Evaluation of a Four Session Version of Stress Control

Dr Evelyn Gibson, Principal Clinical Psychologist, Lead for Governance / Delivery Outcomes, Research & Audit, Hywel Dda University Health Board

***Abstract**: Stress Control is an evidence-based, group intervention for mild to moderate anxiety and depression. Hywel Dda University Health Board (HDUHB) has been running a shortened (four session) version of Stress Control for a number of years. This evaluation aimed to determine the impact of this intervention. Participants were asked to complete session-by-session psychometric measures measuring symptoms of anxiety, depression and levels of general psychological distress. Participants and group facilitators also provided qualitative feedback on their experience of the course. The overall results indicated that there was a statistically significant impact of the intervention on all measures; however, the shortened version of the Stress Control course was found to be significantly less effective than the full version. On an individual level, the majority of participants did not show reliable recovery on any of the outcome measures. Attendance at the Stress Control course resulted in a significant reduction of the development of suicidal plans but not on the level of suicidal thoughts. Individuals who attended three or more sessions had significantly better outcomes than those who attended two or less. Group outcomes indicated that individuals with initial symptoms in the severe range had a significantly greater reduction in symptoms than those with symptoms in the mild to moderate range; however, this was not reflected on an individual level. Recommendations for the future of the Stress Control course in HDUHB are made based on these results.*

Introduction

Anxiety disorders are among the most common mental health problems seen in primary care settings (Kessler *et al*., 2005; Wittchen *et al*., 2011). According to large population-based surveys, up to 34% of the population are affected by an anxiety disorder during their lifetime (Bandelow & Michaelis, 2015). Anxiety disorders substantially impair quality of life (Rapaport, Clary, Fayyad & Endicott, 2005) and are associated with increased use of physical and mental health services (Wittchen *et al*., 2002). Anxiety symptoms often accompany depressive disorders (Clark & Watson, 1991). Epidemiological data indicates that 59% of individuals with generalized anxiety disorder (GAD) meet the criteria for major depressive disorder (MDD; Carter, Wittchen, Pfister & Kessler, 2001).

Psychoeducation has been reported to be effective in treating a range mental health disorders (Lukens & McFarlane, 2004). Stress Control (SC; White, 2000) is a group-based, psychoeducational, didactic intervention that is delivered as a series of six lecture-style sessions based on principles of cognitive behavioural therapy (CBT). SC was originally developed as an intervention for GAD; however, it has since been shown to be effective for other anxiety disorders and for depression (Wood, Kitchiner & Bisson, 2005; Joice & Mercer, 2010).

The National Institute for Health and Clinical Excellence (NICE) recommends the provision of stepped-care service delivery models for the treatment of mild to moderate depression and anxiety disorders (NICE, 2011). Stepped care services adopt an incremental approach to service provision with high volume, low intensity interventions being provided to individuals with the least complex difficulties.

Subsequent steps are defined by increasing levels of case complexity and increasingly intensive forms of treatment (NICE, 2011). Reviews of the stepped-care approach have been positive (Firth, Barkham & Kellett, 2015). When compared to treatment as usual (TAU), individuals in a stepped care system have been found to improve more quickly and have less absenteeism from work (Oosterbaan *et al*., 2013). Welsh guidelines state that "*there is an expectation that psychological therapy services will usually be delivered within a stepped/tiered care model*" (National Psychological Therapies Management Committee / Public Health Wales, 2017, p. 9).

A frequent criticism of mental health services has been the lack of accessibility to evidence-based psychological interventions (Turpin, Richards, Hope & Duffy, 2008). Low intensity, group CBT interventions have sought to address issues of cost, stigma and accessibility to psychological interventions (Lucas & Telch, 1993; White, 2008; Bennett-Levy, Richards, Farrand & Christensen, 2010).

SC has been shown to provide a clinically effective and organisationally efficient approach to treating common mental health problems (White, 2008). To enable uptake and reduce stigma, SC sessions are designed to be delivered in community settings (White, 2008). SC is an approved intervention at Tier 1 (Welsh Government, 2012). SC is also offered in Improving Access to Psychological Therapies (IAPT) low-intensity services in England.

White, Keenan and Brooks (1992) tested the efficacy of SC in a controlled trial. The results indicated that, post-intervention, participants showed significant reductions in symptoms compared to a wait-list control (WLC). A number of practice-based studies have reported that, on average, participants experience a 50% reduction in anxiety and depression following SC (e.g. Joice & Mercer, 2010; Wood *et al.*, 2005). Studies that have looked at long-term outcomes of SC have found that gains are maintained for up to twenty four months (White & Keenan-Ross, 1997; Kellett, Clarke & Matthews, 2007a; Van Daele, Van Audenhove, Vansteenwegen, Hermans & Van den Bergh, 2013; White, 1998). Kellett, Clarke & Matthews (2007b) benchmarked SC outcomes against individual CBT and interpersonal psychotherapy (IPT) and found similar results across all therapeutic modalities. SC users also report extremely high satisfaction rates (e.g. Houghton & Saxon, 2007; Kellett, Newman, Matthews & Swift, 2004; White, 1995).

SC has been running in Hywel Dda University Health Board (HDUHB) for a number of years. A historical decision was made to reduce the number of sessions from six to four in the hope that this would reduce the number of people who dropped out of the intervention prematurely. However, it has been suggested that delivering effective evidence-based interventions requires close adherence to protocols (Roth & Fonagy, 2004; Siev, Huppert & Chambless, 2009) and that differences in clinical outcomes between services may be caused by differing levels of treatment fidelity (Glover, Webb & Evison, 2010). Research suggests that this is the case with SC. Burns, Kellett and Donohoe (2016) reported a dose-response relationship between the number of SC sessions attended and the likelihood of improvement. Delgadillo *et al*. (2016) found that a five session SC course was significantly less effective than the full, six-session course.

