REEACT-2: A large scale pragmatic randomised trial of telephone-supported computerised Cognitive Behaviour Therapy

Simon Gilbody[1][*] Sally Brabyn[1], Karina Lovell[8], David Kessler[7], Thomas Devlin[1], Lucy Smith[1], Ricardo Araya[2], Michael Barkham[3], Peter Bower[4], Cindy Cooper[5, 11], Sarah Knowles[4], Elizabeth Littlewood[1], David A Richards[9], Debbie Tallon[10], David White[11], Gillian Worthy[12], on behalf of the REEACT collaborative

[1]*Department of Health Sciences, University of York, UK*

[2] *Centre of Global Mental Health, London School of Hygiene and Tropical Medicine, UK*

[3]*Centre for Psychological Services Research, University of Sheffield, UK*

[4] *NIHR School for Primary Care Research, University of Manchester, UK*

[5]*School of Health and Related Research, University of Sheffield, UK*

[6]*Centre for Health Economics, University of York, UK*

[7]*Academic Unit of Primary Health Care, University of Bristol, UK*

[8]*School of Nursing, Midwifery and Social Work, University of Manchester, UK*

[9]*University of Exeter Medical School, University of Exeter, UK*

[10]*School of Social and Community Medicine, University of Bristol, UK*

[11]*Clinical Trials Research Unit, University of Sheffield, UK*

[12]*York Trials Unit, University of York, UK*

*Corresponding author (simon.gilbody@york.ac.uk)

Abstract

Background

Computerised cognitive behaviour therapy (cCBT) for depression has the potential to be efficient therapy but engagement is poor in primary care trials.

Aim

We tested the benefits of adding telephone-support to cCBT.

Methods

We compared telephone-facilitated cCBT (MoodGYM) (n=187) to minimally-supported cCBT (MoodGYM) (n=182) in a pragmatic randomised trial (RCT). Outcomes were depression severity (PHQ9), anxiety (GAD7), and somatoform complaints (PHQ15) at 4 & 12 months.

Results

cCBT use increased by a factor of between 1.5 and with telephone-facilitation. At four months PHQ-9 scores were 1.9 points lower (95% CI 0.5 to 3.3) for telephone supported cCBT. At 12 months the results were no longer statistically significant (0.9 PHQ9 points; 95%CI -0.5 to 2.3). There was improvement in anxiety scores and for somatic complaints.

Discussion

Telephone facilitation of cCBT improves engagement and expedites depression improvement. The effect was small to moderate and comparable with other low intensity psychological interventions.

Word count = 150

Introduction

Depression is the most common mental health disorder in community settings and is estimated to become the second largest cause of global disability by 2020.[1] It is one of the most common reasons for consulting with a primary care physician (PCP) and its associated personal and economic burden is considerable.[2]

Whilst antidepressants remain an important treatment option, many patients and healthcare professionals would like to access psychological therapy as an alternative or adjunct to drug therapy.[3] Cognitive behaviour therapy (CBT) has emerged as a leading evidence-supported form of brief psychological therapy for people with depression.[4] However, demand for CBT cannot be met from existing therapist resources.[5] One promising alternative to therapist-delivered CBT is the use of self-help interventions including the provision of therapy via computer.[6] In recent years a number of interactive programmes have been developed which enable CBT to be delivered by computer (computerised CBT or cCBT). If effective, such programmes have the potential to expand the provision of psychological therapy in primary care and may represent an efficient and effective form of care for depression.[7]

In an earlier large scale pragmatic trial (the first REEACT trial)[8, 9] we compared two commonly used cCBT packages (MoodGYM or Beating the Blues) versus usual primary care under real world conditions to test the effectiveness (rather than efficacy) in a pragmatic trial. Participants were proactively offered technical support, and weekly encouragement to use the computer packages, but we purposely did not augment the content of psychological therapy over the telephone. The cCBT in the first REEACT trial was therefore a form of supported self-help, but was not one which was guided by a clinician. The first REEACT trial is at the time of writing the largest publicly funded independently-conducted primary care trial of cCBT. The main finding of the REEACT trial was that for the primary outcome of depression severity at four months there was no significant benefit when participants were offered technically-supported cCBT in addition to usual GP care. The most likely explanatory mechanism of lack of effect was poor uptake and use of computer packages by trial participants under real world conditions.[9]

Systematic reviews have highlighted the potential for cCBT to be effective but have also further demonstrated variable effect sizes and substantial between-study heterogeneity.[10, 11] One important source of between-study heterogeneity is the level of support that is made available to people who are offered treatment with cCBT. Computerised CBT requires a person with depression to engage with a self-help computer-based technology. Research has shown that people with depression often do not engage with cCBT, and only a minority actually complete all of the planned

4

sessions of the computer package.[12] This observation is consistent with a broader body of research into the uptake and effectiveness across the range of self-help interventions for depression such as bibliotherapy (self-treatment using written materials).[13] Research in the area of self-help treatments for depression has demonstrated that entirely self-guided materials (with no professional support) are likely to be less effective than self-help technologies where there is a level of guidance and professional support ('guided self-help'). Unsupported self-help treatment (including unsupported computer-delivered self-help) has been shown in systematic reviews to have minimal or relatively small effect sizes.[13] In contrast, more intensively- and professionally-supported treatments have generally been found in efficacy trials to have moderate effect sizes claimed to be comparable to those achieved with face-to-face therapy.[14] To our knowledge the comparative effectiveness of minimally-supported cCBT versus more intensively supported cCBT has not been directly tested in large-scale, independently-conducted, head-to-head, pragmatic trials (though there are some head to head comparisons in smaller-scale trials [15]).

We postulated on the basis of these findings, and on the basis of emerging trial-based evidence summarised in systematic reviews (e.g. [10]) that people with depression might engage with cCBT and it might show an effect, but only if offered alongside a high level of facilitation and support. We designed the present study (the REEACT 2 trial) to test this hypothesis and to generate trial-based evidence on the best means of delivering cCBT in primary care mental health services.

