Journal of Affective Disorders Reports 6 (2021) 100244

Contents lists available at ScienceDirect

Journal of Affective Disorders Reports

journal homepage: www.sciencedirect.com/journal/journal-of-affective-disorders-reports

Brief Report

Comparing the effectiveness of an 8-week and 12-week cognitive behavioural therapy group for bipolar affective disorder

Michael Kavanagh [a,*], Katharine Brouwer [b], Pete Lawrence [c]

[a] Trainee Clinical Psychologist, School of Psychology, University of Southampton, Building 44 SO17 1BJ, United Kingdom
[b] Highly Specialist Clinical Psychologist, Affective Disorders Pathway Lead, Southampton Psychological Services, Southern Health NHS Foundation Trust, United Kingdom
[c] Centre for Innovation in Mental Health, School of Psychology, University of Southampton, United Kingdom

ARTICLE INFO

Keywords:
Service evaluation
Bipolar affective disorder
Cognitive behavioural therapy

ABSTRACT

Background: Previous systematic reviews have demonstrated efficacy of Cognitive Behavioural Therapy Groups (CBT-G) in treating bipolar affective disorder (BAD). However, effectiveness research of BAD CBT-G (groups delivered in clinical practice rather than a research trial) is sparse. Additionally, the efficacy literature shows variation in the number of sessions delivered, and the number needed for clinically significant change is unclear. Therefore, we examine the effectiveness of a 12-week CBT-G compared to an 8-week group.
Methods: An 8-week vs 12-week CBT-G was delivered in routine practice in an adult community mental health team. We compared pre-post data for $N = 88$ participants with a diagnosis of BAD who attended either an 8-week CBT-G ($n = 43$) or 12-week CBT-G ($n = 45$).
Results: All routine outcome measure scores which included measures related to depression, generalised anxiety, psychological distress and functioning improved significantly from baseline to treatment endpoint for both the 8-week and 12-week CBT-G interventions, with no significant differences between the groups at post-treatment.
Limitations: No measure for manic symptoms was included. No follow-up data was collected. The study lacks a comparator control group.
Conclusions: This research adds to the literature in two ways by: (i) demonstrating the effectiveness of CBT-G for bipolar affective disorder; (ii) being the first to show no significant difference in outcome measures between an 8-week and 12-week group. The findings can be used to inform the provision of both a clinical and cost-effective intervention.

1. Introduction

Bipolar affective disorder (BAD) is characterised by manic, hypomanic, depressive and mixed episodes (Salcedo et al., 2016). With a lifetime prevalence of around 1% (Merikangas et al., 2011), BAD is associated with poor psychosocial functioning and a high economic burden owing to the severity and chronicity of the condition (Fagiolini et al., 2013).

Psychological interventions are recommended as an adjunctive treatment to psychiatric medications in many guidelines for BAD (e.g. National Institute for Health and Care Excellence 2014). Salcedo et al. (2016) reported a systematic review of randomized control trials (RCTs) of psychosocial interventions for BAD which supported CBT-G as an efficacious intervention. However, service evaluation research of effectiveness in routine clinical settings is scarce. This is important because whilst RCTs exclusion criteria maximize opportunities to detect treatment effects, the consequences can be a loss of external validity in clinical practice (Scott et al., 2009). For example, Hoertel et al. (2013) demonstrated that over 50% of individuals with BAD would be excluded from most RCTs because of symptom severity or co-morbidity. People with BAD excluded from RCTs due to complexity are likely to be offered treatment in community mental health teams (CMHTs) (Hoertel et al., 2013). Considering this, it is important to evaluate CBT-G service interventions to identify any efficacy-effectiveness gaps to inform subsequent examination of how to address these (Hoertel et al., 2013).

The scarce service evaluations of group therapy for BAD that have been published are psychoeducation or mixed-model treatment groups, but not CBT-G (Etain et al., 2018). Therefore, this study's primary objective is to evaluate the effectiveness of CBT-G for BAD within an adult CMHT.

* Corresponding author.
E-mail address: m.kavanagh@soton.ac.uk (M. Kavanagh).

https://doi.org/10.1016/j.jadr.2021.100244
Received 17 May 2021; Received in revised form 30 September 2021; Accepted 1 October 2021
Available online 5 October 2021

The controlled efficacy literature has shown BAD CBT-G to be efficacious at different session lengths including at: 20 sessions (Gonzalez-Isai et al., 2012), 14 sessions (Costa et al., 2011) and 12 sessions (Castle et al., 2010). Considering services are typically concerned with delivering cost-efficient treatments (Naylor et al., 2015) and number of treatment sessions is one factor affecting cost, establishing the session-length needed for CBT-G effectiveness could save service resources. Therefore, a secondary objective of this study is to compare whether the outcomes of an 8-week CBT-G differs from a 12-week group. As far as we are aware, this is not something that has been done in the literature.