Given the importance of treatment fidelity on the positive outcomes of evidence-based interventions, it was decided that the adapted version of SC being run in HDUHB should be evaluated. This evaluation addressed the following research questions: (1) How effective is the adapted version of SC being run in HDUHB? (2) Is clinical effectiveness influenced by attendance rates? (3) Are clinical outcomes influenced by initial symptom severity? (4) Does attendance at the SC course reduce risk to self? (5) What are participants' and facilitators' experiences of the four session version of SC?

Method

Intervention

Four SC groups were run across the HDUHB region between October 2019 and December 2019. Data was only collected in Carmarthenshire and Pembrokeshire because SC was not being run in Ceredigion at the time of this evaluation. Just over half of the data (57.8%) was collected from Carmarthenshire. Available demographic and outcome data were aggregated for all individuals who accessed SC in both counties for the analysis.

Participants attended SC through two routes (1) referred to Local Primary Care Mental Health Support Services (LPMHSS) by GPs or (2) self-referral through gaining knowledge of SC online, via posters, through word of mouth or through other services.

Each SC group was facilitated by members of LPMHSS staff and each session lasted for two hours. SC ran weekly over four sessions. Participants were not followed-up if they missed sessions and were not reviewed on completion of the course.

Design

A mixed methods design was employed. The strategy involved primary quantitative methods and secondary qualitative methods used to obtain the views of both course participants and facilitators[1]. The quantitative aspects of the analysis provided statistical and clinically relevant data regarding the outcomes of the intervention, while the qualitative aspects ensured that the complexity inherent in the experiences of participants and facilitators was not lost. Quantitative data was analysed using SPSS Statistics (trial version; IBM Corp., 2017).

A repeated measures design was used for the quantitative evaluation. Repeated measures ANOVAs were used to analyse changes in symptoms and overall distress levels (as indicated by participants' responses on psychometric questionnaires) for each of the four sessions. In addition, dependent t-tests on the results of psychometric measures were used to obtain overall pre- and post-intervention changes. Changes in risk (as indicated by participant responses on psychometric measures) were also analysed over each of the four sessions. Calculations of clinical significance and reliable recovery rates were also carried out. Descriptive statistics were produced for the demographic data.

Content analysis was employed to analyse the Initial and Evaluation Questionnaires returned by the participants. The broad categories of responses were refined to elucidate underlying meanings in an iterative and inductive process to generate higher-order themes.

Questionnaires

<u>Initial and Evaluation Questionnaires</u>

The Initial Questionnaire gathered demographic information, information related to the source of participants' recruitment and the goals that participants had at the start of the course. The Evaluation Questionnaire gathered feedback on participants' experience of attending SC. The Initial Questionnaire required categorical responses to be provided in checkboxes. The Evaluation Questionnaire posed open-ended questions on attendees' experiences of the intervention, alongside Likert scale questions to determine whether participants felt that the course had covered the topics that it was designed to.

[1] Due to time limitations, facilitator feedback was only obtained from LPMHSS staff in Carmarthenshire

CORE-10 (Connell & Barkham, 2007)

The CORE-10 is a brief self-report questionnaire comprising of ten items that are designed to measure an individual's level of global distress, including commonly experienced symptoms of anxiety and depression and aspects of life and social functioning. It is suitable for use as an initial screening tool and as an outcome measure. Each item within the CORE-10 is scored on a five-point Likert scale.

The CORE-10 distinguishes between clinical and non-clinical populations. It has good specificity; acceptability; sensitivity to change; and, very good internal reliability (Barkham *et al.*, 2013).

GAD-7 (Spitzer, Kroenke, Williams & Löwe, 2006)

The GAD-7 was developed as a self-report screening tool for GAD; however, it is now used with individuals with other anxiety disorders (Beard & Björgvinsson, 2014). The GAD-7 consists of seven items each on a four-point Likert scale.

The GAD-7 is considered to be a valid and reliable case identifier and outcome measure for anxiety (Spitzer *et al.*, 2006). Psychometric evaluations suggest that the GAD-7 is a reliable and valid measure of anxiety symptoms in psychiatric samples (Kertz, Bigda-Peyton & Bjorgvinsson, 2013) and in the general population (Löwe *et al.*, 2008). It has good sensitivity and specificity, excellent internal consistency and good test-retest reliability (Spitzer *et al.*, 2006).

PHQ-9 (Kroenke, Spitzer & Williams, 2001)

The PHQ-9 was originally developed as a self-report screening tool for depression in primary care. The PHQ-9 consists of nine items each on a four-point Likert scale.

The PHQ-9 is considered to be a valid and reliable case identifier and outcome measure for depression (Spitzer *et al.*, 2006). It has been reported to have good internal consistency, good sensitivity to change and good specificity (Beard, Hsu, Rifkin, Busch & Björgvinsson, 2016).

Reliable and Clinically Significant Outcomes

Reliable and clinically significant changes were used as measures for course effectiveness. Reliable change refers to the extent to which change falls beyond that which would occur due to the measurement variability of a psychometric instrument (Wise, 2004). The Reliable Change Index (RCI; Jacobson & Truax, 1991) specifies the amount of change a client must show on a specific psychometric measure for that change to be reliable. The RCI for the PHQ-9 is 9, for the GAD-7 the RCI is 6 and for the CORE-10 the RCI is 6.

Clinical significance analyses enable an understanding of the effectiveness of interventions on an individual level. This is in contrast to statistical significance which provides an understanding of the effectiveness of the intervention on a whole group of participants. Clinically significant change indicates whether an individual's symptoms have improved by moving from the clinical to non-clinical range (or vice versa for deterioration). The criteria for clinical change on the GAD-7 is a final score of 8 or less with pre-intervention score of 9 or more. For the PHQ-9, it is a final score of 10 or less with a pre-intervention score of 11 or more. For the CORE-10, it is a final score of 10 or below and a pre-intervention score of 11 or above.

