Internet cognitive behavioural treatment for mixed anxiety and depression: a randomised controlled trial and evidence of effectiveness in primary care

Author:
Newby, Jill; Mackenzie, Anna; Williams, Alishia; McIntyre, Karen; Watts, Sarah; Wong, Nora; Andrews, Andrews

Publication details:
Psychological Medicine
0033-2917 (ISSN)

Publication Date:
2013

Publisher DOI:
http://dx.doi.org/10.1017/S0033291713000111

Internet cognitive behavioural treatment for mixed anxiety and depression: a randomised controlled trial and evidence of effectiveness in primary care

Reference: Newby, J. M., Mackenzie, A., Williams, A. D., McIntyre, K., Watts, S., Wong, N., et al. (2013). Internet cognitive behavioural therapy for mixed anxiety and depression: A randomized controlled trial and evidence of effectiveness in primary care. *Psychological Medicine*, 1-14. doi: 10.1017/s0033291713000111.

Jill M. Newby[1,2]

Anna Mackenzie[1]

Alishia D. Williams[1]

Karen McIntyre[1]

Sarah Watts[1]

Nora Wong[1]

Gavin Andrews [1]

1. Clinical Research Unit for Anxiety and Depression, School of Psychiatry, University of New South Wales at St Vincent's Hospital, Darlinghurst, NSW, Australia.
2. Jill Newby is now at the MRC Cognition and Brain Sciences Unit, Cambridge, United Kingdom

Corresponding Author: Jill Newby, MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge, United Kingdom, CB2 7EF, email: jill.newby@mrc-cbu.cam.ac.uk

Keywords: cognitive behaviour therapy; transdiagnostic; depression; anxiety, internet; effectiveness; treatment outcome

Abstract

Background: Major Depressive Disorder (MDD) and Generalised Anxiety Disorder (GAD) have the highest comorbidity rates within the internalising disorders cluster, yet no internet-based Cognitive Behavioural Treatment (iCBT) exists for their combined treatment. **Methods:** We designed a 6-lesson therapist-assisted iCBT program for mixed anxiety and depression. Study 1 was a Randomised Controlled Trial (RCT) comparing the iCBT program (*n*=46) versus Wait-List Control (WLC, *n*=53) for patients diagnosed by structured clinical interview with MDD, GAD or co-morbid GAD/MDD. Primary outcome measures were the PHQ-9 (depression), GAD-7 (generalised anxiety), K-10 (distress) and WHODAS-II (disability). The iCBT group was followed-up at 3-months post-treatment. In Study 2, we investigated the adherence to, and efficacy of the same program in a primary care setting, where patients (*n*=136) completed the program under the supervision of primary care clinicians. **Results:** The RCT showed that the iCBT program was more effective than WLC, with large within- and between-groups effect sizes found (>.8). Adherence was also high (89%), and gains were maintained at 3 month-follow-up. In Study 2 in primary care, adherence to the iCBT program was low (41%), yet effect sizes were large (>.8). Thirty per cent of non-completers experienced benefit. **Conclusions:** Together the results show that iCBT is effective and adherence is high in research settings, but there is a problem of adherence when translated into the 'real world.' Future efforts need to be placed on developing improved adherence to iCBT in primary care settings.

Major Depressive Disorder (MDD) and Generalised Anxiety Disorder (GAD) are highly comorbid, with 58-70% comorbidity rates (Brown *et al.* 2001). Compared to having either disorder alone, comorbid GAD/MDD is associated with longer episode duration, poorer prognosis and greater impairment (e.g., Kessler *et al.* 1999). Although both disorders can be treated effectively with Cognitive Behavioural Therapy (CBT) (Cuijpers *et al.* 2008; Stewart and Chambless 2009), it remains unclear whether single-disorder CBT protocols provide the most optimal approach for treating individuals with both GAD and MDD. First, there is some evidence that single-disorder treatments may not adequately address comorbidity (van Balkom *et al.* 2008). For example, CBT for GAD reduces depressive symptoms immediately post-treatment, but these gains are not sustained over longer-term follow-up (Newman *et al.* 2010). Second, MDD and GAD share many common biological, environmental and temperamental risk factors (neuroticism, Andrews 1990; Hettema *et al.* 2006), as well as similar cognitive, behavioural and emotional processes that maintain both disorders. For these reasons, CBT protocols for MDD and GAD share many common elements (e.g., cognitive structuring) leading to redundancy across treatment protocols (see Chorpita and Daleiden 2009). Third, the provision of sequential single-disorder treatments for those with GAD and MDD may not always be practical in clinical settings where resources are scarce, the number of trained clinicians in multiple evidence-based protocols is low, and number of sessions are restricted (Norton 2012).

These reasons have motivated the development of new transdiagnostic treatments that distil common elements of CBT to target shared maintaining factors across anxiety and mood disorders, with promising results (Barlow *et al.* 2004; Norton 2012). Such treatments are acceptable to patients, and effective in treating primary and comorbid diagnoses (McEvoy *et al.* 2009). Preliminary evidence suggests that transdiagnostic treatments for anxiety disorders are as effective as single-disorder treatments (Norton 2012). Transdiagnostic treatments have

the added benefit of minimising burden to patients, clinicians and health care systems (e.g. Page & Hooke, 2003; Lumpkin et al, 2002; Norton, 2008), potentially enhance the efficiency of treatment (Chu 2012), and make it easier to disseminate treatments (Weisz *et al.* 2012).

Internet-delivered CBT (iCBT) interventions offer the advantage of being more cost-effective than face-to-face treatment (Hedman *et al.* 2011), and are accessible to many people with anxiety and depression who do not seek face-to-face help (Andrews *et al.* 2001). iCBT is an effective treatment for depression (Andersson and Cuijpers 2009), GAD (Robinson *et al.* 2010), and for mixed anxiety disorders (Titov *et al.* 2010), with large effect sizes (0.88, NNT=2.13) and low drop-out rates (Andrews *et al.* 2010). A recent randomised controlled trial (RCT) by Titov *et al.* (2011) found that an 8-lesson iCBT program for comorbid depression and three anxiety disorders (GAD, Social Phobia and Panic Disorder) was more effective than wait-list control, with moderate effect sizes (.58 and .52 for depression and anxiety respectively) (see Dear *et al.* 2011 for an open trial of a 5-lesson version).

We extended beyond Titov *et al.* (2011), and designed a shorter 6-lesson clinician-assisted iCBT treatment for mixed GAD and MDD, called the *Worry and Sadness Program*. We sought to explore whether their modest effect sizes could be enhanced by restricting the focus on GAD and MDD, and by placing greater emphasis on treating maladaptive rumination and worry, key maintaining factors for depression and anxiety (Nolen-Hoeksema *et al.* 2008). Study 1 was a RCT of the Worry and Sadness Program, compared to Wait-List Control Group. Study 2 investigated the effectiveness of the same program, when prescribed by primary care practitioners to their patients via www.crufadclinic.org (now www.thiswayup.org.au/clinic). We expected that adherence would be better in the RCT compared to primary care, whereas the effectiveness of the program would be similar across both settings.

Study 1: A Randomised Controlled Trial of the Worry and Sadness Program

Method

Design

A CONSORT 2010 compliant (Schulz *et al.* 2010) RCT design was used to compare an immediate treatment group to a deferred-treatment group (wait-list control, WLC). The immediate treatment group was followed up until 3 months post-treatment. The WLC were enrolled in the iCBT course after the treatment group had completed the program. A between-groups effect size (ES) of 0.6 with power of 80% was expected as achieved in prior studies (e.g., Titov *et al.* 2011). A minimum of 45 participants were required in each group, but 110 were recruited to hedge against attrition.