Methods

Study design and patients

The second Randomised Evaluation of Effectiveness and Acceptability of Computer-delivered Therapy (REEACT 2) trial was designed to examine the additional benefits of telephone-facilitation and structured guidance alongside a free to use computer-delivered CBT package (MoodGYM). The comparator was minimally-supported cCBT.

Participants in both arms were given access to a free-to-use cCBT programme (MoodGYM), an accompanying booklet, a Freephone number for technical support and continued with usual GP care. MoodGYM (©ANU http://moodgym.anu.edu.au) is a free-to-use, internet-based, interactive CBT programme for depression developed and copyrighted at the Australian National University Centre for Mental Health Research. The online programme is accompanied by a booklet with exercises and quizzes and consists of five interactive modules released sequentially and lasting approximately 30-45 minutes and a sixth session that is predominantly consolidation and revision. Study participants were asked to complete one session each week. The programme provides patients with CBT techniques to overcome patterns of unhelpful thinking using cartoon characters to represent habits of thought.

Experimental arm: Participants in the telephone-facilitated cCBT (experimental) arm were allocated a telephone support worker (TSW) who provided a programme of weekly telephone calls. The background of TSWs was that of a graduate-level support worker. The telephone facilitation programme comprised eight telephone calls to be completed alongside the cCBT programme within 14 weeks of first contact from the telephone support worker (and before the four-month follow-up time point). The purpose of the first and longest session (30-40 minutes) was to introduce the participant to the principles of CBT and the MoodGYM programme and booklet, explain the process and help the participant identify difficulties and goals, and feel confident about engaging with the intervention. The following six sessions were between 10 and 20 minutes long and were intended to provide motivation and to help participants identify any barriers to engagement with cCBT and to the achievement of their goal(s). The final session helped participants to consolidate what they had learned from cCBT and discuss their next steps and, if appropriate, how they might use the MoodGYM programme in the future. The telephone facilitation programme was delivered according to a manual developed by co-investigator KL in conjunction with the REEACT-2 team. TSWs received one day of training in the delivery of the intervention. Clinical supervision was given to trial telephone support workers by investigators KL, DK and SG.

Comparator arm: All participants in the control group were registered as users of MoodGYM and given a unique password. As with the intervention group, they were supplied with a free helpline number to ring if they had technical problems or needed advice and a booklet explaining MoodGYM, but they did not receive regular phone calls. This comparator intervention replicated United Kingdom National Health Services (NHS) care in most settings and represented what would happen if a patient were given the website of a cCBT package such as MoodGYM by their primary care physician or primary care mental health worker without being offered pro-active support.

The study population comprised patients selected from primary care with depression or low mood as determined by a score of ten or more on the Patient Health Questionnaire (PHQ-9).[16] This cut-off point is known to detect clinical depression (major depression) in a primary care populations[17] with acceptable sensitivity and specificity. The REEACT 2 participants were recruited from a mix of rural and urban UK primary care practices in and around Bristol, Avon, Somerset, Gloucestershire, Manchester, Sheffield, Derbyshire, South Yorkshire, Humberside, East Yorkshire, Durham, Tyneside and Northumberland.

Participants meeting the following criteria were eligible to enter the study:

• Aged 18 or above

• Not currently in receipt of cCBT or specialist psychological therapy

• Score of ≥10 overall (indicating moderate, moderately severe or severe depression) and <3 for question 9 (measuring suicidal thoughts) on the PHQ-9 depression instrument.[16]

Both incident and prevalent primary care cases of depression were included. In line with the pragmatic nature of this trial, patients were eligible to participate whether or not they were in receipt of antidepressant medication or had co-morbid physical illness or non-psychotic functional disorders. We excluded people currently in receipt of psychological therapy.

We also excluded potential participants who:

• were actively suicidal as identified by the Primary Care Physician or as reported by item 9 on the PHQ-9

• had been bereaved within the last year

• had given birth within the last year

• had a diagnosis of psychotic depression

- had a primary diagnosis of alcohol or drug abuse

- were not able to read and write in English

Randomisation and masking

Simple randomisation was performed using a computer-generated random number sequence. At the end of the baseline appointment study researchers telephoned a secure randomisation line at the York Trials Unit and were given participant allocation and MoodGYM log-in details. Participants were informed immediately'

Outcome measures

Primary outcome measure: the pre-specified primary outcome was depression severity and symptomatology as measured on a validated self-report continuous measure (the Patient Health Questionnaire-9 (PHQ9))[16] at four months.

Secondary outcome measures: PHQ-9 at 12 months (as a continuous measure); PHQ-9 at 4 and 12 months (dichotomous measure at cut-point PHQ9 >= 10);[16] anxiety (GAD-7);[18] somatoform complaints (PHQ-15);[19] health-state utility (EuroQol - EQ5D);[20] service use using the adapted Client Service Receipt Inventory (CSRI)[21] at four and 12 months.

Sample size calculation: The REEACT-2 trial was powered on the basis of an ability to detect a between group difference in PHQ scores. We sought to recruit 350 patients with depression - 175 participants per arm. The REEACT-2 trial was designed to have sufficient power to detect a Cohen's d effect size of 0.30 with 80% power allowing for loss to follow up of 20% in line with our empirically-based estimates from the REEACT trial. The final sample size for the two arms was 369 and we exceeded this pre-specified sample size.

Statistical analysis

All outcomes were summarised descriptively by intervention group and at each time point using mean, median, standard deviation (SD), range and number of patients for continuous outcomes and number of patients and percentage for discrete outcomes. The primary outcome was the severity of depression as measured by the Patient Health Questionnaire 9 (PHQ-9) as a continuous measure at four months. Statistical analyses were performed in SAS version 9.3.