Taken together, the RCI and the diagnostic cut-off score can help to define reliable and clinically significant improvement. This has been recommended as a robust method for assessing the degree of recovery following psychological interventions (Evans, Margison & Barkham, 1998; McMillan, Richards & Gilbody, 2010). It should be noted that, in primary care, not all clients will necessarily

meet criteria for a significant mental disorder; in fact, clinical guidelines advocate evidence-based psychological interventions for people with mild or sub-threshold depression and anxiety (NICE, 2011). Although only limited symptom reductions can be expected in clients with sub-clinical symptoms, it is possible that such clients may deteriorate and, therefore, it has been proposed that rates of reliable deterioration should also be reported (Delgadillo, McMillan, Leach, Lucock, Gilbody & Wood, 2014).

Focusing only on recovery fails to recognise significant symptom changes for clients who still meet criteria for common mental health problems but who may feel considerably better. Minimal clinically important differences (MCID) are client derived scores that reflect changes on psychometric measures that are meaningful for the client (Cook, 2008). The MCID for the GAD-7 is four points (Toussainta *et al*., 2020) and for the PHQ-9 it is five points (Löwe, Unützer, Callahan, Perkins & Kroenke, 2004). No MCID has been reported for the CORE-10.

The current evaluation follows the suggestion of Delgadillo *et al*. (2014) who proposed that a comprehensive investigation of outcomes should combine effect sizes (ES) and rates of reliable and clinically significant improvement (RCSI).

Benchmarking is the statistical comparison of results found in routine clinical outcomes against those of clinical trials (high efficacy benchmarks) or those observed in control groups (lower benchmarks for no significant treatment effects) (Lueger & Barkham, 2010; Minami *et al*., 2008). Following the benchmarking methodology suggested by Delgadillo *et al*. (2016), ES in this evaluation were compared to two benchmarks (pre-post Cohen's d). One benchmark was derived from the only controlled trial of SC (White *et al*., 1992) and the second benchmark was derived from a meta-analysis of guided self-help interventions for anxiety and depression (Coull & Morris, 2011). These benchmarks took the GAD-7 as the primary outcome measure (given the main focus on anxiety management in SC). These benchmarks are based on group outcomes.

Another benchmark will be used to look at individual outcomes. This benchmark comes from the IAPT guidelines which state that a minimum of 50% of clients should show clinically significant change (National Collaborating Centre for Mental Health, 2020). Gyani, Shafran, Layard & Clark (2013) proposed that this number should be based on "*reliable recovery*" i.e. individuals who scored above the clinical cut-off at initial assessment, showed reliable improvement (based on the RCI) during treatment and scored below the clinical cut-offs at the end of treatment. It was necessary to use an English benchmark as no Welsh benchmark figures were available.

Results

Participant Demographics

102 participants agreed to complete outcome measures over the course of the SC intervention. 65.4% of participants were female and 34.6% were male.

Age of Participants

The ages of the participants are shown in Figure 1 below.

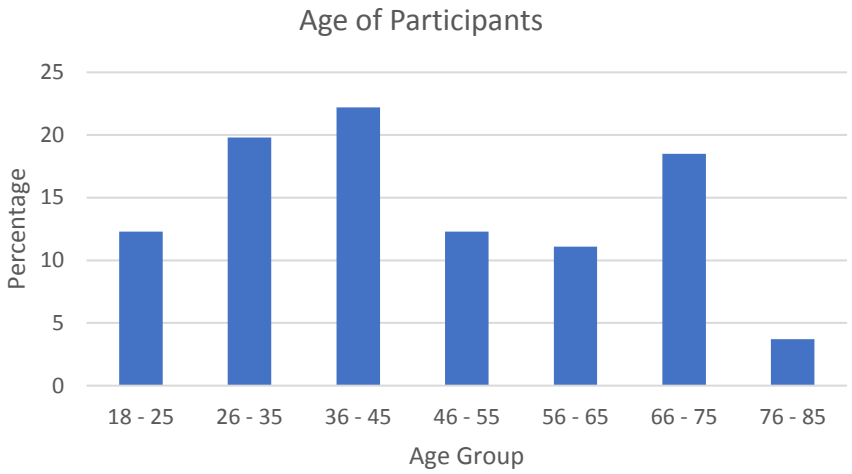

Figure 1: Age of participants who attended the Stress Control course

Ethnicity of Participants

The vast majority (96.3%) of participants were White-British. 2.5% of participants identified as White-Other and 1.3% of participants stated that they were Pakistani. This is in-line with population data which states that 1% of the population in the HDUHB area are from an ethnic minority background (NHS Wales, 2020).

Disability

13.8% of participants described themselves as having a disability and 86.2% did not.

Employment Status of Participants

The employment status of the participants can be seen in Figure 2.