Participants

Participants were recruited from an existing wait-list of individuals who had previously expressed interest in participating in iCBT, and from an online advertisement posted on the www.virtualclinic.org.au website. Participants applied online to www.virtualclinic.org.au after reading details about the study, including the eligibility criteria for inclusion as follows: (i) Aged over 18, (ii) Self-identified as suffering from mild or moderate GAD, MDD or mixed anxiety-depressive disorder (ICD-10 F41.2), and with both GAD-7 and PHQ-9 scores above clinical threshold, (iii) Prepared to provide name, phone number and address, and the name and address of their local general practitioner, (iv) Had access to a phone, computer and printer, (v) Had maintained a stable dosage for at least 2 months prior to participation if they were receiving current pharmacological and/or psychotherapy treatment, and agreed to not make any changes to their treatment during the entire duration of the study.

Details of participant flow are in Figure 1. Eighty seven applicants were excluded after completing initial online screening questions. One hundred and thirty five applicants met the online selection criteria, provided informed consent, and then participated in a brief phone interview (nine additional people passed online screening but were unable to be contacted for telephone interview). Trained interviewers administered the Mini International Neuropsychiatric Interview Version 5.0.0 (MINI) (Sheehan *et al.* 1998) to confirm whether the applicant met DSM-IV criteria for GAD and/or MDD. Twenty-six individuals were excluded after telephone interview, leaving 109 applicants who met inclusion criteria and were randomised. Random numbers were generated via www.random.org by a team member who was not involved in the study. Concealment of allocation was maintained until the applicant met all inclusion criteria and an offer of participation was made. The study was approved by the Human Research Ethics Committee (HREC) of St Vincent's Hospital (Sydney, Australia) (HREC 11/SVH/95), and the trial was registered as ACTRN12611001055998.

[Insert Figure 1 about here]

Primary Outcome Measures

The Mini International Neuropsychiatric Interview Version 5.0.0 (MINI, Sheehan *et al.* 1998) MDD, GAD and risk assessment modules were administered to assess current and lifetime DSM-IV diagnoses. Due to practical constraints, other comorbid diagnoses were not assessed. The MINI possesses excellent inter-rater reliability (k=.88.-1.00) and good concurrent validity with the Composite International Diagnostic Interview (CIDI, World Health Organisation, 1990, see Kessler and Ustun 2004).

The *Patient Health Questionnaire-9* (PHQ-9, Kroenke et al. 2001) and the *Generalised Anxiety Disorder* 7-item scale (GAD-7, Spitzer *et al.* 2006) measured depression and

generalised anxiety symptoms respectively over the past fortnight. Both scales have good psychometric properties, and a cut-off score ≥10 is used to define probable DSM-IV diagnoses of MDD and GAD (Kroenke *et al.* 2007; Wittkampf *et al.* 2007). Internal reliabilities for the current sample were good (current sample α's: PHQ-9 = .75-.87; GAD-7 =.85 -.91). *The Kessler 10-item Psychological Distress scale* (K-10, Kessler *et al.* 2002) measured non-specific psychological distress over the past 30 days. The K-10 has excellent psychometric properties (Furukawa *et al.* 2003), and higher scores indicate higher distress (current sample α = .76-.89).

Secondary Outcome Measures

*The 12-item World Health Organisation Disability Assessment Schedule (*WHODAS - II) (World Health Organization.) measured functional impairment and activity limitation (higher scores indicate higher impairment). The WHODAS-II possesses good psychometric properties (Andrews *et al.* 2009) (current sample α=.85 - .87). The *Beck Depression Inventory – Second Edition* (BDI-II, Beck *et al.* 1996) measured depressive symptoms in the past fortnight, and has good psychometrics, including high internal consistency (Beck *et al.* 1996) and comparable psychometric properties for online versus pen-and-paper administration (Hollandare *et al.* 2010) (α =.84 -.92 for the current sample). *The Penn State Worry Questionnaire* (PSWQ, Meyer *et al.* 1990) measured trait worry, and has good reliability and validity (Brown *et al.* 1992) (α =.87- .88 for the current sample). The *NEO-Five Factor Inventory – Neuroticism Subscale* (NEO-FFI-N, Costa and McCrae 1985) measured the personality dimension of Neuroticism, and has good psychometric properties (Cuijpers *et al.* 2005) (current sample α = .58-.78).

Description of Treatment

The Worry and Sadness Program, delivered via www.virtualclinic.org.au, consists of six online lessons to be completed over a 10 week period (19 October to 23 December, 2011). Lesson content is presented in the form of an illustrated story about two fictional characters who experience anxiety and depression, and gain mastery over their symptoms using CBT techniques (e.g., activity scheduling[1]; see Table 1 for course content). Following each lesson, participants download and print out a lesson summary, which includes practical homework (e.g., graded exposure tasks). Participants have access to (i) frequently asked questions for each lesson, (ii) "Patient Recovery Stories" from former patients of www.virtualclinic.org.au, and (iii) extra resources on: good sleep, activity planning, assertiveness, pleasant events, conversational skills, hunt for the positives, medications, panic attacks, structured problem solving, thought challenging, worry time, and worry stories (imaginal worry exposure).

[Insert Table 1 about here]

Outcome Measurement

The MINI was administered to all participants at pre-treatment, and at 3-month follow-up for the treatment group (assessors were not blinded to treatment condition at 3-month-follow-up). Self-report outcome measures were administered at pre-treatment (prior to lesson 1), before lesson 4, at post-treatment (one week after the treatment group finished the program), and 3-month follow-up (treatment group only), or at matched time points for WLC. The treatment group completed the K-10 before they commenced each lesson, as a measure to alert the clinician if participants' scores rose by more than 0.5SD between lessons, indicating a significant increase in distress.

[1] To view a demonstration of the lesson content of a similar program (the Depression course), visit: https://thiswayupclinic.org.au/demo/all

Clinical Contact with Clinician and Therapist

The treatment group participants received regular email and/or phone contact with their Clinician (KM, the practice manager) until they completed lesson 2, after which contact was made in response to patient request or if the clinician initiated contact because of a deterioration in the K-10, or PHQ-9 score. KM was supervised by a Therapist (JN, a PhD Clinical Psychologist). All emails requiring clinical advice were responded to by the Therapist. If clinically indicated, or if patients' K-10 and/or PHQ-9 scores deteriorated, the Therapist would make telephone contact with the participant. The Clinician, Therapist, and Research Support Officer (AM) participated in weekly meetings to assess and discuss participants' progress.

Statistical Analyses

Significance testing of group differences regarding demographic data and pre-treatment measurements was conducted using independent samples *t*-tests, and $\chi 2$ where the variables consisted of nominal data. Intent-to-treat (ITT) marginal model analyses using the restricted maximum likelihood (REML) method were used to account for missing data due to participant drop-outs. This approach is appropriate for RCTs with multiple time points (Salim *et al.* 2008), and does not assume that the last measurement was stable (an assumption of the the last observation carried forward approach, Gueorguieva and Krystal 2004). As the primary outcomes measures (GAD-7, PHQ-9, K-10) were also collected at mid-point, effects for the primary measures were modelled with an autoregressive (AR1) covariance structure to account for the correlation between the time-points. Effects for the secondary measures were modelled using an unstructured (UN) covariance structure. Model fit was evaluated using Schwarz's Bayesian Criterion (BIC). Significant effects were followed up with pairwise contrasts comparing pre-treatment to post-treatment scores. Analyses were performed in

SPSS version 20. We applied Hedges *g* adjustment to calculate the between-group effect sizes, and adjusted for the correlations between repeated measurement time-points when calculating the within-group pre- to post-treatment effect sizes.