PHQ-9 as a continuous outcome: The PHQ-9 score was summarised and analysed as a continuous outcome. This was summarised for each assessment time point (baseline, four and 12 months) using mean, SD, median and range, and the number of missing values. Plots were presented showing the mean and 95% CI at each time point. A repeated measures mixed regression model was used to analyse the change in PHQ-9 score over time. This included all randomised participants (intention to treat) and provides reliable estimates assuming the data are missing at random (MAR). The outcome was the PHQ-9 score at four and 12 months and the model included the baseline PHQ-9 score, treatment group, age, gender, baseline GAD-7 score and time. The treatment x time interaction was included to evaluate if the difference between treatments changed over time. The mean difference, 95% CI and p-values are presented for all terms in the model. Effect sizes (Cohen's d) were calculated for the between group differences in mean PHQ-9 score at four and 12 months using the difference between the means and corresponding standard errors from the mixed model. The standard errors were converted to standard deviations using the corresponding sample size in each treatment group.

PHQ-9 as a dichotomous outcome: The dichotomous analysis (not depressed (PHQ-9 < 10)/depressed (PHQ 9 ≥ 10)) compared minimally-supported cCBT with telephone-facilitated cCBT using a logistic regression model adjusting for the baseline PHQ-9 score, age, gender, baseline GAD-7 score and treatment. The dichotomous analysis was on a complete case basis (only including those with a four-month assessment). A sensitivity analysis was performed using simple imputation and a worst case scenario. This assumed that all participants with a missing outcome were still depressed with a PHQ-9 score ≥ 10.

Other secondary outcomes: GAD-7 and the PHQ-15 scores were analysed as continuous outcomes using the same repeated measures mixed models described for PHQ-9 above.

Resource use data and health state utilities: (derived from the EQ5D) formed the basis of a full economic evaluation and are described in the full study report.[22]

Adherence: Adherence by participants to the computer programme was measured by requesting information from the website providing MoodGYM (hosted by the developers of MoodGYM at the Australian National University – ANU). We obtained computer usage data on the number of times each participant logged on to the MoodGYM programme and whether each module was 25, 50, 75 or 100 per cent complete.

Adverse events: were classified according to their seriousness and relationship to the intervention.

Results

A total of 369 participants were randomised to the two-armed comparison of minimally-supported cCBT with telephone-facilitated cCBT, n=182 and n=187 respectively. The first participant was randomised on the 24th June 2011 and the last on the 25th April 2013. The flow of participants through the trial is shown in the CONSORT diagram (Figure 1).

The two groups were well balanced at baseline for gender, age, ethnicity and education. The mean age of participants was 40.6 years (sd13.8). The study population was mostly white British (94%) and 64.5% were female. The minimally-supported cCBT and telephone-facilitated cCBT groups were balanced at baseline for employment. The majority (61.5%) of participants were employed and of these 23.6% were absent from work by reason of depression at the time of their baseline assessment (Table 1). The severity of depression at baseline (as ascertained by the median PHQ9 score) was 16 (range 10-25) which corresponds with a moderate to high level of severity.

<Figure 1 about here>

<Table 1 about here>

PHQ-9 as a continuous outcome: At the four month primary outcome the between-group difference in PHQ-9 scores was 1.9 points (95% CI 0.5 to 3.3) in favour of telephone-facilitated cCBT, with a standardised effect size (Cohen's d) of 0.32 (p=0.009). At 12 months there was no longer evidence of a between-group difference in PHQ-9 scores (0.9 95% CI -0.5 to 2.3). Using a repeated measures analysis over the whole trial period the between-group difference in PHQ-9 scores was 1.4 (95% CI 0.2 to 2.6) in favour of telephone-facilitated cCBT with a standardised effect size (Cohen's d) of 0.27 (p=0.0253) (see figure 2 and table 2 & 3).

PHQ-9 as a dichotomous outcome: After four months 66 (50.30%) of the 128 participants in the minimally-supported cCBT group and 51 (36.2%) of the 141 in the telephone-facilitated cCBT group had a PHQ-9 score greater than or equal to 10. The odds of no longer being depressed (defined as PHQ-9 < 10) at four months were increased twofold in the telephone-facilitated cCBT group compared to minimally supported cCBT (odds ratio [OR] 2.05 95% confidence interval [CI] 1.23 to 3.42). The benefit of telephone-facilitated cCBT was no longer significant at 12 months ([OR 1.63 95% CI 0.98 to 2.71 p=0.06). See table 2.

<Figure 2 about here>

<Table 2 & 3 about here>

Other secondary outcomes: For secondary outcomes there was a significant between group difference in anxiety scores (GAD-7) in favour of telephone-facilitated cCBT when all time points were considered (between group difference 1.2, 95% CI 0.1 to 2.3; p=0.037) (see figure 3 and table 4). For somatic complaints there was a borderline significant difference in favour of telephone–facilitated cCBT when all time points were considered (between group difference 1.1, 95% CI 0.0 to 1.8; p=0.051) (see figure 4 and table 5).

<Figure 3 & 4 about here>

<Table 4 & 5 about here>

Adherence: When computer records were scrutinised there were few participants who completed all five sessions in either minimally-supported (10.4%) or telephone-facilitated cCBT (19.4%). Usage was generally increased by a factor of between 1.5 and two when telephone-facilitation was offered, with 46.2% of participants in receipt of telephone facilitation completing two or more sessions versus 29.1% of participants with minimal support. See figure 5 and table 6 for a detailed summary of cCBT programme usage.

<Figure 5 and table 6 about here>

Adverse events: There were a total of ten serious adverse events, none of which was thought to be related to the trial. All were reviewed by the Trial Steering Committee (TSC) and the Data Monitoring and Ethics Committee (DMEC).