Considering the efficacy literature related to previous CBT-G interventions, we hypothesised that (i) there will be improved mean scores from pre- to-post intervention on routine outcome measures (ROMs) for both the 12-week and 8-week CBT-G and; (ii) there will be no significant difference in the mean pre-post ROMs scores between the 12-week and 8-week CBT-G post treatment.

In addition to statistically comparing the mean pre-post scores of both groups, the criteria of reliable and clinically significant improvement (RCSI) will be used. These criteria are increasingly taken as a credible index of psychological recovery in the literature (Barkham et al., 2012).

2. Methods

2.1. Design

The study used a 2 × 2 mixed model design. The between subject factor levels were the 8-week and 12-week CBT-G. The within subject factor levels were time points of pre and post CBT-G. Approval was given as a service evaluation by the hosting National Health Service (NHS) trust. Ethical approval was also given by The University Of Southampton, Ethics and Research Governance Online code 55,666). Participants gave written informed consent for their data to be used for a service evaluation.

2.2. Measures

Outcome measures used were those routinely given to clients receiving care from the CMHT psychology service. This was so participants were not asked to provide more data than usual in treatment. Completing routine outcome measures (ROMs) was optional and not completing ROMs had no impact on the care provided.

2.2.1. CORE-OM: clinical outcomes in routine evaluation outcome measure (Evans et al., 2002)

The CORE-OM is a 34-item self-report questionnaire measuring psychological distress in four dimensions: subjective well-being, problems/symptoms, life functioning and risk/harm. Total mean scores above 1 indicate clinically significant distress. The CORM-OM has demonstrated good reliability (Cronbach's alpha =0.75–0.94) and validity, with large differences between clinical and non-clinical samples (Evans et al., 2002).

2.2.2. GAD7: generalised anxiety disorder assessment-7 (Spitzer et al., 2006)

Because anxiety disorders are the most prevalent comorbidity in BAD (Ott, 2018), it was important to include an anxiety measure. The GAD-7 is a seven-item self-report screening tool for measuring generalised anxiety disorder severity. Scores of 5, 10, and 15 are taken as cut-off points for mild, moderate and severe anxiety, respectively. Using a threshold score of 10, the GAD-7 has a sensitivity of 89% and a specificity of 82%. The GAD-7 has also demonstrated good reliability (Cronbach's alpha =0.92) (Spitzer et al., 2006).

2.2.3. PHQ9: patient health questionnaire-9 (Kroenke et al., 2001)

The PHQ-9 is a nine-item self-report screening tool for measuring depression severity. Scores of 5, 10, 15, and 20 represent mild, moderate, moderately severe, and severe depression, respectively. Using the threshold score of 10, the PHQ-9 has a sensitivity of 88% and a specificity of 88% for major depression. The PHQ-9 has also demonstrated good reliability (Cronbach's alpha =0.89) (Kroenke et al., 2001).

2.2.4. Work and Social Adjustment Scale (WSAS) (Mundt et al., 2002)

The WSAS is a five-item self-report measure of impact on functioning from the reported problem. Scores of < 10, 10–20 and > 20 represent subclinical populations, functional impairment but less severe symptomology, and moderately severe or worse psychopathology respectively. The WSAS has demonstrated good reliability (Cronbach's alpha =0.70–0.94) and is sensitive to differences in disorder severity (Mundt et al., 2002).

2.3. Participants

All participants ($N = 96$) had a diagnosis of BAD. Diagnoses were made by a psychiatrist via medical consultation in line with the DSM-5 or ICD-10 criteria. Referrals to the group were made by staff in the NHS secondary care CMHT for working age adults with severe and enduring mental health problems. Participants were then assessed for suitability by a clinical psychologist via clinical interview (assessing for clear periods of hypomania/mania and depression using a CBT 5-areas formulation (Greenberger and Padesky, 1995) and offered a place in the CBT-G as part of routine treatment.