Results

Baseline

The mean age of participants was 44.3 years (*SD*=12.2, range=21-80), and 77 were female (77.8% of the sample). The majority were married or living in a de-facto relationship (*n*=62, 62.6%), were educated with a postgraduate degree (*n*=58, 58.6%) and were in full-time (*n*=36, 36.4%) or part-time paid work (*n*=23, 23.2%). Participants reported moderate levels of depression and anxiety on the PHQ-9 and GAD-7 at the start of the program (see Table 2 for sample characteristics).

[Insert Table 2 about here]

Diagnostic Status according to MINI Interviews

In the treatment group, 19 had GAD/MDD (41.3%), 21 had GAD (45.7%) (but with sub-threshold MDD), and 6 met criteria for MDD (13.0%) and had sub-threshold GAD. In the WLC group, 28 met criteria for co-morbid GAD/MDD (52.8%), 16 (30.2%) had GAD, and 9 (17.0%) had MDD, but each had sub-threshold GAD or MDD respectively. Diagnostic status did not differ between groups at baseline ($\chi^2(2)=2.52$, $p>.05$).

Baseline Between-Group Comparisons

There were no differences between the groups on age, pre-treatment BDI-II, GAD-7, NEO, PHQ-9, K-10, nor PSWQ scores (*ps*>.05). However, the control group reported significantly higher disability on the WHODAS-II *(t*(97)=2.35, *p*=.02). Chi-square analyses

demonstrated that there were no between-group differences in any other demographic characteristics such as gender, marital status, educational status, nor employment status (see Tables 2 and 3).

Adherence Results

Forty-nine individuals were randomised into the treatment group. Of these, 46 completed pre-treatment questionnaires and were eligible for analysis, and 41/46 completed the total six lessons (89% adherence). Post-treatment and 3-month follow-up data was collected on 43/46 and 40/46 participants, respectively. Of the participants who were considered drop-outs, one completed one lesson only, one completed two lessons, two completed three lessons, and one completed four lessons. Sixty participants were randomly allocated to the WLC group. Of these, 54 completed pre-treatment questionnaires. One of these withdrew to seek alternative treatment, leaving 53 eligible for the analysis. Post-treatment data was collected on 53/53 participants (see Figure 1 for further details).

Primary Outcome Measures and Effect Sizes

Marginal models with group as a fixed factor and time as a repeated factor were conducted separately for each of the dependent measures (see Table 3 for results). All main effects for PHQ-9, GAD-7, and K-10 scores were qualified by significant Group x Time interactions, all F's (df's 1, 181.84-185.61)>7.91, all p's$\leq$.001. Between-group comparisons on the PHQ-9, GAD-7 and K-10 scores revealed that post-treatment scores were significantly lower in the Treatment group relative to WLC, with large observed effect sizes (.85-1.4). Within-group comparisons for the Treatment group revealed large effect sizes. The reductions in the Control group on these measures were not significant.

[Insert Table 3 about here]

Secondary Outcome Measures and Effect Sizes

All main effects for BDI-II, PSWQ, WHODAS-II and NEO scores were qualified by significant Time x Group interactions, all *F*'s (df's 1, 91.53-93.56) >5.17, all *p*'s≤.001(WHODAS-II, *p*=.02)[2].Between-group comparisons revealed that all post-treatment scores were significantly lower in the Treatment group relative to the Control group, with medium (.56, PSWQ) to large (1.13, BDI-II) between-subjects effect sizes being found. This corresponded to large within-group effects in the Treatment group. The reductions in the Control group were not significant, with the exception of WHODAS-II scores (Effect size=.30).

Primary Outcome Measures and Effect Sizes at 3-Month Follow-Up

Marginal models with group as a fixed factor and time as a repeated factor were conducted separately to compare mean reductions in scores from post-treatment to 3-month follow-up for the treatment group. Baseline scores were entered as a covariate for their respective analysis. For GAD-7 scores (*r*=.67), the main effect of Time was significant, *F*(1,37.36)=4.39,*p*=.04. The reduction corresponded to a small effect size (.26). For PHQ-9 scores (*r*=.64) the main effect of Time was not significant, *F* (1,36.56)=3.04, *p*=.09. For K-10 (*r*=.61) the main effect of Time was significant, *F* (1,36.06)=5.89,*p*=.02. The reduction corresponded to a small effect size (.33) (see Table 3 for means and standard deviations).

Clinical significance

Three criteria for clinical significance were employed: first, pre- and post-treatment PHQ-9 and GAD-7 scores were compared with cut-off scores (≥10) to provide an index of remission. For the treatment group, 26/46 (56.5%) versus 11 (23.9%) met criteria on the GAD-7 at pre-treatment and post-treatment respectively, and for the PHQ-9, 27 (58.7%) at pre-treatment reduced to 8 (17.4%) above cut-off at post-treatment. For the WLC group,

[2] Given there were significant pre-treatment differences on the WHODAS-II score, we also conducted a separate analysis evaluating the between-group differences at post-treatment on the WHODAS-II, controlling for baseline WHODAS-II scores. This difference remained significant, *F*(91)=18.24,*p* < .001.

27/53 (50.9%) met criteria on the GAD-7 at pre-treatment, compared to 24 (45.3%) at post-assessment. Thirty-nine (73.6%) were above cut-off on the PHQ-9 at pre-treatment, versus 30 (56.6%) post-assessment. The proportion of participants who were in the clinically depressed and anxious range differed according to group at post-treatment (PHQ-9: $\chi2(1)=16.01, p<.001$; GAD-7: $\chi2(1)=4.92, p<.05$) but not at pre-treatment (PHQ-9: $\chi2(1)=2.46, p>.05$; GAD7:$\chi2(1)=.38, p>.05$).

Second, following Jacobson and Truax (1991), we calculated reliable change index (RCI) values for the PHQ-9 and GAD-7 scores. RCI values were calculated using test-retest reliability values of .83 and .84 from Spitzer *et al.* (2006) for GAD-7 scores, and Kroenke *et al.* (2001) for PHQ-9 scores respectively. In order to calculate standard error of measurement values, standard deviations were derived from current sample (GAD-7 pre-treatment pooled SD = 4.44; PHQ-9 pre-treatment pooled SD = 4.43). For PHQ-9 scores, a change score (from pre-post treatment) greater than 4.91 was considered reliable change, and for the GAD-7, a change score greater than 5.07 was considered reliable change. Of the treatment group, 21/46 (45.73%) showed reliable improvements on the PHQ-9 compared to 13/53 (24.5%) in the WLC group ($\chi2\ (2,93)=10.00, p<.001$). For the GAD-7, 18/46 (39.1%) treatment group participants reliably improved, compared to 5/53 (9.4%) in the WLC group ($\chi2\ (2,93)=18.13, p<.001$). Notably, 5 (9.4%) and 6 (11.3%) WLC participants reported deteriorated PHQ-9 and GAD-7 scores respectively at post-treatment.

Third, we compared the proportion of participants with principal diagnoses at pre-treatment and compared this with follow-up diagnoses for the treatment group. At 3-month follow-up, 32/46 participants (69.6%) no longer met criteria for either disorder. 7 participants

were diagnosed with GAD at 3-month follow-up. Of these, 4 had initially been diagnosed with co-morbid MDD/GAD.[3]

Time spent per participant and patient satisfaction

The clinician and therapist combined spent on average 23.37 minutes per participant (*SD*=12.15, range=7-60 minutes) on email and telephone contact in the treatment group over the course of the program. Treatment group participants were asked to provide a rating ranging from 1 to 10 (where 10=high level of agreement) about how logical the program was, their confidence that the program was successful at teaching them techniques for managing symptoms, and their confidence in recommending the program to a friend with similar concerns. The results were combined to derive a total treatment satisfaction score, which was high on average (*M*=25.64, *SD*=3.58, where the highest possible combined total score was 30).