Discussion

REEACT 2 is one of the largest trials of computer delivered CBT to date. The trial tested whether the addition of structured telephone facilitation substantially increased engagement with computer-mediated CBT and resulted in improved outcomes. We purposely designed a pragmatic trial to test effectiveness under real world conditions rather than efficacy under ideal but restrictive conditions in order to maximise the external validity of our results.[23] The main finding of the REEACT 2 trial is that the addition of structured telephone facilitation resulted in significant reductions in depression severity compared to cCBT with technical support alone. The effect size was moderate and was most evident in the short term (4 months) and had diminished by 12 months. Telephone facilitation of cCBT therefor expedited depression improvement, though the absence of benefit at 12 months is unsurprising given the average duration of an episode of depression is less than 12 months. When depression was considered as a binary outcome, the odds of no longer being depressed were twice as high in the telephone facilitated group at 4 months. Benefits in terms of psychological outcomes were also observed using a validated anxiety scale and for somatoform complaints when outcomes were averaged over a 12 month period. Engagement with the technology was increased through the addition of telephone facilitation.

The REEACT 2 trial drew upon a manualised form of telephone support, which can readily be delivered after a relatively brief period of training. At present computerised CBT is offered by many healthcare systems as a minimally-supported low-intensity psychological intervention and as part of a stepped care framework. The intervention trialed in the REEACT 2 study therefore represents an enhancement of care that can be readily delivered at scale in primary care settings. The results of the REEACT 2 trial should be considered alongside other trials and systematic reviews of cCBT and low intensity interventions for common mental disorders. Our earlier study (the REEACT trial)[8] was similarly a large scale pragmatic trial of cCBT where one arm included the free to use cCBT package (MoodGYM). In this previous trial we offered a low intensity form of technical telephone support and found that usage was low and there were no additional clinical benefits of cCBT when it was added to usual primary care. This led us to speculate that an enhancement in the level of support and guidance might increase uptake and effectiveness. Evidence that the addition of guidance to cCBT is associated with a greater level of effectiveness comes from systematic reviews, where

pooled estimates of the effect size of trials with therapist guidance are larger than the pooled effect size obtained from unsupported cCBT.[10] Evidence also comes from a systematic review of small-scale head to head comparisons of unsupported versus supported cCBT in a range of common mental disorders.[15] This hypothesis has now been directly tested in the present randomised controlled trial which, to our knowledge, is the first test of this in a large scale (adequately-powered) direct randomised head-to-head comparison. The results of the REEACT 2 trial are also comparable to other primary care based psychological treatments,[24] but the effect size observed in REEACT 2 is smaller compared to other developer-led trials of cCBT.[10] The additional benefit of guided support is in line with the results of a systematic review of three small scale studies in depression.[15]

The REEACT 2 trial has several strengths in its design. First the REEACT 2 trial was pragmatic in design and recruited from primary care, whereas most trials to date have recruited from online populations or by participant advertisement. This addresses a major shortcoming of the literature identified by Andersson and Cuijpers in their 2009 review.[10] The results of REEACT 2 are therefore more generalisable to clinical populations encountered in primary care. Second, the trial was significantly larger than other trials to date (see Baumeister et al[15]) and had sufficient power to detect more modest effect sizes. Third, we conducted a pragmatic trial of effectiveness rather than efficacy by trialling a low-intensity enhancement to cCBT that could be delivered at distance to a range of people fulfilling very broad depression inclusion criteria (typical of those encountered in primary care). Fourth, the period of follow up was one year and this allows some conclusions to be drawn about the durability of effect. Finally we were able to study the actual use of computer technology in our trial participants with reference to computer records.

There were limitations to the REEACT 2 study. First, in view of the pragmatic nature of the design there was loss to follow up of around one quarter of the participants overall, and we know very little about the outcomes of these participants. Second, we did not measure outcome with a clinical interview to establish the presence of depression according to accepted classification systems. Instead we rely on self-report measures of depression severity; though these are well-validated against diagnostic systems.[25] Third, even with the provision of telephone facilitation, only a small proportion of participants in either arm completed all sessions of the cCBT programme. There is possibly more still that can be offered to enhance the uptake of computer therapy. Finally the level of depression severity at baseline was moderate to high, and this is at the upper range of severity recommended in some stepped care systems.[4] However the positive results of the REEACT 2 trial provide supportive evidence that low intensity interventions can be offered to this group and will

results in improved outcomes. This finding is consistent with recent reviews of the effectiveness of low intensity interventions across the range of severities of depression.[26]

The implications for practice and policy that emerge from the first REEACT and REEACT 2 trials are twofold. The first is that minimally supported cCBT results in very low levels of uptake and confers little over usual care. We would therefore suggest that healthcare systems do not offer this form of unsupported treatment as part of stepped care. However unsupported cCBT should still be offered as a form of direct access treatment to non-clinical populations, though the benefits that might be expected are likely to be small. The second implication is that the addition of structured telephone facilitation (such as that designed in REEACT 2 to work alongside MoodGYM) will result in greater levels of engagement with computer technology. In turn this will produce moderate clinical improvements and reductions in the proportion of people who continue to experience depression over a 4 to 12 month period. Telephone support is a low intensity enhancement of care that can be offered at scale and could be readily implemented in most healthcare settings as part of a stepped care system.

References

1.   Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. The Lancet. 2013; **382**(9904): 1575-86.

2.   Layard R. The case for psychological treatment centres. BMJ. 2006; **332**: 1030-2.

3.   Roth A, Fonagy P. What works for whom?: A critical review of psychotherapy research: The Guilford Press; 2005.

4.   National Institute for Health and Clinical Excellence. Depression in Adults: The Treatment and Management of Depression in Adults (Update - NICE clinical guideline 90). Manchester: National Institute for Health and Clinical Excellence; 2009.

5.   Kaltenthaler E, Brazier J, De Nigris E, Tumur I, Ferriter M, Beverley C, et al. Computerised cognitive behaviour therapy for depression and anxiety update: a systematic review and economic evaluation. Health Technol Assess. 2006; **10**: 33.

6.   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**(10): e13196.

7.   Christensen H, Griffiths KM, Jorm AF. Delivering interventions for depression by using the internet: randomised controlled trial. BMJ. 2004; **328**: 265.

8.   Gilbody S, Littlewood E, Hewitt C, Brierley G, Tharmanathan P, Araya R, et al. Computerised cognitive behaviour therapy (cCBT) as treatment for depression in primary care (REEACT trial): large scale pragmatic randomised controlled trial. BMJ. 2015; **351**: h5627.