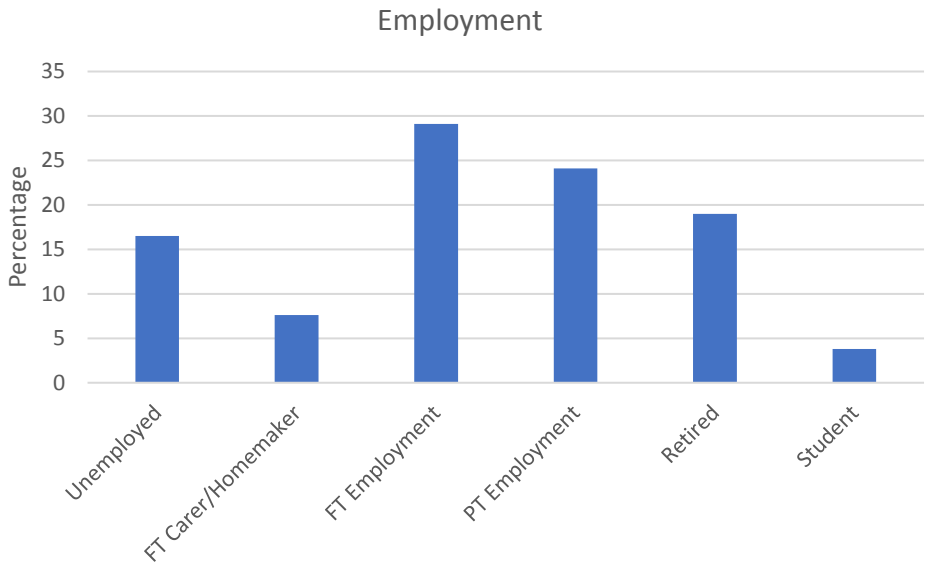

Figure 2: Employment status of participants who attended the Stress Control course

How Participants Heard about the Stress Control Course

The majority of participants (51.3%) heard about the Stress Control course from their GP. Figure 3 shows the breakdown of all the sources that provided participants with information about the course.

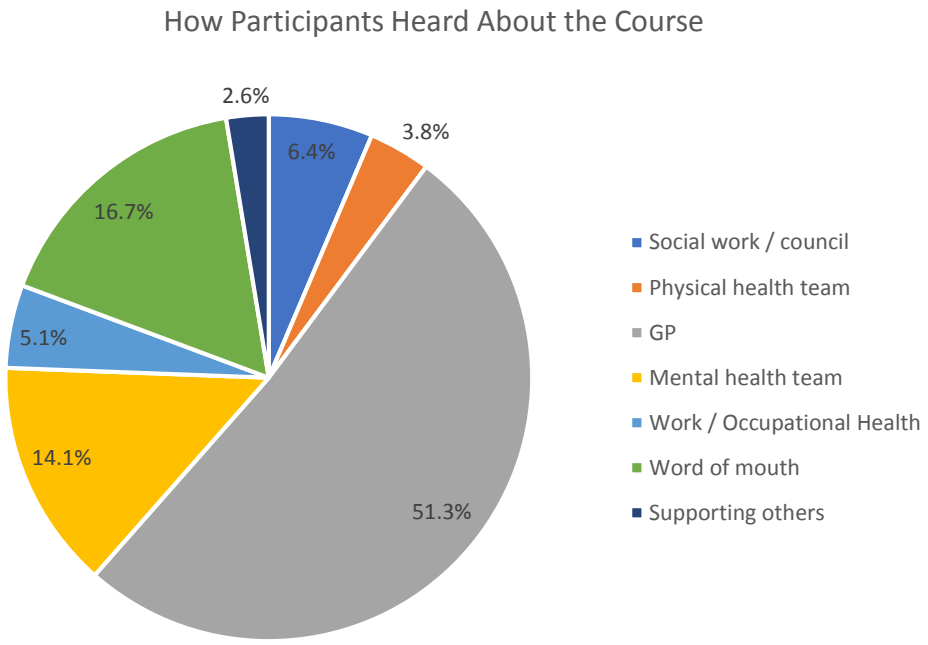

Figure 3: How participants heard about the Stress Control Course

Participants' Goals for the Course

Participants were asked to identify their goals for the course at the start of the first session. These are shown in Figure 4.

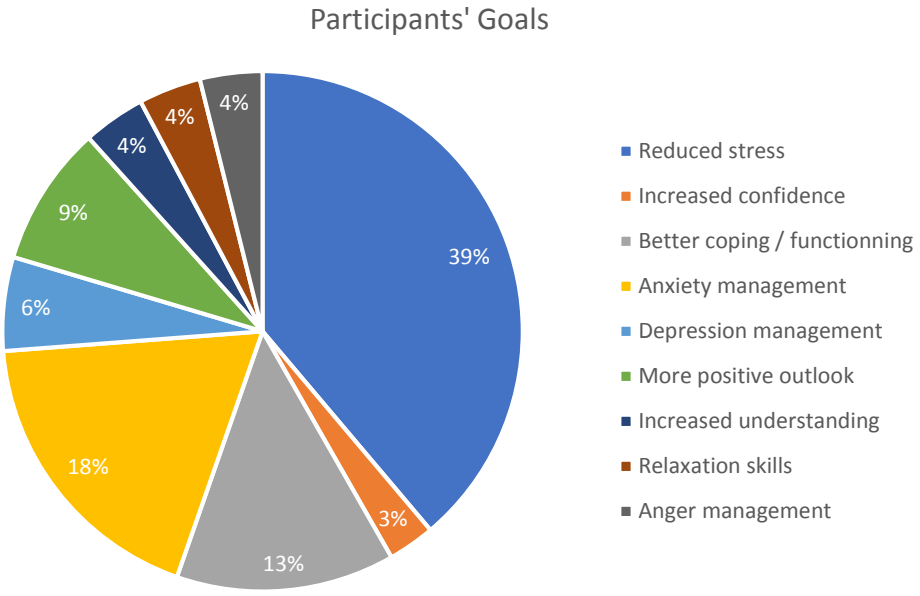

Figure 4: Participants' goals for the course

Attendance

Figure 5 provides information about the number of participants in each session. It should be noted that this is an approximation based on the number of questionnaires completed because it was not mandatory for participants to complete the questionnaires. 30.4% of participants attended all four sessions. A cut-off of attendance at three or more sessions was used to distinguish "*attenders*" from "*drop-outs*". This cut-off point has been reported in past studies of SC (e.g. Burns *et al*., 2016); however, other studies have suggested that attendance at least four sessions is necessary to receive an "*adequate dose*" of SC (e.g. Firth, Delgadillo, Kellett & Lucock, 2020). Based on the first of these cut-offs, 53.9% of participants were considered to be attenders and 46.1% were drop-outs.