Discussion

This study compared a 6-lesson clinician-assisted iCBT program for mixed depression and anxiety to a WLC group. Adherence was high (89%), and the iCBT program was more efficacious than WLC on all primary and secondary measures of depression, generalised anxiety and functional impairment. Between 40-45 per cent of participants in the treatment group showed reliable improvements immediately following treatment. Importantly, gains were maintained at 3-month follow-up for the treatment group, with evidence of further improvements (albeit small effects) in GAD symptoms and general distress between post-treatment and follow-up. Approximately 70% of participants no longer met diagnostic criteria on structured interview at 3 month follow-up. The impact of our iCBT program on other comorbid anxiety disorders (e.g., Panic Disorder and Social Phobia) is unknown and awaits

[3] The remaining 7 individuals (15.2%) were unable to be contacted for phone interview.

further investigation. Nevertheless, these results support growing evidence for the efficacy of iCBT for mixed anxiety and depression (Andersson *et al.* 2011), and for the primary disorders (Perini *et al.* 2009). Future comparisons with single-disorder treatments will test whether this transdiagnostic iCBT program provides added benefits in terms of efficacy, acceptability, cost-effectiveness and efficiency.

The non-blinded structured interviews at 3-month follow-up, reliance on self-report measures, and short follow-up period are some limitations of Study 1. In addition, the use of a wait-list control comparison group (instead of an active treatment control condition) is a limitation of this study. Of note, approximately 10 per cent of the WLC group deteriorated whilst waiting for treatment. Active treatment controls could be used in future to minimise the possibility of deterioration occurring, and would offer the advantage of ruling out the role of non-specific therapeutic factors accounting for symptom improvement. Another issue is that it is unknown whether these results would generalise to community patient samples outside of a strictly controlled research setting. In order to address this question, we conducted an effectiveness trial of the Worry and Sadness Program in Study 2.

Study 2

Adherence and effectiveness of the *Worry and Sadness Program* in a primary care setting

Method and Procedure

We aimed to test the effectiveness of the Worry and Sadness Program in www.crufadclinic.org by making it available to the 1800 clinicians registered with CRUfADclinic from August 8, 2011 to 15 December 2011. Registered clinicians identify patients that are likely to benefit from the iCBT courses and prescribe a course to their

patient, with a prescription that tells the patient how to enrol, and provides a secure passcode linking the patient to the clinician (participants termed "*prescription patients*").

The iCBT course in CRUfADclinic was the same as in Study 1, with the following exceptions: patients had 30 days to complete the first two lessons, and were allowed 90 days to complete the entire course. Once patients completed each lesson and downloaded the homework, they were required to book in a date on which they would commence the following lesson. Reminder emails were sent if patients missed the date. Patients completed the PHQ-9, GAD-7, and WHODAS-II prior to commencing lesson 1, and prior to commencing lesson 6. The K-10 was also administered prior to each lesson. Automatic emails were sent to a patients' supervising clinician to report lesson-by-lesson progress on the K-10. Clinicians were also alerted via email if their patients' K-10 scores rose over 30 (severe range), rose by more than 0.5SD, or if patients missed their nominated lesson date.

Results

Participants

There were 136 prescription patients prescribed the *Worry and Sadness Program* between 8 August, 2011 and the 15 December 2011. Patients' mean age was 39.27 (*SD*=13.05, range=18-78), 88 were female (64.7%), and 75 (55.1%) were living in rural Australia. Prescribing clinicians were: general practitioners (*n*=59, 43.4%), 39 (28.7%) were psychologists, 23 were medical specialists (16.9%), 6 were nurses (4.4%), and 9 were other allied health specialists (6.6%). Participants' depression (*M*=14.14, *SD*=6.33) and anxiety (*M*=12.19, *SD*=5.43) symptoms were on average in the moderate range.

Primary Outcome Measures, Effect Sizes, and Adherence

Prescription patients completed on average, 4.21 lessons (*SD*=1.87, range = 1-6). Out of 136 participants who began the course, 56 completed all six lessons (41.2% adherence). Completers of the program were highly satisfied. Of the non-completers 17 (12.5%) completed one lesson only, 17 (12.5%) completed two lessons, 15 (11.0%) completed three lessons, 14 (10.3%) completed four lessons, and 17 (12.5%) completed five lessons.

Intent-to-treat (ITT) marginal model analyses with Time as the repeated variable were conducted separately for each of the dependent measures. All main effects of time for each of the dependent variables were significant, *Fs*(1, 57.14-69.49) >50.47, all *p*s< .001. Due to the high level of attrition, estimated marginal means and standards errors are reported here (see Table 4 for observed means and standard deviations). For GAD-7 scores (pre-treatment: *M*=12.14, *SE*=.47; post-treatment: *M*=6.90, *SE*=.60, *r*=.61), the pairwise comparison was significant, *t*(1,65.35)=8.94,*p*<.001, corresponding to a large effect (1.05, 95%CI=.67-1.42). For PHQ-9 scores (pre-treatment: *M*=14.04, *SE*=.56; post-treatment: *M*=8.32, *SE*=.78, *r*=.57) the pairwise comparison was significant, *t*(1,62.19)=7.61,*p*<.001, with a large corresponding effect size (.94, 95%CI=.56-1.31). For K-10 scores (pre-treatment: *M*=30.41, *SE*=.67; post-treatment: *M*=21.94, *SE*= .99, *r*=.62) the pairwise comparison was significant, *t* (1, 60.67) = 9.21, *p* < .001, with a large corresponding effect size (1.07, 95%CI = .69-1.44). For WHODAS-II scores (pre-treatment: *M*=16.20, *SE*=.85; post-treatment: *M*=11.26, *SE*=.99, *r*=.66) the pairwise comparison was significant, *t*(1,57.14)=7.10, *p*<.001. The reduction corresponded to a medium effect size (.78, 95%CI = .40-1.15).

[Insert Table 4 about here]

Completers versus non-completers

Adherence to the *Worry and Sadness Program* by prescription patients was low (41.2%). Therefore, we compared completers versus non-completers to indirectly investigate possible contributing factors for drop-out. Independent samples *t*-tests showed there were no significant differences between completers versus non-completers of the program in terms of age, pre-treatment K-10, PHQ-9, GAD-7, and WHODAS-II scores (see Table 4). Chi-square analyses showed no differences in the type of clinicians who prescribed the course to completers versus non-completers ($\chi^2(5)=5.49$, $p>.05$). These variables were also entered along with rurality as predictors in a multivariate logistic regression model predicting completion status (0=completed<6 lessons, 1=completed 6 lessons). No significant predictors emerged, all *p*'s>.05. Independent samples *t*-tests compared the completers and non-completers on lesson-by-lesson K-10 scores. There were no differences, suggesting that patients did not drop-out because they were not getting benefit from the course (see Table 4).

Using a conservative SD score of 7.5, we measured the number of people who had experienced benefit of at least one SD on the K-10 prior to drop-out (this analysis only included individuals who completed ≥ 2 lessons) (*n*=63). Twenty-four (30%) of the non-completers had experienced greater than 7.5 points reduction on the K-10 prior to drop-out, suggesting they had experienced benefit from the iCBT program prior to drop-out.