9.   Littlewood E, Duarte A, Hewitt C, Knowles S, Palmer S, Walker S, et al. A randomised controlled trial of computerised cognitive behaviour therapy for the treatment of depression in primary care:The Randomised Evaluation of the Effectiveness and Acceptability of Computerised Therapy (REEACT) trial. Health Technology Assessment. 2015; **Volume 19**(Number 101).

10.   Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: A meta-analysis. Cog Behav Ther. 2009; **38**(4): 196-205.

11.   Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. Psychol Med. 2007; **37**(3): 319-28.

12.   Waller R, Gilbody S. Barriers to the uptake of computerised cognitive behavioural therapy: A systematic review of the quantitative and qualitative evidence. Psychol Med. 2008; **39**(5): 705-12.

13.   Cuijpers P, Donker T, Johansson R, Mohr DC, van Straten A, Andersson G. Self-guided psychological treatment for depressive symptoms: a meta-analysis. PloS ONE. 2011; **6**(6): e21274.

14.   Andersson G, Cuijpers P, Carlbring P, Riper H, Hedman E. Guided Internet-based vs. face to face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis. World Psychiatry. 2014; **13**(3): 288-95.

15.   Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on Internet-based mental health interventions—A systematic review. Internet Interventions. 2014; **1**(4): 205-15.

16.   Kroenke K, Spitzer RL. The PHQ-9: A new depression and diagnostic severity measure. Psychiatr Ann. 2002; **32**: 509-21.

17.   Moriarty AS, Gilbody S, McMillan D, Manea L. Screening and case finding for major depressive disorder using the Patient Health Questionnaire (PHQ-9): a meta-analysis. General Hospital Psychiatry. 2015; **37**(6): 567-76.

18.   Spitzer RL, Kroenke K, Williams JBW, Lowe B. A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. Arch Intern Med. 2006; **166**(10): 1092-7.

19.   Kroenke K, Spitzer RL, Williams JB. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. Psychosom Med. 2002; **64**: 258–66.

20.   The EuroQol Group. EuroQOL - a new facility for the measurement of health related quality of life. Health Policy. 1990; **16**: 199-208.

21.   Chisholm D, Knapp MRJ, Knudsen HC, Amaddeo F, Gaite L, Van Wijngaarden B. Client Socio-Demographic and Service Receipt Inventory-European Version: development of an instrument for international research EPSILON Study 5. The British Journal of Psychiatry. 2000; **177**(39): s28-s33.

22.   Brabyn S, Araya R, Barkham M, Bower P, Cooper C, Duarte A, et al. The second randomised evaluation of the effectiveness, cost-effectiveness and acceptability of computerised therapy trial (REEACT-2) – Does the provision of telephone support enhance the effectiveness of computer-delivered Cognitive Behaviour Therapy? Health Technol Assess. 2016; **In Press**.

23.   Thorpe KE, Zwarenstein M, Oxman AD, Treweek S, Furberg CD, Altman DG, et al. A pragmatic–explanatory continuum indicator summary (PRECIS): a tool to help trial designers. J Clin Epidemiol. 2009; **62**(5): 464-75.

24.   Cuijpers P, van Straten A, van Schaik A, Andersson G. Psychological treatment of depression in primary care: a meta-analysis. Brit J Gen Pract. 2009; **59**: e51-e60.

25.   Manea L, Gilbody S, McMillan D. Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. Can Med Assoc J. 2012; **184**(3): E191-E6.

26.   Bower P, Kontopantelis E, Sutton A, Kendrick T, Richards DA, Gilbody S, et al. Influence of initial severity of depression on effectiveness of low intensity interventions: meta-analysis of individual patient data. BMJ: British Medical Journal. 2012; **346**: f540-f.

Figure 1: Consolidated standards of reporting trials (CONSORT) diagram

Returned Consent to Contact
n=602

Excluded n=233

Ineligible (n=151); Declined to participate (n=52);
Unable to be contacted (n=30);

Randomised n= 369

Minimally-supported CCBT n=182

Telephone-facilitated cCBT n=187

4-month follow up n=128 (70.3%)

Did not return questionnaire (n=42);

Withdrawn from study (n=5)

Withdrawn from follow-up (n=7)

4-month follow up n=141 (75.4%)

Did not return questionnaire (n=36);

Withdrawn from study (n=6)

Withdrawn from follow-up (n=4)

12 months follow up n=132 (72.5%)

Did not return questionnaire (n=36);
Withdrawn from study (n=5)
Withdrawn from follow-up (n=9)

12 months follow up n=142 (75.9%)

Did not return questionnaire (n=33)
Withdrawn from study (n=6)
Withdrawn from follow-up (n=6)