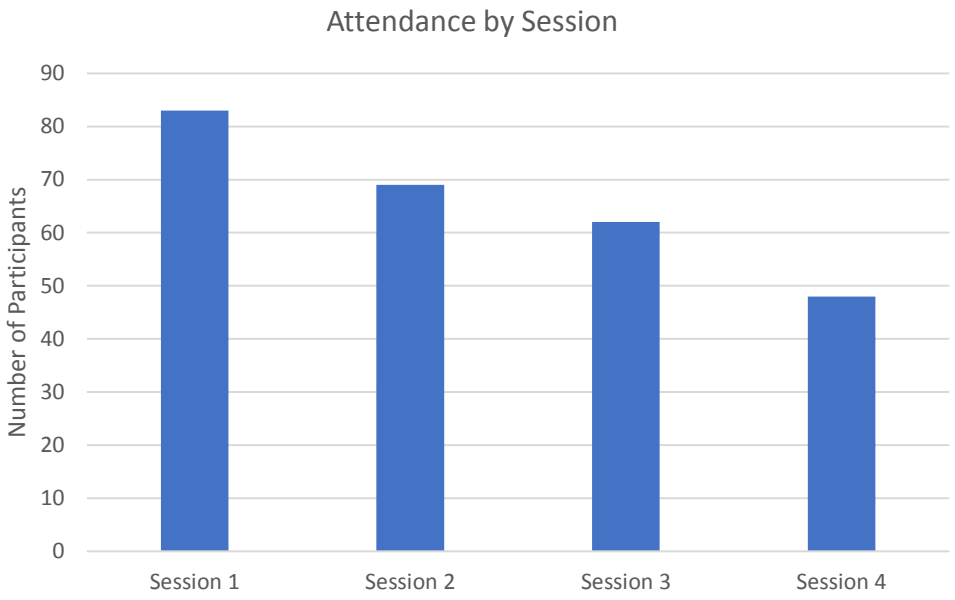

Figure 5: Number of participants in each Stress Control session

Pre-Intervention Symptoms

The breakdown of GAD-7, PHQ-9 and CORE-10 scores by severity range is shown in Table 1.

Questionnaire	Non-Clinical	Mild	Moderate	Moderate-Severe	Severe
GAD-7	12 (11.8%)	33 (32.4%)	18 (17.6%)	N/A	39 (38.2%)
PHQ-9	14 (13.7%)	19 (18.6%)	29 (28.4%)	25 (24.5%)	15 (14.7%)
CORE-10	16 (15.7%)	18 (17.6%)	29 (28.4%)	17 (16.7%)	22 (21.6%)

Table 1: Breakdown of GAD-7, PHQ-9 and CORE-10 scores by severity range

Table 2 indicates the presentation (anxiety, depression, co-morbid) for attenders and drop-outs. Participants scoring above clinical cut-offs on both the PHQ-9 and the GAD-7 were considered to have comorbid anxiety and depression. If a participant scored above clinical cut-off on the GAD-7 and not the PHQ-9, they were considered to have an anxiety disorder (and vice versa for the PHQ-9 and depression).

Attendance	Presentation		Mean GAD-7 Score	Mean PHQ-9 Score	Mean CORE-10 Score	Mean Risk CORE	Mean Risk PHQ-9
Attenders (≥3 sessions)	Anxiety	5 (9.0%)	11.38 (SD = 5.17)	12.53 (SD = 5.73)	19.18 (SD = 7.62)	0.53 (SD = 1.07)	0.51 (SD = 0.90)
	Depression	6 (10.9%)					
	Co-morbid	33 (60.0%)					
	None	11 (20.0%)					
Drop-outs (≤2 sessions)	Anxiety	5 (10.6%)	11.32 (SD = 5.86)	12.74 (SD = 7.34)	17.85 (SD = 7.16)	0.17 (SD = 0.52)	0.40 (SD = 0.71)
	Depression	2 (4.3%)					
	Co-morbid	28 (59.6%)					
	None	12 (25.5%)					
Total	Anxiety	10 (9.8%)	11.35 (SD = 5.47)	12.63 (SD = 6.49)	18.57 (SD = 7.40)	0.37 (SD = 0.88)	0.46 (SD = 0.82)
	Depression	8 (7.8%)					
	Co-morbid	61 (59.8%)					
	None	23 (22.5%)					

Table 2: Clinical presentation and risk for individuals who drop-out and those who attended three or more sessions

There was no significant difference in CORE-10 scores (t(100) = 0.469, ns), GAD-7 scores (t(100) = 0.954, ns) or PHQ-9 (t(100) = 0.865, ns) at pre-intervention between attenders and drop-outs.

Pre-Intervention Risk

There were no significant differences between attenders and drop-outs on PHQ-9 risk scores (t(93) = .526, ns) or CORE-10 risk scores (t(99) = .0481, ns).

Overall Group Outcomes

Table 3 shows the group outcomes for the all participants, attenders and drop-outs.

A repeated measures ANOVA was conducted to determine the impact of the number of sessions attended on the level of symptoms (divided into anxiety, depression and general distress). The main effect of number of sessions attended was significant: $F(3, 84) = 6.790$, $p < .001$. The main effect of symptom type (as determined by scores on each of the three questionnaires) was significant: $F(2, 56) = 106.635$, $p < .001$. The interaction between the number of sessions attended and symptom type was not significant: $F(6,168) = 0.989$, $p = .434$. Pairwise comparisons indicated that individuals who attended three or more sessions had significantly better outcomes than those who attended less than three sessions. In addition, outcomes on all three questionnaires were significantly different to each other, with the most change being found on the CORE-10 and the least change being found on the GAD-7. These results are displayed in Figure 6.