Discussion

Study 2 was the first investigation of the effectiveness of our iCBT course for mixed anxiety and depression in patients in primary care. Patients who completed the course were highly satisfied with the program when supervised by their primary care clinician. Although we do not know whether patients in this study registered for the iCBT course as adjuncts to additional treatments (e.g., psychological therapy and/or medications), patients reported similar reductions in depression and anxiety symptoms, distress and disability as in the RCT

in Study 1, suggesting that the efficacy of this iCBT program is generalisable to community settings.

Adherence to the program was much lower for patients supervised in primary care (41% versus 89% in the RCT [Study 1]). This result mirrors the pattern of poorer adherence to psychological treatments in primary care settings (Cuijpers *et al.* 2009). The completion rates in Study 2 are slightly lower than median completion rates (56%) reported in a meta-analysis of computerised CBT treatments (Waller and Gilbody 2009), and are lower than program adherence for our iCBT programs for depression (54%) (Williams and Andrews 2012) and GAD (55%) in primary care (Mewton *et al.* 2012). However, interestingly, Mason & Andrews (2012) recently found better adherence to the Worry and Sadness Program in a sample of patients who were referred by primary care practitioners to be monitored and supervised over the web by Clinical Psychologists and Psychiatrist at CRUfAD (60%, Mason & Andrews, 2012). CRUfAD's clinicians are highly familiar with the program and closely monitor and supervise iCBT patients following a standard protocol[4]. It is possible that due to the higher level of complexity (and broader scope) of the transdiagnostic program, patients need a greater level of monitoring, frequent reminders, and feedback from primary care clinicians to continue with the program. Future work is needed to improve techniques primary care clinicians use to encourage adherence to iCBT courses in community settings (Hilvert-Bruce *et al.* 2012), particularly for individuals with complex and comorbid emotional disorders.

[4] CRUfAD's present policy in supervising referral patients is to contact the patient after the first two lessons to encourage progress, and following this, contact if the patient initiated it, or contact if the patients' K-10 scores deteriorate, and/or if patients miss their nominated lesson date. Although we recommend to primary care clinicians using our programs that contact by them improves adherence, it is unclear to what extent clinicians adhere to our recommendations.

It is noteworthy that we found that 30% of the drop-outs of the program had experienced significant benefit (on the K-10); this suggests people may drop-out from iCBT in community settings after experiencing some benefit. It is also possible that participants who experienced benefit may have dropped out because they attributed their improvements to other concurrent treatments (e.g., psychotherapy or medications), thereby influencing their decision to discontinue the program. A limitation of our study was that we did not assess participants' use of concurrent treatments. Closer monitoring of the use of concurrent treatments whilst taking part in iCBT will shed light on this possibility. Finally, other limitations of Study 2 were the reliance on self-report measures and short follow-up period. Without a control group in Study 2, we also cannot rule out the possibility that symptom reductions can be accounted for by other factors (e.g., spontaneous recovery).

General Discussion

In two studies, we investigated the efficacy of the Worry and Sadness Program - an internet-delivered CBT program for mixed anxiety and depression - in both a research RCT setting and in primary care. A key strength of this study was the measurement of both the effectiveness and efficacy of the same iCBT program across different settings. Overall, results showed that iCBT for depression and anxiety was effective compared to wait-list control, and the efficacy of this program was generalisable to patients in who completed the program in primary care, a "real world setting" supervised by busy practitioners (effect sizes>1.0). Not surprisingly, adherence was better in the RCT, and relatively low (41%) in primary care. This finding may be attributable to observed sample differences between the studies (e.g., lower depression severity, higher mean age and greater proportion of females in the RCT). Individuals who volunteer to participate in RCTs may also be more motivated to adhere and engage in treatment, more receptive to psychotherapeutic interventions and more willing to try psychological techniques suggested in the program, than those who are

prescribed the program by their primary care practitioner (who may not have even been interested in receiving psychological treatment). Participants in the RCT are also are more rigorously assessed at the outset, and are closely supervised compared to those who referred to treatment from their primary care practitioner, potentially minimising drop-out. In summary, our findings suggest that we have an effective and accessible iCBT program that reduces symptoms of co-morbid depression and anxiety. However, adherence differs dramatically depending on clinician guidance and treatment setting (research versus primary care). In future, greater emphasis needs to be placed on changes in the delivery and design of the program, and education of primary care practitioners to support patients throughout their prescribed program in an attempt to improve adherence rates so that we can maximise benefit of iCBT in the "real-world."

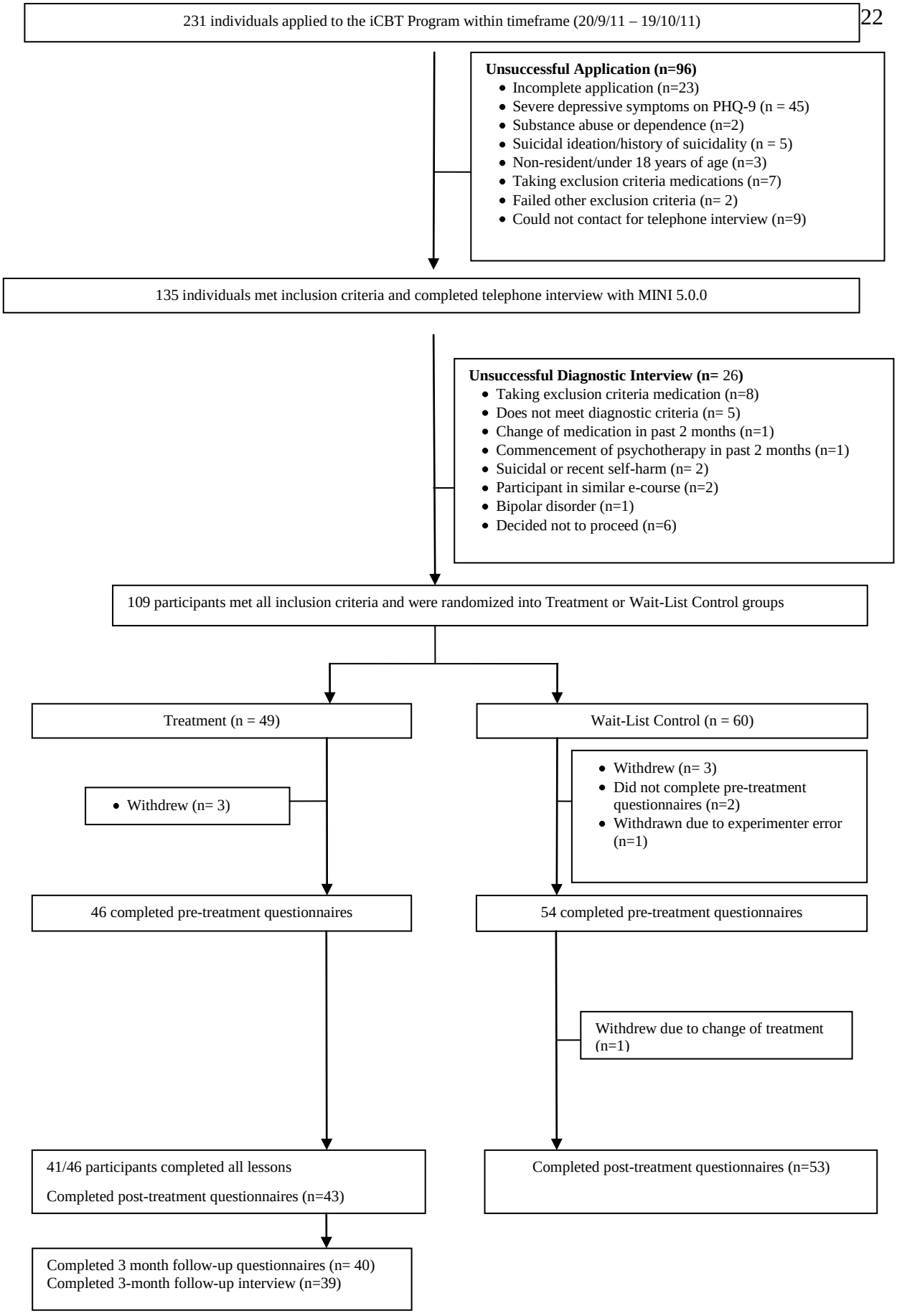

Figure 1. Participant Flow Diagram for Study 1 (RCT)