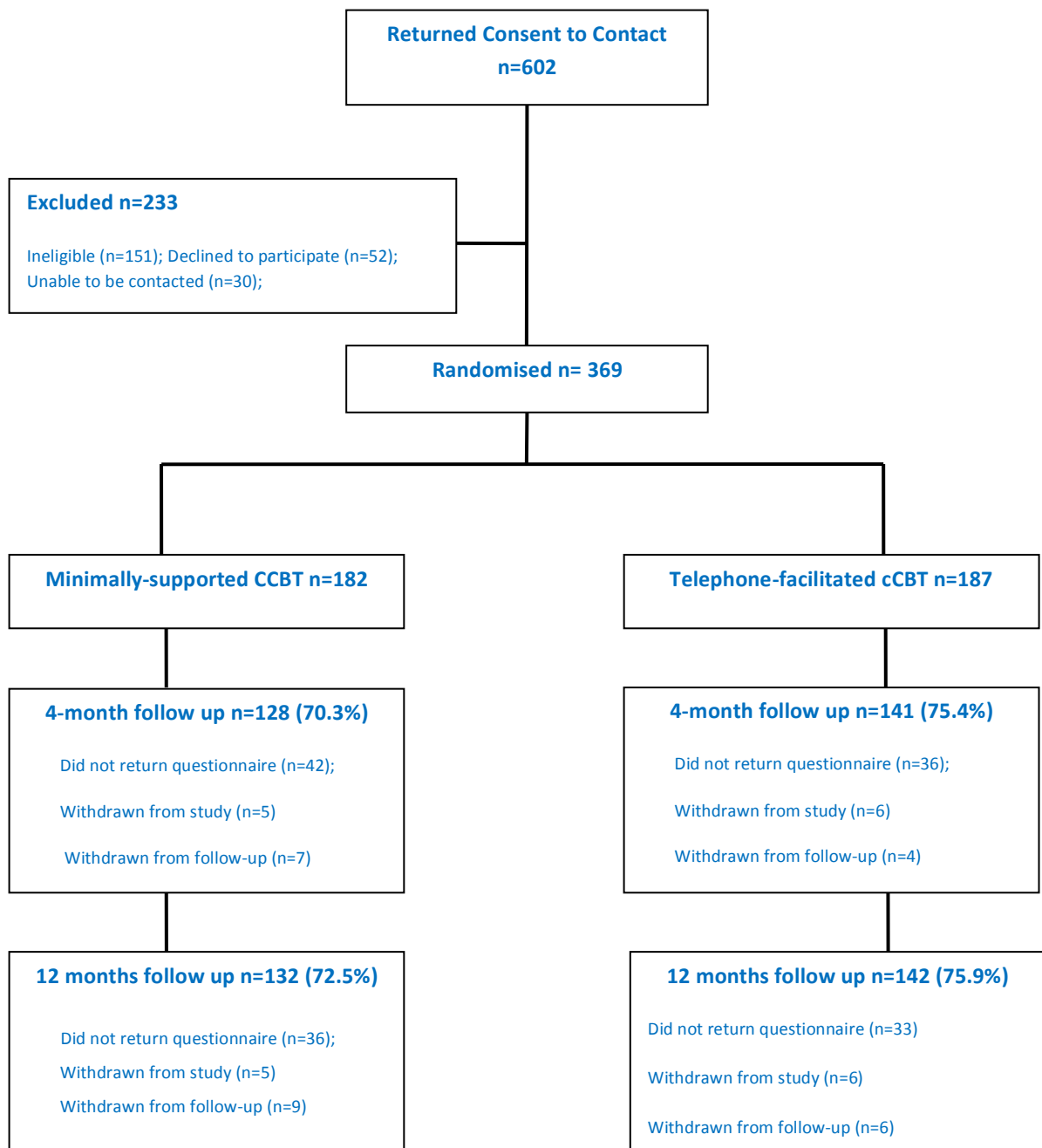**Figure 2: Depression severity (mean and 95% CI PHQ-9 score*) at each assessment**

*Results from a repeated measures, mixed model adjusting for baseline score, age, gender, baseline GAD-7 score and time

Figure 3: Anxiety severity (mean and 95% CI GAD7 score*) at each assessment

*Results from a repeated measures, mixed model adjusting for baseline score, age, gender, and time