Of the $n = 71$ people that were referred to the 12-week CBT-G, $n = 6$ did not respond to be seen for the suitability assessment. Those that did attend the suitability assessment, $n = 4$ were unable to attend the group and offered a 1:1 intervention instead, $n = 4$ were referred to an emotional coping skills group as a more appropriate intervention, $n = 2$ declined the group, $n = 1$ had no experience of hypomania/mania and offered CBT for depression, $n = 1$ were already receiving private CBT for BAD. Therefore, $n = 53$ accepted a place on the CBT-G after being deemed suitable.

Of the $n = 54$ people that were referred to the 8-week CBT-G, $n = 6$ did not respond to be seen for the suitability assessment. Those that did attend the suitability assessment, $n = 1$ was unable to attend the group and offered a 1:1 intervention instead, $n = 1$ requested a 1:1 intervention rather than a CBT-G, $n = 1$ was referred on to an emotional coping skills group as a more appropriate intervention, $n = 1$ had no experience of hypomania/mania and offered CBT for depression, $n = 1$ opted for a counselling intervention. Therefore, $n = 43$ accepted a place on the CBT-G after being deemed suitable.

The 12-week group $n = 53$ (24.5% ($n = 13$) male, 75.5% ($n = 40$) female), had a mean age of 41.3 years (SD=10.89). The majority identified as White British (66%, $n = 35$), 4% ($n = 2$) as Black, 4% ($n = 2$) as Asian, 2% ($n = 1$) as Mixed Black and White, 2% ($n = 1$) as White Other and 23% ($n = 12$) as undisclosed.

The 8-week group $n = 43$ (30.2% ($n = 13$) male, 69.8% ($n = 30$) female), had a mean age of 39.49 years (SD=9.79). The majority identified as White British (70%, $n = 30$), 9% ($n = 4$) as White Other, 5% ($n = 2$) as Black, 2% ($n = 1$) as Mixed Race and 14% ($n = 6$) as undisclosed.

2.4. Group content

Groups ran weekly for two hours and were facilitated by two qualified Clinical Psychologists. The data used are from seven 12-week groups delivered from 2013 to 2016 and four 8-week groups delivered from 2017 to 2019. Participants completed ROMs pre-post CBT-G. Table 1 summarizes the content of each of the 8-week and 12-week group sessions. The group followed some didactic delivery of information but also focused on sharing experiences and asking questions.

Table 1
Group session content.

	12-week Group	8-week Group
Session 1	Introduction to the group. Group Rules. Discussion of diagnosis. Acknowledge people are at different stage of journey. Mood chart.	Introduction to bipolar disorder. What is depression and Mania? Causes of bipolar disorder.
Session 2	What is depression and Mania? Causes of bipolar disorder. Stress vulnerability model. Goals.	Depression – triggers, warning signs and symptoms. Hot cross bun. Goals.
Session 3	Depression – triggers, warning signs and symptoms. Hot cross bun.	Interventions for lows. Activity Scheduling. Acting Opposite.
Session 4	Interventions for lows. Activity Scheduling. Acting Opposite.	Interventions for lows. Recognising NATs. Thought Challenging. Positive data diary.
Session 5	Interventions for lows. Recognising NATs. Thought Challenging. Positive data diary.	Mania – triggers, warning signs and symptoms. Hot cross bun. Pros and cons.
Session 6	Mania – triggers, warning signs and symptoms. Hot cross bun. Pros and cons.	Interventions for highs. Activity Scheduling. Living by values. Acting Opposite.
Session 7	Interventions for highs. Activity Scheduling. Living by values. Acting Opposite.	Interventions for highs. Relaxation. Mindfulness. Sleep hygiene. Thought balancing.
Session 8	Interventions for highs. Relaxation. Mindfulness. Sleep hygiene. Thought balancing.	Relapse prevention and maintaining wellness. Summary, *Re*-Cap, plans for the future.
Session 9	Life balance and managing relationships.	N/A
Session 10	Problem-solving, prioritising.	N/A
Session 11	Relapse prevention and maintaining wellness.	N/A
Session 12	*Re*-Cap, plans for the future.	N/A

3. Results

Intent to Treat Analysis (ITTA) and per protocol analysis (PPA) can each introduce different biases into conclusions (Brody, 2016). Therefore, researchers often conduct both types of analysis (Brody, 2016) as was done in this study. For both the PPA and ITTA, six participants (n = 4 for the 8-week CBT-G and n = 2 for the 12-week CBT-G) had one item missing for the WSAS measure pre or post group and pro-rating was used for these missing items.