Table 1. Lesson Content of the Worry and Sadness Program

Lesson Number	Content	Homework Tasks
1	Psychoeducation about anxiety and depression, the fight or flight response, controlled breathing, and physical exercise	Controlled breathing, physical exercise
2	Cognitive therapy components: education about the cognitive model, cognitive distortions, and introduction to thought monitoring; activity planning	Thought monitoring, activity planning
3	Thought challenging/cognitive restructuring; challenging positive and negative meta-cognitive beliefs about repetitive thinking; shifting attention, hunt for positives	Thought challenging, hunt for positives
4	Education about avoidance and safety behaviours; graded exposure and structured problem solving	Graded exposure and structured problem solving
5	Advanced graded exposure (imaginal exposure, interoceptive exposure); troubleshooting difficulties with graded exposure	Graded exposure
6	Relapse prevention	Relapse prevention plan

Table 2 Baseline Demographics and Sample Characteristics for the Treatment and Wait-List Control Groups

	Treatment Group *n*=46		WLC Group *n*=53		Statistic
Age (years) [*M/SD*]	43.6	12.44	44.9	12.11	$t(97) = .52$, $p = .61$
Gender					$\chi2\ (1, 93) = .01$, $p = .91$
Male	10	21.7	12	22.6	
Female	36	78.3	41	77.4	
Marital Status					$\chi2\ (7, 93) = 6.65$, $p = .47$
Single/Never Married	8	17.4	12	22.6	
Married/De-Facto	33	71.7	29	54.7	
Separated/Divorced/Widowed	2	4.4	8	15.1	
No answer provided	3	6.6	0	0	
Educational Status					$\chi2\ (4, 93) = .51$, $p = .97$
High School	5	10.9	7	13.2	
Tertiary (Undergraduate)	24	52.2	26	49.1	
Tertiary (Postgraduate)	3	6.5	5	9.4	
Other Certificate	11	23.9	11	20.1	
No answer provided	3	6.6	0	0	
Employment Status					$\chi2\ (6, 93) = 8.78$, $p = .19$
At home parent	3	6.5	6	11.3	
Full-time paid work	22	47.9	14	26.4	
Part-time paid work	11	23.9	13	24.5	
Unemployed	1	2.2	5	9.4	
Student	3	6.6	2	3.8	
Retired	1	2.2	7	13.2	
Disabled	2	4.3	2	3.8	
No answer provided	3	6.6	0	0	
Previous Mental Health Treatment	36	78.3	41	77.4	$\chi2\ (1, 93) = .04$, $p = .98$
Age at First Consultation for Mental Health	32.75	13.09	30.63	11.67	$t(75) = .75$, $p = .46$
Hours of Internet Use [*M/SD*]	14.2	17.19	12.4	12.16	$t(90) = .58$, $p = .56$
Confidence Using Internet					$\chi2\ (5, 93) = 2.12$, $p = .83$
Not Confident	0	0	1	1.9	
Average	2	4.3	5	9.4	
Confident	14	30.4	15	28.3	
Mildly Confident	3	6.5	4	7.5	
Very Confident	24	52.2	24	45.3	
Current Medications	19	41.3	21	39.6	$\chi2\ (1\ 93) = .37$, $p= .54$
Current Psychological Treatment	2	4.3	4	7.5	$\chi2\ (1, 93) = .44$, $p= .51$

Note. Except where noted, values refer to number and percentage scores. Educational Status = highest level of education received. *M* = mean, *SD* = standard deviation. WLC = Wait-List Control.

Table 3 Observed means and standard deviations for primary and secondary outcome measures and within and between-group effect sizes following treatment in Study 1

Measure	Group	Pre-Treatment		Post-Treatment		3-month follow-up		Pre-treatment between-group comparisons	Post-treatment between-group comparisons	Between-group effect sizes Hedges *g* (95% CI)	Pre-to post-treatment within-group comparisons	*r*	Within-group effect sizes, Cohen's *d* (95% CI)
		M	*SD*	*M*	*SD*	*M*	*SD*	*t* (df)	*F* (df)		*t* (df)		
PHQ-9	Treatment	10.39	3.90	5.76	4.24	4.05	3.79	$t(158.34) = 1.31$, $p = .18$	$F(1, 166.32) = 26.51$, $p < .001$	1.00 (.59-1.40)	$t(229.84) = 7.05$, $p < .001$	.52	1.05 (.62-1.47)
PHQ-9	WLC	11.62	4.80	10.41	4.88	-	-				$t(228.66) = 1.94$, $p = .15$	.61	.26 (-.12-.64)
GAD7	Treatment	10.37	3.74	5.93	4.28	4.39	3.71	$t(178.02) = .06$, $p = .94$	$F(1, 187.16) = 19.36$, $p < .001$	.85 (.43-1.27)	$t(240.55) = 5.95$, $p = .001$	.44	.96 (.53-1.38)
GAD7	WLC	10.43	5.00	9.92	4.90	-	-				$t(239.31) = .72$, $p = .99$	.70	.07 (-.31-.45)
K-10	Treatment	25.43	5.14	18.78	5.74	15.46	7.59	$t(183.17) = .84$, $p = .40$	$F(1, 192.15) = 44.52$, $p < .001$	1.40 (.99-.1.80)	$t(243.181) = 6.18$, $p < .001$	.45	.98 (.55-1.40)
K-10	WLC	26.51	6.30	27.51	6.64	-	-				$t(241.99) = 1.03$, $p = .90$	.62	.12 (-.26-.50)
BDI-II	Treatment	21.24	6.98	10.48	8.30			$t(140.76) = .67$, $p = .50$	$F(1, 147.93) = 48.02$, $p < .001$	1.13 (.72-1.53)	$t(96.00) = 10.21$, $p < .001$	.26	1.89 (1.46-2.31)
BDI-II	WLC	22.41	9.17	21.24	10.56						$t(91.94) = 1.09$, $p = .27$	.73	.11 (-.27-.49)
PSWQ	Treatment	64.22	8.67	57.00	10.98			$t(130.65) = .52$, $p = .60$	$F(1, 137.46) = 10.79$, $p = .001$	.56 (.15-.96)	$t(95.13) = 6.39$, $p < .001$	.60	.87 (.44-1.29)

PSWQ	WLC	63.11	11.77	62.96	9.99				$t(91.80) = .13$, $p = .89$	.75	.01 (-.30-.39)
WHODAS-II	Treatment	24.35	6.38	20.17	6.47	$t(144.51) = .65$, $p = .51$	F (1, 132.95) = 17.00, $p < .001$	.76 (.34-1.18)	$t(93.42) = 5.58$, $p < .001$	.66	.70 (.27-1.12)
WHODAS-II	WLC	27.89	8.27	25.66	7.82				$t(90.34) = 2.93$, $p = .004$	.72	.30 (-.08-.68)
NEO	Treatment	31.28	5.38	26.17	7.44	$t(144.51) = .65$, $p = .51$	F (1, 151.71) = 27.44, $p < .001$	.80 (.39-1.20)	$t(94.96) = 6.51$, $p < .001$	.45	1.04 (.61-1.46)
NEO	WLC	32.11	6.03	32.09	7.07				$t(90.57) = .02$, $p = .98$	.62	.00 (-.38-.38)

Note. PHQ-9 = The Patient Health Questionnaire-9, GAD-7 = Generalised Anxiety Disorder 7-item scale, WHODAS-II = The 12 item World Health Organisation Disability Assessment Schedule, PSWQ = The Penn State Worry Questionnaire, BDI-II = Beck Depression Inventory – Second Edition, NEO = NEO-Five Factor Inventory – Neuroticism Subscale. Treatment = treatment group, WLC = wait-list control group, *M* = Mean, *SD* = Standard Deviation, CI = Confidence Interval. *r* = inter-correlation between pre- and post- treatment scores used to calculate within-group effect sizes.