	N	Pre-Intervention Mean (SD)	Post-Intervention Mean (SD)	Pre-Post Change Mean (SD)	95% CI range	t	Cohen's d	Effect Size
Whole sample (attended 2 – 4 sessions)[2]								
PHQ-9	74	11.88 (5.50)	9.82 (5.79)	-3.23 (4.80)	1.00 – 3.17	3.893**	0.365	Small
GAD-7	74	11.83 (5.28)	7.76 (5.76)	-1.97 (4.36)	2.42 – 4.82	6.007**	0.656	Moderate
CORE-10	73	19.23 (6.87)	14.99 (6.49)	-4.52 (5.93)	2.84 – 5.65	6.029**	0.635	Moderate
Attended 3 – 4 sessions[1]								
PHQ-9	54	12.20 (5.39)	9.83 (5.27)	-2.35 (4.87)	1.03 – 2.71	3.544**	0.445	Small
GAD-7	54	11.52 (5.11)	7.85 (5.33)	-3.67 (5.28)	2.27 – 5.11	5.106**	0.703	Moderate
CORE-10	53	19.96 (7.14)	14.91 (6.64)	-5.06 (6.27)	3.33 – 6.78	5.876**	0.732	Moderate
Completed all 4 sessions[1]								
PHQ-9	30	13.03 (5.52)	9.83 (4.27)	-2.04 (4.50)	1.40 – 5.00	3.639**	0.390	Small
GAD-7	30	11.8 (5.16)	9.83 (4.27)	-3.65 (5.17)	.34 – 3.59	2.471*	0.416	Small
CORE-10	29	20.0 (6.70)	15.52 (5.65)	-4.26 (6.02)	2.26 – 6.77	4.101**	0.613	Moderate

* p < 0.05, ** p < 0.001

Table 3: Group outcomes on all measures for attenders and drop-outs

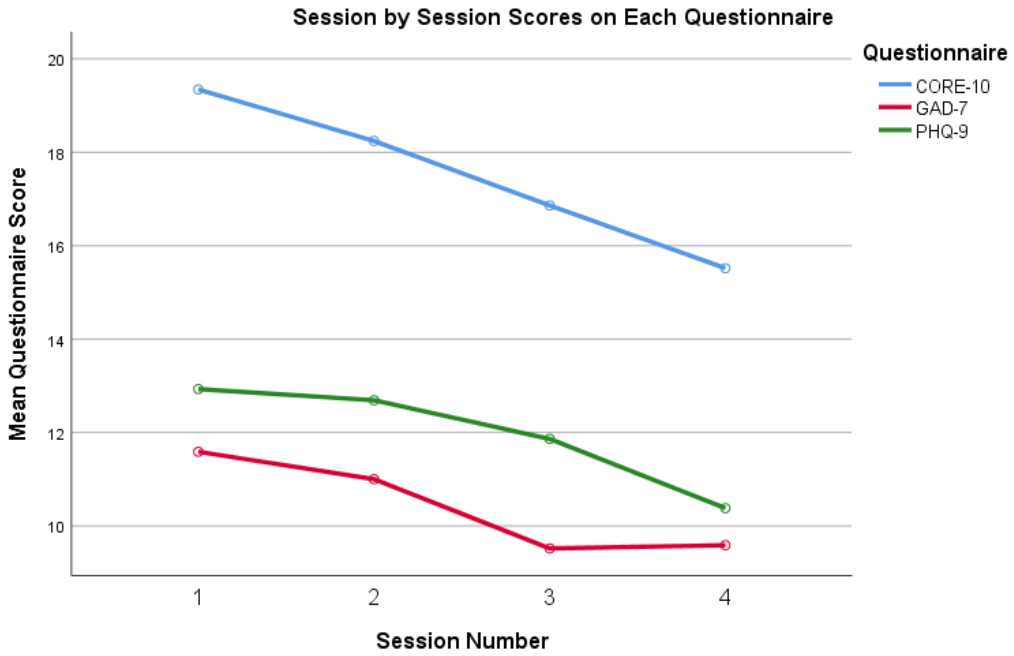

Figure 6: Session by session scores for each out the outcome measures

[2] Only individuals with pre and post intervention scores on all measures have been included in the analysis

Overall Individual Outcomes

Individual outcomes were calculated using a number of measures: clinically reliable change, clinically significant change and MCID. These results can be seen in Table 4.

	N	Positive Clinically Significant Change	Positive MCID	Reliable Recovery	Stasis	Reliable Deterioration	Negative MCID
	Whole sample (attended 2 – 4 sessions)[3]						
PHQ-9	74	18 (24.3%)	23 (31.1%)	10 (13.5%)	54 (73.0%)	2 (2.7%)	3 (4.1%)
GAD-7	74	31 (41.9%)	31 (41.9%)	20 (27.0%)	40 (54.1%)	2 (2.7%)	3 (4.1%)
CORE-10	73	27 (37.0%)	N/A	9 (12.3%)	44 (60.3%)	1 (1.4%)	N/A
	Attended 3 – 4 sessions[2]						
PHQ-9	54	16 (29.6%)	20 (37.0%)	10 (18.5%)	36 (66.7%)	2 (4.7%)	3 (5.6%)
GAD-7	54	24 (44.4%)	24 (44.4%)	15 (27.8%)	27 (50.0%)	2 (4.7%)	2 (4.7%)
CORE-10	53	23 (43.4%)	N/A	8 (15.1%)	29 (54.7%)	1 (1.9%)	N/A
	Completed all 4 sessions[2]						
PHQ-9	30	10 (33.3%)	13 (43.3%)	7 (23.3%)	19 (63.3%)	1 (3.3%)	1 (3.3%)
GAD-7	30	9 (30.0%)	9 (30.0%)	5 (16.7%)	19 (63.3%)	1 (3.3%)	1 (3.3%)
CORE-10	29	12 (41.4%)	N/A	4 (13.8%)	16 (55.2%)	1 (3.4%)	N/A