Drop-out across both groups was defined as missing two or more sessions. This drop-out definition was arrived at through the clinical experience of the Clinical Psychologists who facilitated the groups, believing that missing two or more sessions would impact the effectiveness of the intervention received. For example, missing two sessions could mean fully missing the depression intervention modules. Additionally, sessions build on the previous week's material and attendance ensures a cohesive core group is formed.

Out of the n = 43 participants who were assigned to the 8-week CBT-G, n = 16 (37.2%) dropped-out. Here n = 2 missed two sessions, n = 2 missed three sessions, n = 1 missed four sessions and n = 11 missed 5 or more sessions including not attending any sessions. Of the n = 16 who dropped out, drop out reasons included mental health crisis (n = 2), social stressors e.g. court case, housing, (n = 5), did not respond to contact (n = 9). Out of the n = 52 participants who started the 12-week CBT-G, n = 19 (35.8%) dropped-out. With the data available for the 12-week CBT-G it was not possible to distinguish the number of sessions attended before drop-out or the reasons for drop-out.

Out of the n = 35 participants that dropped-out from the study, n = 8 were missing both pre and post data (data was missing because completing the outcome measures was not mandatory in order to attend the group or they did not attend any sessions) and therefore were not included in the ITTA. Therefore, a modified-ITTA was used as the primary analysis. Three participants that dropped-out had pre-post data that was used in the modified-ITTA. For the remaining n = 24 participants, pre-data was used as post-data. Using a modified-ITTA, a 2 × 2 mixed model Multivariate Analysis of Variance (MANOVA) was performed with group length (8-week vs 12-week CBT-G) as the between-subjects factor and time point (pre and post group) as the within-subjects factor. The MANOVA included the WSAS, PHQ-9, GAD-7 and CORE-OM measures. The MANOVA showed a significant main effect of therapy time point (V = 0.322, F(4,83)=9.86, p < 001, ηp^2=0.322) indicating participants improved over time. Univariate Analysis of Variance (ANOVA) results showed that there were significant differences between the time points for the CORE-OM (F(1,86)=37.61, p < 001, ηp^2=0.304), the PHQ-9 (F(1,86)=27.84, p < 001, ηp^2=0.245), the GAD-7 (F(1,86)=22.22, p<.001, ηp^2=0.205) and the WSAS (F(1,86)= 10.21, p < 001, ηp^2=0.106). However, the MANOVA main effect of group length was not significant (V = 0.043, F(4,83)=0.93, p=.45, ηp^2=0.043) which indicated there was no significant difference between the 8-week and 12-week groups over time. Additionally, the group length X time point interaction was not significant (V = 0.030, F(4,83)= 0.633, p=.64, ηp^2=0.030).

A secondary PPA was conducted using a 2 × 2 MANOVA again with the same between-subjects and within subject factors and the same dependant variables. The MANOVA showed a significant main effect of therapy time point (V = 0.433, F(4,56)=10.68, p < 001, ηp^2=0.433) indicating participants improved over time. Univariate ANOVA results showed that there were significant differences between the time points for the CORE-OM (F(1,59)=39.97, p < 001, ηp^2=0.404), the PHQ-9 (F (1,59)=32.61, p < 001, ηp^2=0.356), the GAD-7 (F(1,59)=24.80, p < 001, ηp^2=0.296) and the WSAS (F(1,59)=12.75, p < 001, ηp^2=0.178). However, the MANOVA main effect of group length was not significant (V = 0.077, F(4,56)=1.16, p=.338, ηp^2=0.077) which indicated there was no significant difference between the 8-week and 12-week groups over time. Additionally, the group length X time point interaction was not significant (V = 0.060, F(4,56)=0.892, p=.475, ηp^2=0.060).

The results of both the modified-ITTA and the PPA mixed model MANOVAs are consistent with the hypotheses, that (i) scores will improve from pre-treatment to post-treatment on ROMs for both the 8-week and 12-week BAD CBT-G interventions and (ii) there will be no difference in improved ROMs scores between the 8-week and 12-week group post treatment.

Mean ROMs scores pre-post 8-week and 12-week CBT-G for the modified-ITTA are shown in Table 2. The same is shown for the PPA in Table 3. Higher mean scores indicate more severe clinical scores. The mean scores indicate that ROMs improved for both the 8-week and 12-week CBT-G from pre-post group.