Table 4 Observed means and standard deviations for primary outcome measures for total sample, completer and non-completer samples in Study 2

Measure	Lesson Number	Total Sample			Within-subjects effect size: Cohen's *d* (95% CI)	Completers		Non-completers		Between-group comparison (completer versus non-completer)
		N	*M*	*SD*		*M*	*SD*	*M*	*SD*	Statistic
K-10	Lesson 1	136	30.57	7.90		30.05	8.14	30.93	7.75	t (134) = .63, p > .05
	Lesson 2	119	26.98	7.71		26.82	7.95	27.13	7.55	t (117) = .21, p > .05
	Lesson 3	106	24.11	7.95		24.09	8.48	24.14	7.40	t (104) = .03, p > .05
	Lesson 4	88	23.86	8.51		23.66	7.71	24.22	8.26	t (73) = .29, p > .05
	Lesson 5	75	22.60	8.26		22.18	8.46	23.84	7.71	t (56) = .75, p > .05
	Lesson 6	57	21.81	8.56	1.07 (.69-1.44)	21.81	8.56	n/a	n/a	-
WHODAS-II	Lesson 1	136	16.23	9.43		16.80	8.32	15.83	9.54	t (134) = .59, p > .05
	Lesson 6	57	11.77	9.40	.78 (.40-1.15)	11.77	9.40	n/a	n/a	
GAD-7	Lesson 1	135	12.19	5.43		11.79	5.80	12.48	5.18	t (134) = .72, p > .05
	Lesson 6	57	6.71	5.15	1.05 (.67-1.42)	6.71	5.15	n/a	n/a	
PHQ-9	Lesson 1	135	14.14	6.33		14.23	6.55	14.08	6.20	t (134) = .14, p > .05
	Lesson 6	57	8.43	6.50	94 (56-1.31	8.43	6.50	n/a	n/a	
Patient Satisfaction	Lesson 6	57	27.30	5.69						

Note. K-10 = Kessler-10 item psychological distress scale, WHODAS-II = World Health Organisation Disability Assessment Schedule, GAD-7 = Generalised Anxiety Disorder 7-item scale, PHQ-9 = Patient Health Questionnaire 9-item scale. Completers: n=57, non-completers: n =79. *M* = Mean, SD = Standard Deviation.

Acknowledgements

We wish to express our gratitude to Dr Matthew Boulton, Dr Alison Mahoney, and Dr Elizabeth Mason for assistance with the design of the treatment protocol. We also wish to thank the three anonymous reviewers for their comments on the original version of the manuscript. Alishia D. Williams is supported by a National Health and Medical Research Council of Australia Fellowship (630746).

Declaration of Interest

None.

References

Andersson, G., Cuijpers, P. (2009). Internet-Based and Other Computerized Psychological Treatments for Adult Depression: A Meta-Analysis. *Cognitive Behaviour Therapy* 38(4): 196-205.

Andersson, G., Estling, F., Jakobsson, E., Cuijpers, P., Carlbring, P. (2011). Can the patient decide which modules to endorse? An open trial of tailored Internet treatment of anxiety disorders. *Cognitive Behaviour Therapy* 40(1): 57-64.

Andrews, G. (1990). Classification of neurotic disorders. *Journal of the Royal Society of Medicine* 83: 606-607.

Andrews, G., Cuijpers, P., Craske, M. G., McEvoy, P., Titov, N. (2010). Computer Therapy for the Anxiety and Depressive Disorders Is Effective, Acceptable and Practical Health Care: A Meta-Analysis. *PLoS ONE* 5(10): e13196.

Andrews, G., Henderson, S., Hall, W. (2001). Prevalence, comorbidity, disability and service utilisation. Overview of the Australian National Mental Health Survey. *British Journal of Psychiatry* 178: 145-153.

Andrews, G., Kemp, A., Sunderland, M., Von Korff, M., Ustun, T. B. (2009). Normative data for the 12 item WHO Disability Assessment Schedule 2.0. *PLoS ONE* 4(12): e8343.

Barlow, D. H., Allen, L. B., Choate, M. L. (2004). Toward a unified treatment for emotional disorders. *Behavior Therapy* 35(2): 205-230.

Beck, A. T., Steer, R. A., Ball, R., Ranieri, W. F. (1996). Comparison of Beck Depression Inventories-IA and -II in Psychiatric Outpatients. *Journal of Personality Assessment* 67(3): 588.

Beck, A. T., Steer, R. A., Brown, G. K. (1996). Beck Depression Inventory (2nd Edition)-Manual. San Antonio, TX, The Psychological Corporation.

Brown, T. A., Antony, M. M., Barlow, D. H. (1992). Psychometric properties of the Penn state worry questionnaire in a clinical anxiety disorders sample. *Behaviour Research and Therapy* 30(1): 33-37.

Brown, T. A., Campbell, L. A., Lehman, C. L., Grisham, J. R., Mancill, R. B. (2001). Current and lifetime comorbidity of the DSM-IV anxiety and mood disorders in a large clinical sample. *Journal of Abnormal Psychology* 110(4): 585-599.

Chorpita, B. F., Daleiden, E. L. (2009). Mapping evidence-based treatments for children and adolescents: application of the distillation and matching model to 615 treatments from 322 randomized trials. *Journal of Consulting and Clinical Psychology* 77(3): 566-579.

Chu, B. C. (2012). Translating Transdiagnostic Approaches to Children and Adolescents. *Cognitive and Behavioral Practice* 19(1): 1-4.

Costa, P. T. J., McCrae, R. R. (1985). The NEO personality inventory manual. Odessa, FL, Psychological Assessment Resources.

Cuijpers, P., van Straten, A., Andersson, G., van Oppen, P. (2008). Psychotherapy for depression in adults: A meta-analysis of comparative outcome studies. *Journal of Consulting and Clinical Psychology* 76(6): 909-922.

Cuijpers, P., van Straten, A., Donker, M. (2005). Personality traits of patients with mood and anxiety disorders. *Psychiatry Research* 133(2–3): 229-237.

Cuijpers, P., van Straten, A., van Schaik, A., Andersson, G. (2009). Psychological treatment of depression in primary care: a meta-analysis. *The British journal of general practice : the journal of the Royal College of General Practitioners* 59(559): e51-60.

Dear, B. F., Titov, N., Schwencke, G., Andrews, G., Johnston, L., Craske, M. G., McEvoy, P. (2011). An open trial of a brief transdiagnostic internet treatment for anxiety and depression. *Behaviour Research and Therapy* 49(12): 830-837.

Furukawa, T., Kessler, R., Slade, T., Andrews, G. (2003). The performance of the K6 and K10 screening scales for psychological distress in the Australian National Survey of Mental Health and Well-Being. *Psychological Medicine: A Journal of Research in Psychiatry and the Allied Sciences* 33(2): 357-362.