Table 4: Individual outcomes for all participants based on number of sessions attended

Benchmarking

The outcomes for the participants who completed all four sessions of SC were compared against two benchmarks from previous studies (one specifically for SC and the other for guided self-help interventions). Both of these comparisons indicated that the current evaluation showed a significantly smaller effect size on changes on anxiety symptoms as measured by the GAD-7. These results are shown in Figure 7.

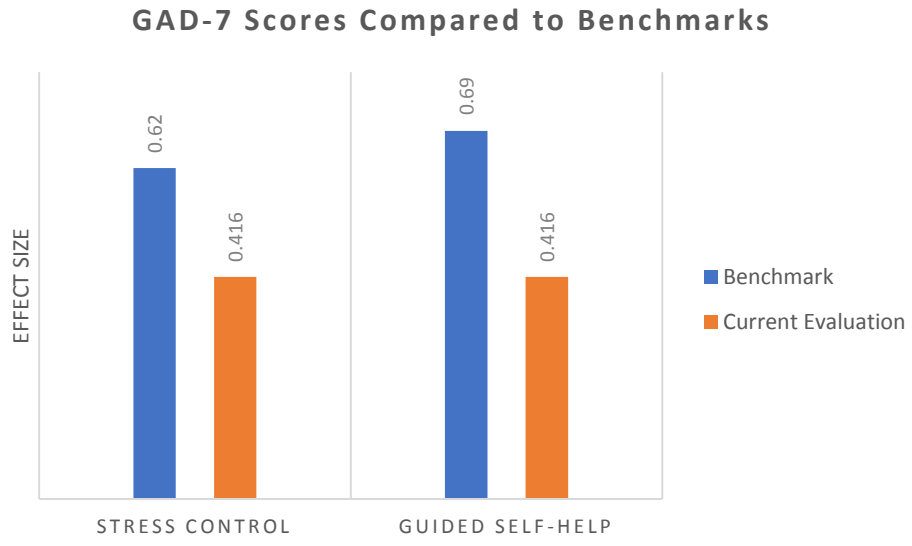

Figure 7: GAD-7 change effect sizes for participants who attended all four sessions compared to two benchmark effect size.

[3] Only individuals with pre and post intervention scores on each all measures have been included in the analysis

The benchmark provided by the English IAPT services is that 50% of clients should show reliable recovery. A person demonstrates reliable recovery if their symptoms were in the clinical range at the start of their treatment and below the clinical range at the end of their treatment and the degree of change is over the RCI. For the current evaluation, this figure is below 30% for all outcome measures regardless of the number of sessions attended by the participants (see Table 4).

The Impact of Initial Symptom Severity on Outcome

Group Outcomes by Initial Symptom Severity

Pre and post intervention scores were analysed to determine whether initial symptom severity had an impact in the effectiveness of the SC intervention for anxiety (as measured by GAD-7), depression (as measured by PHQ-9) and levels of general distress (as measured by CORE-10). This information is displayed in Table 5.

The results of dependant t-tests on the GAD-7 indicated that individuals who scored as severely anxious prior to intervention showed a significantly greater reduction in anxiety than those categorised with mild to moderate anxiety ($t(20) = 2.704$, $p<.05$).

The results of dependant t-tests on the PHQ-9 showed that individuals who scored as severely depressed prior to the intervention showed a significantly greater reduction in depression than those categorised with mild to moderate anxiety ($t(17) = 2.956$, $p<.05$).

The results of dependant t-tests on the CORE-10 showed that individuals whose symptoms were severe prior to the intervention showed a significantly greater reduction in symptoms than those categorised with mild to moderate symptoms ($t(31) = 4.359$, $p<.001$).

Severity	N	Pre-Intervention Mean (SD)	Post-Intervention Mean (SD)	Pre-Post Change Mean (SD)	95% CI range	t	Cohen's d	Effect Size
GAD-7								
Mild - Moderate	34	7.94 (3.17)	7.21 (3.82)	-1.53 (4.50)	- 0.82 – 2.29	0.961	0.16	Small
Severe	21	16.95 (1.66)	12.33 (4.16)	-4.62 (4.71)	2.48 – 6.76	4.50**	0.98	Large
PHQ-9								
Mild - Moderate	37	9.49 (4.15)	8.22 (5.43)	-1.35 (4.50)	-0.15 – 2.69	1.817	0.30	Small
Severe	18	18.78 (2.51)	14.22 (4.82)	-4.56 (4.03)	2.55 – 6.56	4.793**	1.13	Large
CORE-10								
Mild - Moderate	42	13.71 (4.54)	12.45 (5.69)	1.62 (5.51)	-0.46 – 2.98	1.485	0.23	Small
Severe	32	24.91 (3.96)	18.16 (6.05)	6.75 (6.11)	4.55 – 8.95	6.253**	1.11	Large

** $p<0.01$

Table 5: Post-intervention group outcomes by initial symptom severity

Individual Outcomes by Initial Symptom Severity

A variety of individual outcome measures were used to determine whether initial symptom severity had an impact in the effectiveness of the SC intervention for anxiety (as measured by GAD-7), depression (as measured by PHQ-9) and levels of general distress (as measured by CORE-10). This information is displayed in Table 6.