3.1. Reliable and clinically significant improvement (RCSI)

For the CORE-OM, a clinical cut-off of ≥ 1 was used and a reliable change index (RCI) of ≥ 0.5 as recommended by Connell et al. (2007). For the PHQ-9, a clinical cut-off of ≥ 10 was used and an RCI of ≥ 6 as recommended by Clark and Oats (2014). For the GAD-7, a clinical cut-off of ≥ 10 was used and an RCI of ≥6 as recommended by Bischoff et al. (2020).

For the WSAS, a clinical cut-off of ≥10 was used as recommended by

Table 2
Modified-ITTA mean outcome measure scores for the 8-week and 12-week CBT-G pre-post treatment with RCSI of the sample's clinical population.

Time Point	Group Length	Outcome Measure	*N*	*Mean*	*SD*	Participants in the Clinical PopulationPre-Group% (n)	Reliable Change Improvementof the Clinical Population%(n)	RCSI of the Clinical Population% (n)
Pre Group	8-Week	CORE-OM	43	1.65	0.83	77% (33)		
	12-Week	CORE-OM	45	1.85	0.62	91% (41)		
	Total	CORE-OM	88	1.75	0.73	84% (74)		
Post Group	8-Week	CORE-OM	43	1.25	0.75		35% (15)	14% (6)
	12-Week	CORE-OM	45	1.53	0.69		40% (18)	16% (7)
	Total	CORE-OM	88	1.39	0.73		38% (33)	15% (13)
Pre Group	8-Week	PHQ-9	43	14.60	7.94	72% (31)		
	12-Week	PHQ-9	45	15.64	6.71	65% (32)		
	Total	PHQ-9	88	15.14	7.31	72% (63)		
Post Group	8-Week	PHQ-9	43	10.79	7.53		28% (12)	23% (10)
	12-Week	PHQ-9	45	12.82	6.39		29% (13)	18% (8)
	Total	PHQ-9	88	11.83	7		28% (25)	20% (18)
Pre Group	8-Week	GAD-7	43	10.09	6.8	48% (21)		
	12-Week	GAD-7	45	11.56	5.11	69% (31)		
	Total	GAD-7	88	10.84	6	59% (52)		
Post Group	8-Week	GAD-7	43	7.98	6.1		26% (11)	21% (9)
	12-Week	GAD-7	45	9.13	5.31		24% (11)	22% (10)
	Total	GAD-7	88	8.56	5.7		25% (22)	22% (19)
Pre Group	8-Week	WSAS	43	17.88	9.63	77% (33)		
	12-Week	WSAS	45	20.09	9.06	84% (38)		
	Total	WSAS	88	19.01	9.35	81% (71)		
Post Group	8-Week	WSAS	43	15.4	8.52		21% (9)	9% (4)
	12-Week	WSAS	45	17.84	7.56		16% (7)	4% (2)
	Total	WSAS	88	16.64	8.09		18% (16)	7% (6)

CORE-OM: Clinical Outcomes in Routine Evaluation Outcome Measure (Evans et al., 2002). PHQ-9: Patient Health Questionnaire-9 (Kroenke et al., 2001). GAD-7: Generalised Anxiety Disorder Assessment-7 (Spitzer et al., 2006). WSAS: Work and Social Adjustment Scale (Mundt et al., 2002). RCSI: Reliable and clinically significant improvement.

Mundt et al. (2002). An RCI of ≥10 was calculated for the WSAS based on a reliability score of 0.86 (Mundt et al., 2002) and the *SD* of this study's sample (9.35). To establish the RCI for the WSAS, the standard error of measurement (*SEM*) was needed, which is $SD(9.35)X\sqrt{1}$ – the reliability of the measure (0.86) =3.46 (Jacobson and Traux, 1992). Next the standard error of difference (*Sdiff*) was calculated which is $\sqrt{2}$ $(SEM\ (3.46))^2$ =4.89 (Jacobson and Traux, 1992). The RCI was then calculated as *Sdiff*(4.89)X1.96 (1.96 used for a 95% confidence interval) (Jacobson and Traux, 1992) =9.58 (rounded up to 10 for this study).

The percentage of the study's sample that were in the clinical population at pre-group, as well as their reliable change and RCSI experienced from the 8-week and 12-week CBT-G are in Table 2 (modified-ITTA) and Table 3 (PPA).