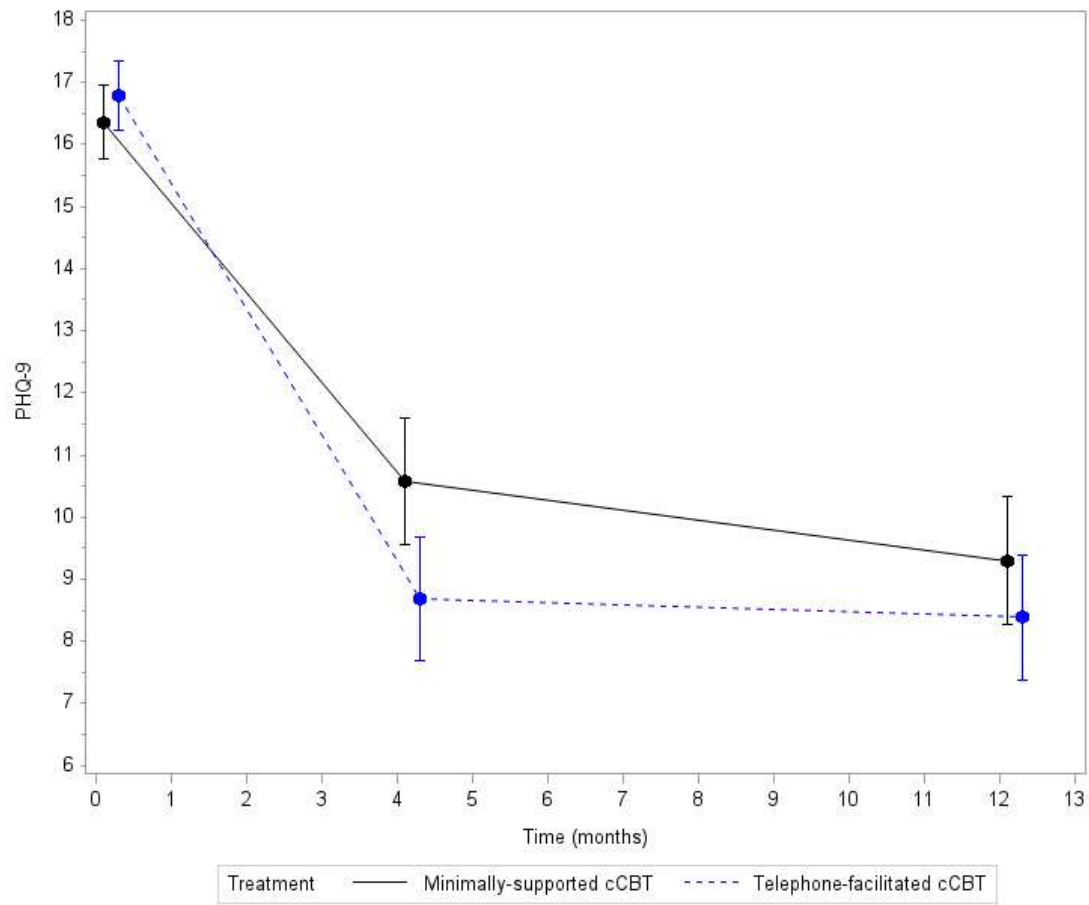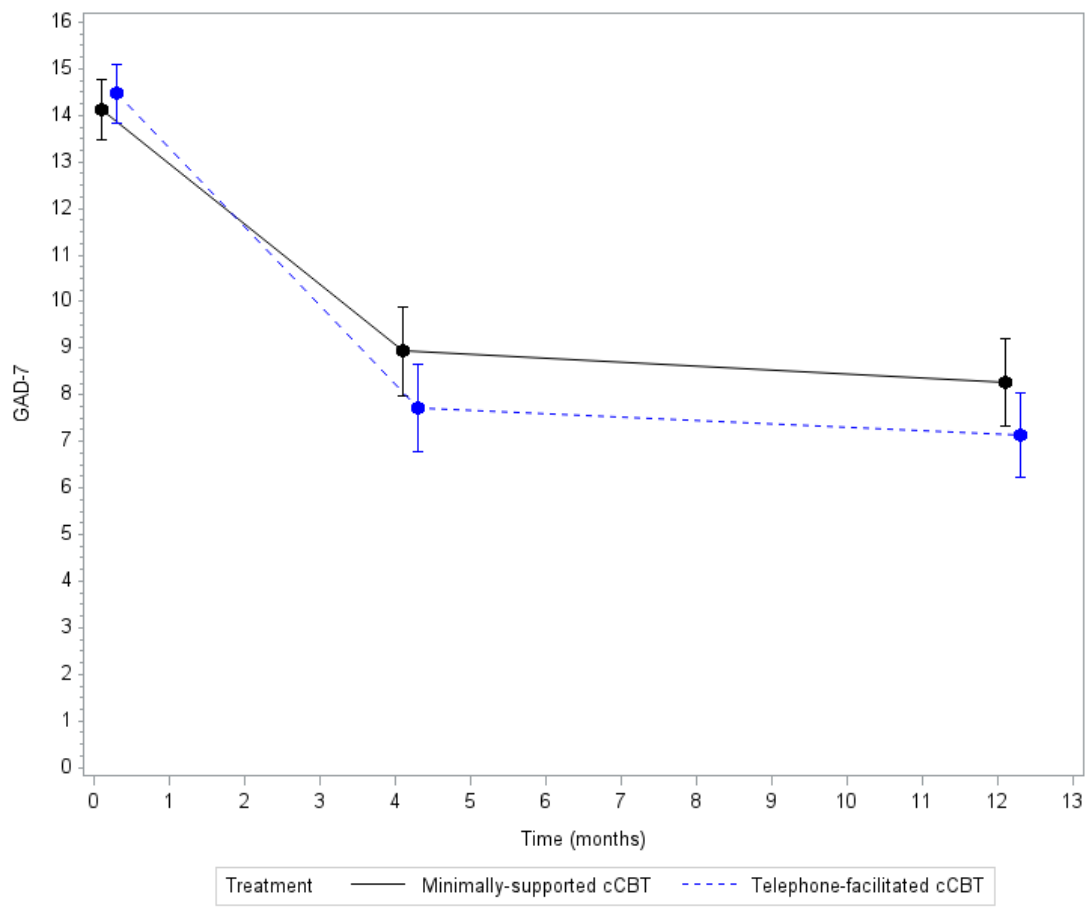**Figure 4: Severity of somatoform complaints (mean and 95% CI PHQ15 score) at each assessment**
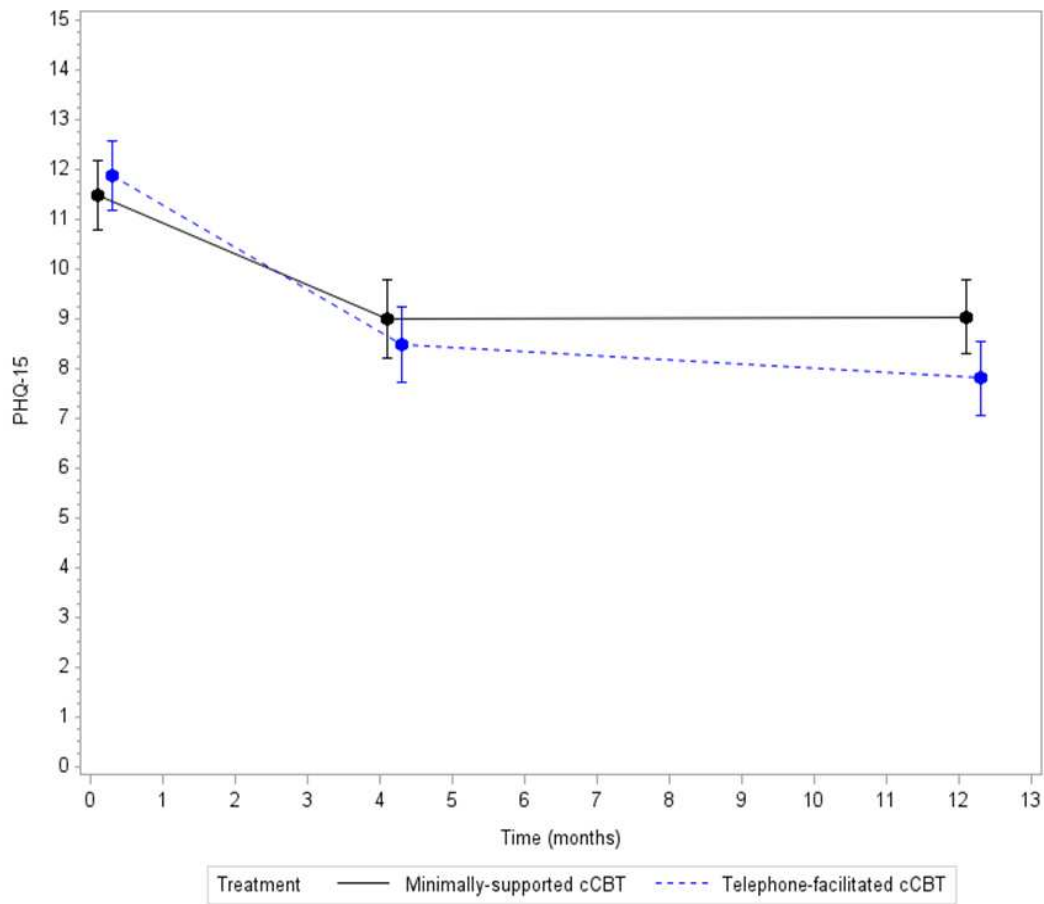

