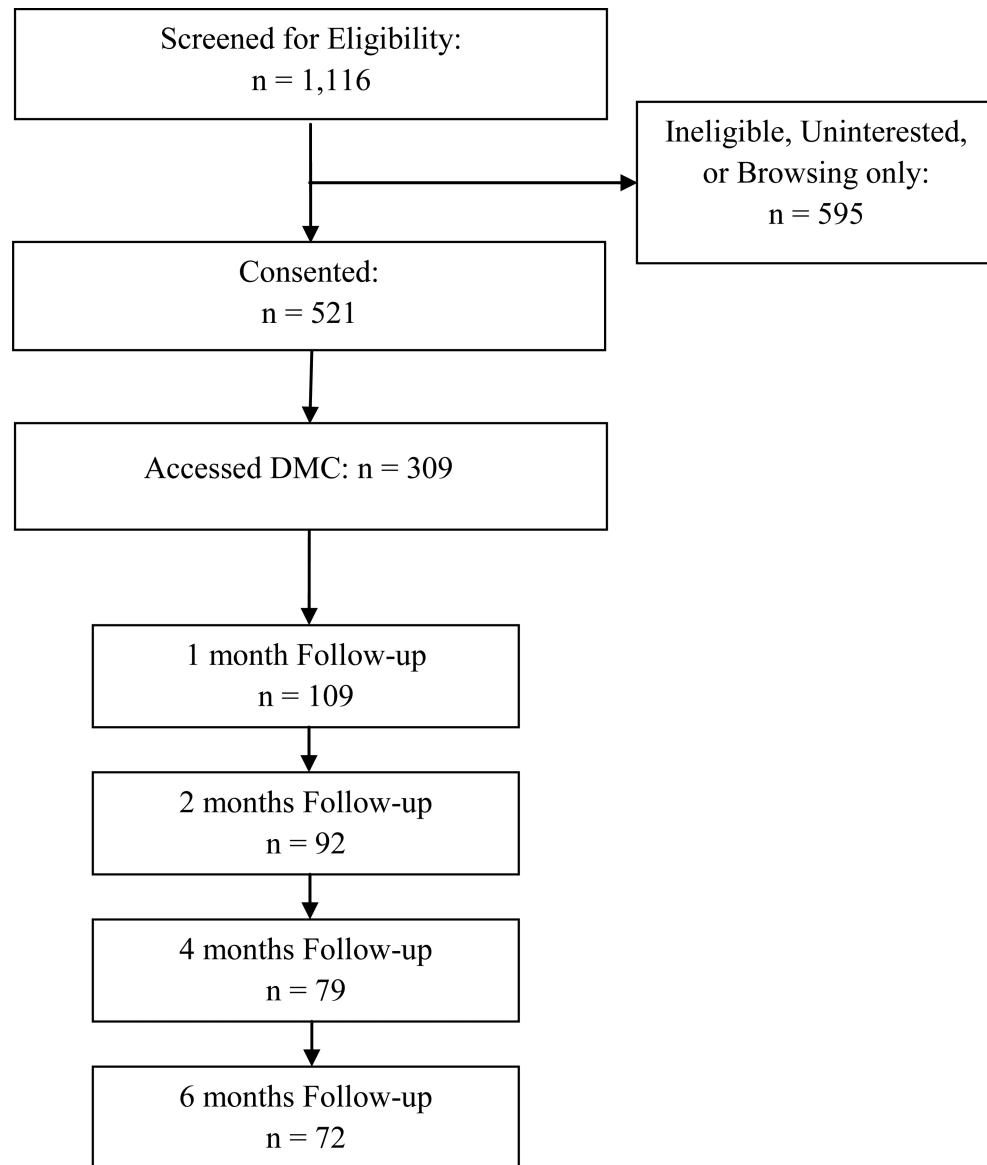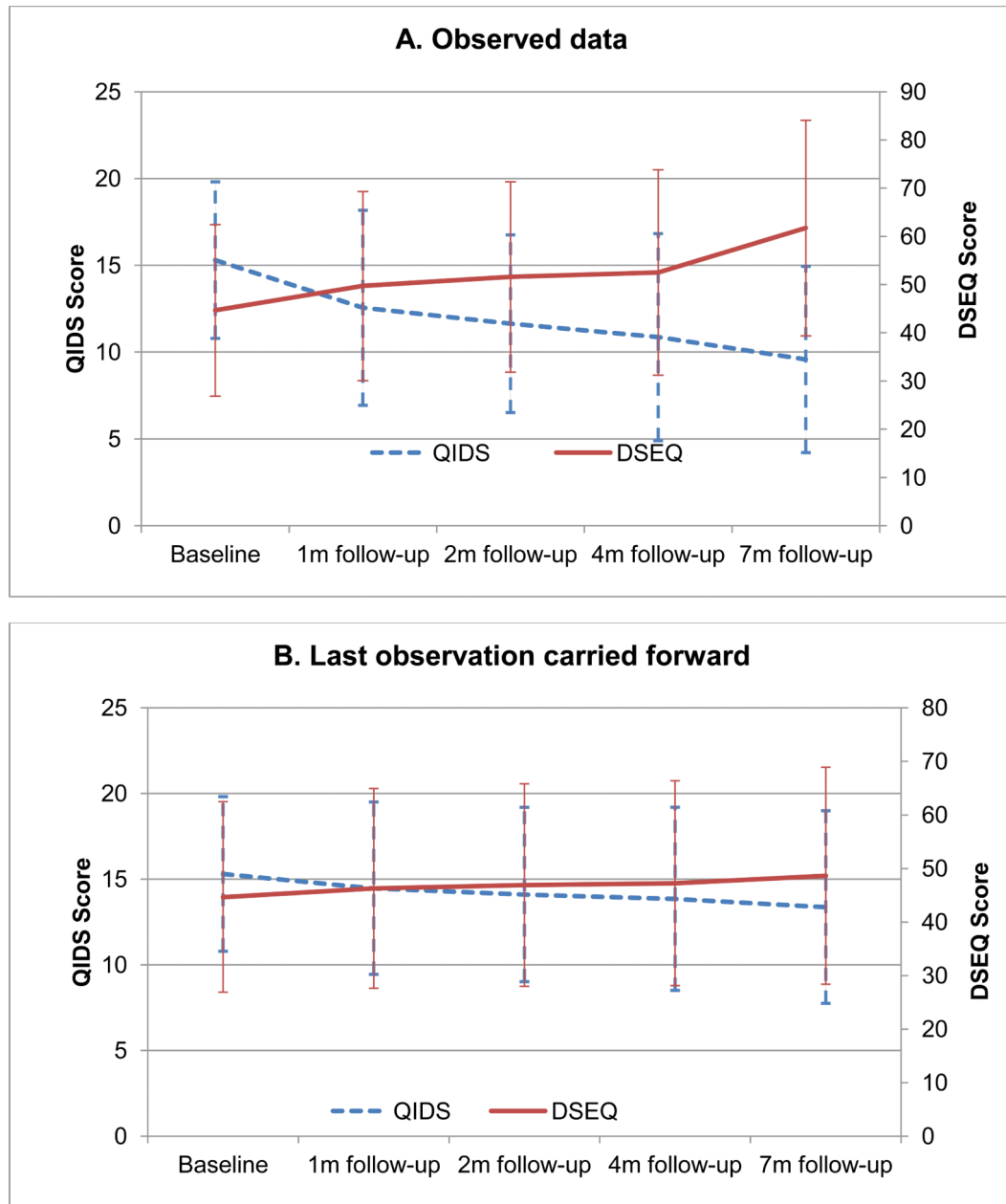- Results remained significant using LOCF convention, and with the higher severity subsample

Figure 1.
CONSORT diagram

Note: The error bars denote standard deviation. Broken line error bars denote SD of DESQ mean scores; solid line error bars denote SD of QIDS mean scores.

Figure 2.
Depression symptom and depression self-efficacy scores throughout the study, using A. observed data, and B. Last observation carried forward.