The RCSI associations between groups were tested with chi-squared, and no significant difference was found for either a modified-ITTA (CORE-OM X^2 (1,N = 74)=0.019, p=.901; PHQ-9 X^2 (1,N = 65)= 0.618, p=.432; GAD-7 X^2 (1,N = 53)=0.518, p = 419; WSAS X^2 (1,N = 71)=1.07, p=.3) or a PPA (CORE-OM X^2 (1,N = 50)=0.333, p=.564; PHQ-9 X^2 (1,N = 43)=1.623, p=.203; GAD-7 X^2 (1,N = 37)=0.504, p = 478; WSAS X^2 (1,N = 46)=0.103, p=.749). This further supports the secondary hypothesis.

4. Discussion

The aims of this study were to investigate the effectiveness of BAD CBT-G and whether a shorter 8-week CBT-G differs in outcomes from a 12-week group. The findings suggest that CBT-G in CMHT settings are an effective treatment for BAD when assessed with ROMs of depression, generalised anxiety, impact on functioning and global distress. This supports previous research which has found efficacy for CBT-G interventions in RCT conditions (Salcedo et al., 2016). The findings also indicated no significant difference between an 8-week and 12-week CBT-G in treating BAD at the end of treatment. This is in line with previous studies that have found efficacy for CBT-G at different session lengths (Salcedo et al., 2016).

This study builds on previous research by investigating the effectiveness of CBT-G for BAD in a CMHT setting. It is also the first study to evaluate ROMs differences between a shorter 8-week CBT-G intervention and a 12-week intervention.

Considering there was no significant difference in outcomes between the 8-week and 12-week CBT-G, the implications are that CMHT service leads can run more cost-efficient services, saving budget-costs, clinician time and cut waiting lists. This is also potentially important for commissioners of healthcare services who are likely to be making restrictive budget cuts in a resulting Covid-19 recession. It might also be important for future NICE guideline reviews which currently recommend a treatment length of 16–20 sessions. Future research could investigate whether group delivery is the significant variable in change (e.g. shared experiences, normalising with others) or if individual 8-week vs 12-week sessions would have the same impact. An RCT could further investigate equivalence and non-inferiority between the groups.

4.1. Limitations

The study has several limitations. First, a measure for manic/hypomanic symptoms was not included and so an important element of BAD was not evaluated. However, the study did include measures of functioning which Jonas et al. (2012) note are often missing in efficacy studies, yet most meaningful to patients. Second, the study lacks a comparator control group. As the study is effectively a comparison of an 8-week CBT-*G* + treatment as usual (TAU) vs 12 CBT-*G* +TAU, a TAU condition without CBT is a limitation and means improvements could be

Table 3
The PPA mean outcome measure scores for the 8-week and 12-week CBT-G pre-post treatment as well as RCSI of the sample's clinical population.

Time Point	Group Length	Outcome Measure	*N*	*Mean*	*SD*	Participants in the Clinical PopulationPre-Group% (n)	Reliable Change Improvementof the Clinical Population%(n)	RCSI of the Clinical Population% (n)
Pre Group	8-Week	CORE-OM	27	1.60	0.85	70% (19)		
	12-Week	CORE-OM	34	1.84	0.68	91% (31)		
	Total	CORE-OM	61	1.73	0.76	82% (50)		
Post Group	8-Week	CORE-OM	27	1.02	0.6		69% (13)	26% (5)
	12-Week	CORE-OM	34	1.45	0.72		55% (17)	19% (6)
	Total	CORE-OM	61	1.26	0.69		60% (30)	22% (11)
Pre Group	8-Week	PHQ-9	27	14.11	8.33	70% (19)		
	12-Week	PHQ-9	34	15.26	7.04	65% (22)		
	Total	PHQ-9	61	14.75	7.60	67% (41)		
Post Group	8-Week	PHQ-9	27	8.78	6.38		58% (11)	47% (9)
	12-Week	PHQ-9	34	11.29	5.79		59% (13)	36% (8)
	Total	PHQ-9	61	10.21	6.13		59% (24)	41% (17)
Pre Group	8-Week	GAD-7	27	9.56	7.14	48% (13)		
	12-Week	GAD-7	34	11.79	5.03	71% (24)		
	Total	GAD-7	61	10.80	6.10	61% (37)		
Post Group	8-Week	GAD-7	27	6.59	5.49		69% (9)	54% (7)
	12-Week	GAD-7	34	8.38	4.91		46% (11)	42% (10)
	Total	GAD-7	61	7.59	5.21		54% (20)	46% (17)
Pre Group	8-Week	WSAS	27	17.93	9.77	70% (19)		
	12-Week	WSAS	34	20.44	8.77	79% (27)		
	Total	WSAS	61	19.33	9.23	75% (46)		
Post Group	8-Week	WSAS	27	13.96	7.75		47% (9)	21% (4)
	12-Week	WSAS	34	17.11	7.48		25% (7)	7% (2)
	Total	WSAS	61	15.22	7.7		35% (16)	13% (6)