Gueorguieva, R., Krystal, J. (2004). Move over ANOVA: progress in analyzing repeated-measures data and its reflection in papers published in the Archives of General Psychiatry. *Archives of General Psychiatry* 61(3): 310-317.

Hedman, E., Andersson, E., Ljótsson, B., Andersson, G., Rück, C., Lindefors, N. (2011). Cost-effectiveness of Internet-based cognitive behavior therapy vs. cognitive behavioral group therapy for social anxiety disorder: Results from a randomized controlled trial. *Behaviour Research and Therapy* 49(11): 729-736.

Hettema, J. M., Neale, M. C., Myers, J. M., Prescott, C. A., Kendler, K. S. (2006). A Population-Based Twin Study of the Relationship Between Neuroticism and Internalizing Disorders *American Journal of Psychiatry* 163(5): 857-864.

Hilvert-Bruce, Z., Rossouw, P. J., Wong, N., Sunderland, M., Andrews, G. (2012). Adherence as a determinant of effectiveness of internet cognitive behavioural therapy for anxiety and depressive disorders. *Behaviour Research and Therapy* 50(7–8): 463-468.

Hollandare, F., Andersson, G., Engstrom, I. (2010). A comparison of psychometric properties between internet and paper versions of two depression instruments (BDI-II and MADRS-S) administered to clinic patients. *Journal of medical internet research* 12(5): e49.

Jacobson, N. S., Truax, P. (1991). Clinical significance: a statistical approach to defining meaningul change in psychotherapy research. *Journal of Consulting and Clinical Psychology* 59(1): 12-19.

Kessler, R., Andrews, G., Colpe, L., Hiripi, E., Mroczek, D., Normand, S., Walters, E., Zaslavsky, A. (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine: A Journal of Research in Psychiatry and the Allied Sciences* 32(6): 959-976.

Kessler, R. C., DuPont, R. L., Berglund, P., Wittchen, H.-U. (1999). Impairment in Pure and Comorbid Generalized Anxiety Disroder and Major Depression at 12 Months in Two National Surveys. *American Journal of Psychiatry* 156(12): 1915-1923.

Kessler, R. C., Ustun, T. B. (2004). The World Mental Health (WMH) Survey Initiative Version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatry Research* 13(2): 93-121.

Kroenke, K., Spitzer, R., Williams, J. (2001). The PHQ-9: Valdity of a Brief Depression Severity Measure. *Journal of General Internal Medicine* 16(9): 606-613.

Kroenke, K., Spitzer, R. L., Williams, J. B. W., Monahan, P. O., Löwe, B. (2007). Anxiety Disorders in Primary Care: Prevalence, Impairment, Comorbidity, and Detection. *Annals of Internal Medicine* 146(5): 317-325.

Mason, E., Andrews, G. (2012). Automated and specialist assessments lead to the same outcomes in internet-based CBT for anxiety and depression: The computer will be with you shortly. *Manuscript Under Review.*

McEvoy, P. M., Nathan, P., Norton, P. J. (2009). Efficacy of Transdiagnostic Treatments: A Review of Published Outcome Studies and Future Research Directions. *Journal of Cognitive Psychotherapy* 23(1): 20-33.

Mewton, L., Wong, N., Andrews, G. (2012). The effectiveness of internet cognitive behavioural therapy for generalized anxiety disorder in clinical practice. *Depression and Anxiety* 29(10): 843-849.

Meyer, T. J., Miller, M. L., Metzger, R. L., Borkovec, T. D. (1990). Development and validation of the penn state worry questionnaire. *Behaviour Research and Therapy* 28(6): 487-495.

Newman, M. G., Przeworski, A., Fisher, A. J., Borkovec, T. D. (2010). Diagnostic comorbidity in adults with generalized anxiety disorder: impact of comorbidity on psychotherapy outcome and impact of psychotherapy on comorbid diagnoses. *Behavior Therapy* 41(1): 59-72.

Nolen-Hoeksema, S., Wisco, B. E., Lyubomirsky, S. (2008). Rethinking rumination. *Perspectives on Psychological Science* 3(5): 400-424.

Norton, P. J. (2012). A Randomized Clinical Trial of Transdiagnostic Cognitve-Behavioral Treatments for Anxiety Disorder by Comparison to Relaxation Training. *Behavior Therapy* 43(3): 506–517.

Perini, S., Titov, N., Andrews, G. (2009). Clinician-Assisted Internet-Based Treatment is Effective for Depression: Randomized Controlled Trial. *Australian and New Zealand Journal of Psychiatry* 43(6): 571-578.

Robinson, E., Titov, N., Andrews, G., McIntyre, K., Schwencke, G., Solley, K. (2010). Internet treatment for generalized anxiety disorder: A randomized controlled trial comparing clinician vs. technician assistance. *PLoS ONE* 5(6).

Salim, A., Mackinnon, A., Christensen, H., Griffiths, K. (2008). Comparison of data analysis strategies for intent-to-treat analysis in pre-test–post-test designs with substantial dropout rates. *Psychiatry Research* 160(3): 335-345.

Schulz, K. F., Altman, D. G., Moher, D., for the, C. G. (2010). CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomised Trials. *PLoS Med* 7(3): e1000251.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry* 59(Suppl 20): 22-33.

Spitzer, R. L., Kroenke, K., Williams, J. B., Lowe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine* 166(10): 1092-1097.

Stewart, R. E., Chambless, D. L. (2009). Cognitive–behavioral therapy for adult anxiety disorders in clinical practice: A meta-analysis of effectiveness studies. *Journal of Consulting and Clinical Psychology* 77(4): 595-606.

Titov, N., Andrews, G., Johnston, L., Robinson, E., Spence, J. (2010). Transdiagnostic Internet treatment for anxiety disorders: A randomized controlled trial. *Behaviour Research and Therapy* 48(9): 890-899.

Titov, N., Dear, B. F., Schwencke, G., Andrews, G., Johnston, L., Craske, M. G., McEvoy, P. (2011). Transdiagnostic internet treatment for anxiety and depression: A randomised controlled trial. *Behaviour Research and Therapy* 49(8): 441-452.

van Balkom, A. J., van Boeijen, C. A., Boeke, A. J., van Oppen, P., Kempe, P. T., van Dyck, R. (2008). Comorbid depression, but not comorbid anxiety disorders, predicts poor outcome in anxiety disorders. *Depression and Anxiety* 25(5): 408-415.

Waller, R., Gilbody, S. (2009). Barriers to the uptake of computerized cognitive behavioural therapy: A systematic review of the quantitative and qualitative evidence. *Psychological Medicine* 39(5): 705-712.

Weisz, J. R., Chorpita, B. F., Palinkas, L. A., Schoenwald, S. K., Miranda, J., Bearman, S. K., Daleiden, E. L., Ugueto, A. M., Ho, A., Martin, J., Gray, J., Alleyne, A., Langer, D. A., Southam-Gerow, M. A., Gibbons, R. D. (2012). Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth: a randomized effectiveness trial. *Archives of General Psychiatry* 69(3): 274-282.

Williams, A. D., Andrews, G. (2012). The effectiveness of internet cognitive behavioural therapy (iCBT) for depression in primary care: A quality assurance study. *Manuscript under review*.

Wittkampf, K. A., Naeije, L., Schene, A. H., Huyser, J., van Weert, H. C. (2007). Diagnostic accuracy of the mood module of the Patient Health Questionnaire: a systematic review. *General Hospital Psychiatry* 29(5): 388-395.

World Health Organization. World Health Organization Disability Assessment Schedule II (WHO-DAS II).(http://www.who.int/icidh/whodas/).