*Results from a repeated measures, mixed model adjusting for baseline score, age, gender, baseline GAD-7 score and time

Figure 5: MoodGYM usage as ascertained by computer login records

Table 1: Baseline characteristics of randomised participants

	Minimally-supported cCBT (n=182)	Telephone-facilitated cCBT (n=187)	Total (n=369)
SEX			
Female	113 (62.1%)	125 (66.8%)	238 (64.5%)
PATIENT AGE (YEARS)			
Mean (SD)	40.3 (13.7)	41.0 (13.8)	40.6 (13.8)
PREVIOUS EPISODES OF DEPRESSION WHEN HELP WAS SOUGHT?			
Yes	127 (69.8%)	134 (71.7%)	261 (70.7%)
CURRENTLY TAKING MEDICATION FOR DEPRESSION			
Yes	71 (39.0%)	72 (38.5%)	143 (38.8%)
No	111 (61.0%)	115 (61.5%)	226 (61.2%)
Mean PHQ9 score (SD)	16.4 (4.1)	16.8 (3.9)	16.6 (4.0)
Mean GAD7 score (SD)	14.1 (4.4)	14.5 (4.4)	14.3 (4.4)
Mean PHQ15 score (SD)	11.5 (4.8)	11.9 (5.0)	11.7 (4.9)

Table 2: Depression severity (PHQ-9 scores) at each time point

	Baseline		Month 4		Month 12	
	Minimally-supported	Telephone-facilitated	Minimally-supported	Telephone-facilitated	Minimally-supported	Telephone-facilitated
PHQ-9 continuous score						
Mean (SD)	16.4 (4.1)	16.8 (3.9)	10.4 (6.4)	8.5 (6.3)	9.2 (6.2)	8.2 (6.4)
Median	16	17	10	7	9	7
Range	10-25	10-26	0-27	0-24	0-25	0-27
n	182	187	128	141	132	142
PHQ-9 dichotomised						
Depressed N (%)	182 (100%)	187 (100%)	66 (51.6%)	51 (36.2%)	57 (43.2%)	46 (32.4%)
Not depressed N (%)	0	0	62 (48.4%)	90 (63.8%)	75 (56.8%)	96 (67.6%)
Missing	0	0	54 (30%)	46 (25%)	50 (27%)	45 (24%)

Table 3: Between group differences in depression severity (PHQ9) at four and 12 months and mixed repeated measure across all time points

Mean differences in PHQ-9 score						
Effect	Cohen's d effect size	Estimate	Lower 95% CI	Upper 95% CI	t Value	p-value
Telephone-facilitated cCBT vs. minimally-supported cCBT Month 4	0.324	-1.8923	-3.2969	-0.4877	2.65	0.0085
Telephone-facilitated cCBT vs. minimally-supported cCBT Month 12	0.155	-0.9192	-2.3341	0.4957	1.28	0.2020
Telephone-facilitated cCBT vs. minimally-supported cCBT (over all time points)	0.274	-1.4057	-2.6336	-0.1748	2.25	0.0253

Table 4: Between group differences in anxiety severity (GAD7) at four and 12 months and mixed repeated measure across all time points

Mean differences in GAD-7 score						
Effect	Cohen's d effect size	Estimate	Lower 95% CI	Upper 95% CI	t Value	p-value
Telephone-facilitated cCBT vs. minimally-supported cCBT Month 4	0.236	-1.2291	-2.4374	0.1425	1.85	0.0659
Telephone-facilitated cCBT vs. minimally-supported cCBT Month 12	0.166	-1.1269	-2.3122	0.1676	1.75	0.0819
Telephone-facilitated cCBT vs. minimally-supported cCBT (over all assessments)	0.255	-1.1780	-2.2813	-0.0747	2.10	0.0365

Table 5: Between group differences in severity of somatoform complaints (PHQ15) at four and 12 months and mixed repeated measure across all time points

Mean differences in PHQ-15 score						
Effect	Cohen's d effect size	Estimate	Lower 95% CI	Upper 95% CI	t Value	p-value
Telephone-facilitated cCBT vs. minimally-supported cCBT Month 4	0.121	-0.5088	-1.5701	0.5526	0.94	0.3460
Telephone-facilitated cCBT vs. minimally-supported cCBT Month 12	0.300	-1.2410	-2.2692	-0.2127	2.38	0.0182
Telephone-facilitated cCBT vs. minimally-supported cCBT (over all assessments)	0.303	-1.1099	-1.7521	0.0024	1.96	0.0506

Table 6: MoodGYM usage data

	Module 1		Module 2		Module 3		Module 4
	Minimally-supported	Telephone-facilitated	Minimally-supported	Telephone-facilitated	Minimally-supported	Telephone-facilitated	Minimally-supported
Logged on but didn't complete 25%			7.1%	5.4%	6%	8.1%	1.1%
25% complete	1.6%	2.2%	2.2%	3.8%	2.7%	2.7%	
50% complete	1.6%	2.2%	3.8%	3.8%	0.5%	4.3%	0.5%
75% complete	2.2%	4.8%	0.5%	3.8%	0.5%	1.1%	
100% complete	45.1%	64.5%	29.1%	46.2%	17.6%	28.5%	14.3%