CORE-OM: Clinical Outcomes in Routine Evaluation Outcome Measure (Evans et al., 2002). PHQ-9: Patient Health Questionnaire-9 (Kroenke et al., 2001). GAD-7: Generalised Anxiety Disorder Assessment-7 (Spitzer et al., 2006). WSAS: Work and Social Adjustment Scale (Mundt et al., 2002). RCSI: Reliable and clinically significant improvement.

due to variables other than treatment. Third, as measures were collected at pre-post time points only, the lack of more frequent outcome assessment during the group is a limitation. Fourth, it is unknown if positive changes were sustained as no follow-up data was collected.

5. Conclusions

This study adds to the literature by providing evidence consistent with: (i) CBT-G for BAD being effective in a CMHT setting, (ii) no significant difference in ROMS between a shorter 8-week vs 12-week CBT-G.

CRediT authorship contribution statement

Michael Kavanagh: Writing – original draft, Writing – review & editing, Formal analysis. **Katharine Brouwer:** Data curation, Writing – original draft, Writing – review & editing. **Pete Lawrence:** Formal analysis, Writing – review & editing.

Declaration of Competing Interest

Michael Kavanagh carried out a training placement at the National Health Service secondary care community mental health team where the study was conducted. Katharine Brouwer is employed by the National Health Service secondary care community mental health team where the study was conducted and has facilitated the Bipolar Disorder Cognitive Behavioural Therapy Group. Pete Lawrence has no conflicts of interest to declare.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Acknowledgements

We would like to acknowledge the valuable contributions of Katherine Bradbury for facilitating the bipolar affective disorder cognitive behavioural therapy groups.

References

Barkham, M., Stiles, W.B., Connell, J., Mellor-Clark, J., 2012. Psychological treatment outcomes in routine NHS services: what do we mean by treatment effectiveness? Psychol. Psychother. Theory Res. Pract. 85 (1), 1–16. https://doi.org/10.1111/j.2044-8341.2011.02019.x.

Bischoff, T., Anderson, S.R., Heafner, J., Tambling, R., 2020. Establishment of a reliable change index for the GAD-7. Psychol., Commun. Health 8 (1), 176–187. https://doi.org/10.5964/pch.v8i1.309.

Brody, T., 2016. Clinical Trials: Study Design, Endpoints and Biomarkers, Drug safety, and FDA and ICH Guidelines. Academic press, London, UK eBook ISBN: 9780128042588.

Castle, D., White, C., Chamberlain, J., Berk, M., Berk, L., Lauder, S., Gilbert, M., 2010. Group-based psychosocial intervention for bipolar disorder: randomised controlled trial. Br. J. Psychiatry 196 (5), 383–388. https://doi.org/10.1192/bjp.bp.108.058263.

Clark, D.M., Oates, M., 2014. Improving access to psychological therapies: measuring improvement and recovery adult services. Lond. NHS Engl. https://doi.org/10.1093/med:psych/9780195389050.003.0004.

Connell, J., Barkham, M., Stiles, W.B., Twigg, E., Singleton, N., Evans, O., Miles, J.N., 2007. Distribution of CORE–OM scores in a general population, clinical cut-off points and comparison with the CIS–R. Br. J. Psychiatry 190 (1), 69–74. https://doi.org/10.1192/bjp.bp.105.017657.

Costa, R.T.D., Cheniaux, E., Rosaes, P.A.L., Carvalho, M.R.D., Freire, R.C.D.R., Versiani, M., Nardi, A.E., 2011. The effectiveness of cognitive behavioral group

therapy in treating bipolar disorder: a randomized controlled study. Braz. J. Psychiatry 33 (2), 144–149. https://doi.org/10.1590/S1516-44462011000200009.
Etain, B., Scott, J., Cochet, B., Bellivier, F., Boudebesse, C., Drancourt, N., Richard, J.R., 2018. A study of the real-world effectiveness of group psychoeducation for bipolar disorders: is change in illness perception a key mediator of benefit? J. Affect. Disord. 227, 713–720. https://doi.org/10.1016/j.jad.2017.11.072.
Evans, C., Connell, J., Barkham, M., Margison, F., McGRATH, G.R.A.E.M.E., Mellor-Clark, J., Audin, K., 2002. Towards a standardised brief outcome measure: psychometric properties and utility of the CORE–OM. Br. J. Psychiatry 180 (1), 51–60. https://doi.org/10.1192/bjp.180.1.51.
Fagiolini, A., Forgione, R., Maccari, M., Cuomo, A., Morana, B., Dell'Osso, M.C., Rossi, A., 2013. Prevalence, chronicity, burden and borders of bipolar disorder. J. Affect. Disord. 148 (2–3), 161–169. https://doi.org/10.1016/j.jad.2013.02.001.
González-Isasi, A., Echeburúa, E., Limiñana, J.M., González-Pinto, A., 2012. Predictors of good outcome in patients with refractory bipolar disorder after a drug or a drug and cognitive-behavioral treatment. Compr. Psychiatry 53 (3), 224–229. https://doi.org/10.1016/j.comppsych.2011.05.001.
Greenberger, D., Padesky, C., 1995. Mind Over mood: Change How You Feel by Changing the Way you Think, 2nd ed. The Guildford Press, New York. https://doi.org/10.3399/bjgp16x685477.
Hoertel, N., Le Strat, Y., Lavaud, P., Dubertret, C., Limosin, F., 2013. Generalizability of clinical trial results for bipolar disorder to community samples: findings from the national epidemiologic survey on alcohol and related conditions. J. Clin. Psychiatry. https://doi.org/10.4088/JCP.12m07935.
Jacobson, N.S., Truax, P., 1992. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J. Consult. Clin. Psychol. 59, 12–19. https://doi.org/10.1037/10109-042.
Jonas, D.E., Mansfield, A.J., Curtis, P., Gilmore, J.H., Watson, L.C., Brode, S., Shetiman, B., 2012. Identifying priorities for patient-centered outcomes research for serious mental illness. Psychiatr. Serv. 63 (11), 1125–1130. https://doi.org/10.1176/appi. ps.201100369.
Kroenke, K., Spitzer, R.L., Williams, J.B., 2001. The PHQ-9: validity of a brief depression severity measure. J. Gen. Intern. Med. 16 (9), 606–613. https://doi.org/10.1046/j.1525-1497. 2001.016009606.x.
Merikangas, K.R., Jin, R., He, J.P., Kessler, R.C., Lee, S., Sampson, N.A., Ladea, M., 2011. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. Arch. Gen. Psychiatry 68 (3), 241–251. https://doi.org/10.1001/archgenpsychiatry.2011.12.
Mundt, J.C., Marks, I.M., Shear, M.K., Greist, J.M., 2002. The work and social adjustment scale: a simple measure of impairment in functioning. Br. J. Psychiatry 180 (5), 461–464. https://doi.org/10.1192/bjp.180.5.461.
Naylor, C., Imison, C., Addiott, R., Buck, D., Goodwin, N., Harrison, T., …. .& Curry, N. (2015). Transforming our health care system. The King's Fund. Retrieved from: https://dataingovernment.blog.gov.uk/2020/11/10/we-want-your-feedback-on-using-the-doi-standard-in-government/.
National Institute for Health and Clinical Excellence (2014). Bipolar disorder: Assessment and Management (Clinical Guideline no. 185). Updated: February 2020. Retrieved from: guidance.nice.org.uk/CG185.
Ott, C.A., 2018. Treatment of anxiety disorders in patients with comorbid bipolar disorder. Ment. Health Clin. 8 (6), 256–263. https://doi.org/10.9740/mhc.2018.11.256.
Salcedo, S., Gold, A.K., Sheikh, S., Marcus, P.H., Nierenberg, A.A., Deckersbach, T., Sylvia, L.G., 2016. Empirically supported psychosocial interventions for bipolar disorder: current state of the research. J. Affect. Disord. 201, 203–214. https://doi.org/10.1016/j.jad.2016.05.018.
Scott, J., Colom, F., Popova, E., Benabarre, A., Cruz, N., Valenti, M., Vieta, E., 2009. Long-term mental health resource utilization and cost of care following group psychoeducation or unstructured group support for bipolar disorders: a cost-benefit analysis. J. Clin. Psychiatry 70 (3), 378. https://doi.org/10.4088/JCP.08m04333.
Spitzer, R.L., Kroenke, K., Williams, J.B., Löwe, B., 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch. Intern. Med. 166 (10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092.