	N	Positive Clinically Significant Change	Positive MCID	Reliable Recovery	Stasis	Reliable Deterioration	Negative MCID
	GAD-7						
Mild - Moderate	34	7 (20.6%)	7 (20.6%)	4 (11.8%)	23 (67.6%)	4 (11.8%)	3 (8.8%)
Severe	21	10 (47.6%)	10 (47.6%)	3 (14.3%)	11 (52.4%)	0 (0.0%)	0 (0.0%)
	PHQ-9						
Mild - Moderate	37	4 (10.8%)	6 (16.2%)	4 (10.8%)	28 (75.7%)	2 (5.4%)	2 (5.4%)
Severe	18	10 (55.6%)	10 (55.6%)	3 (16.7%)	8 (44.4%)	0 (0.0%)	0 (0.0%)
	CORE-10						
Mild - Moderate	42	9 (21.4%)	N/A	6 (14.3%)	30 (71.4%)	2 (4.8%)	N/A
Severe	32	17 (53.1%)	N/A	2 (6.3%)	15 (48.9%)	0 (0.0%)	N/A

Table 6: Post-intervention individual outcomes by initial symptom severity

Chi-squared tests were used to determine whether reliable recovery rates were impacted by the participants' initial symptom severity.

Differences between reliable recovery rates for individuals in the severe anxiety category compared to those in the mild-moderate anxiety category were not statistically significant ($\chi^2(1) = 0.074$, p= .785). See Figure 8.

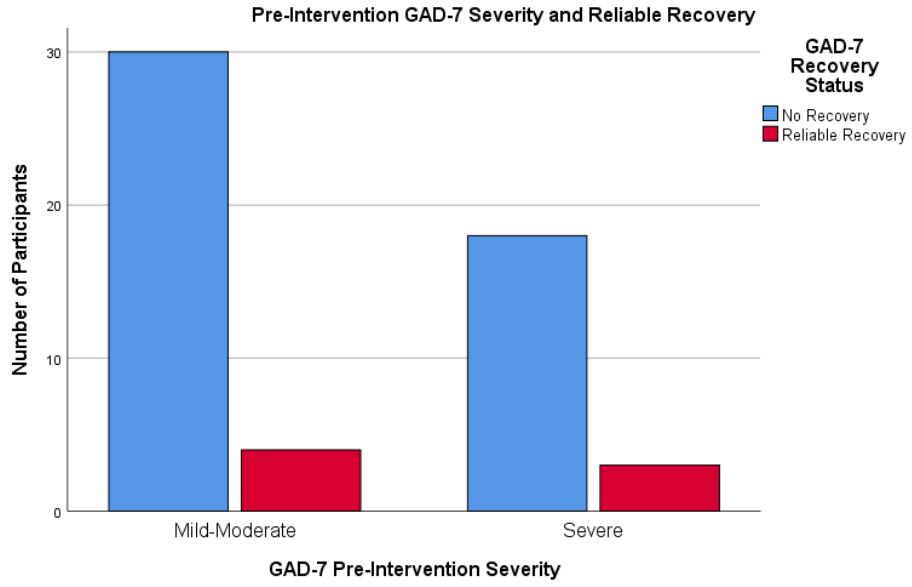

Figure 8: Number of participants who demonstrated reliable recovery as calculated by initial symptom severity on the GAD-7

Differences between reliable recovery rates for individuals with severe depression were not statistically different to those for individuals in the mild-moderate depression category ($\chi^2(1) = 0.374$, $p = .541$). See Figure 9.

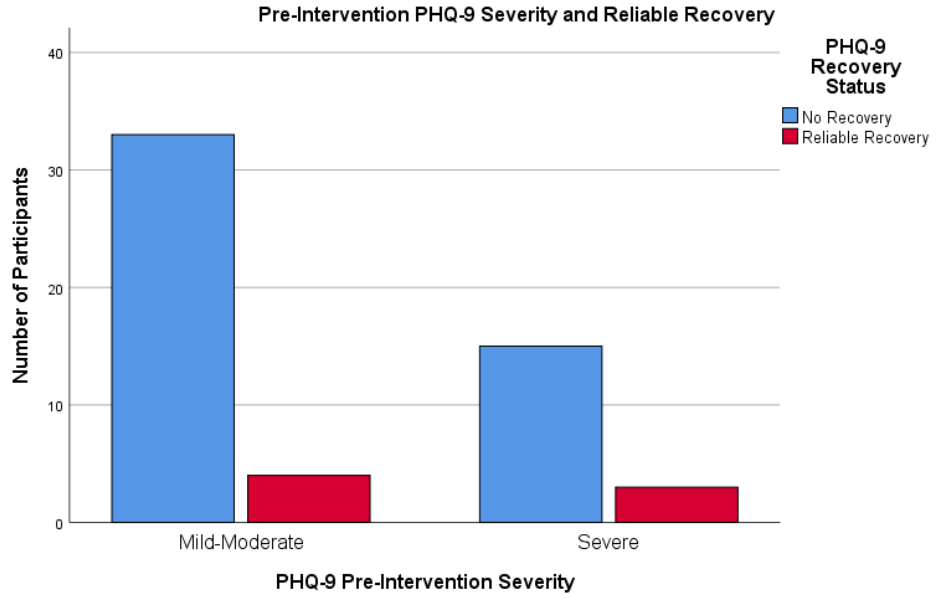

Figure 9: Number of participants who demonstrated reliable recovery as calculated by initial symptom severity on the PHQ-9

In terms of general psychological distress as measured by the CORE-10, reliable recovery rates between individuals in the severe category and those in the mild-moderate category were not statistically significant ($\chi^2(1) = 1.216$, $p = .270$). See Figure 10.

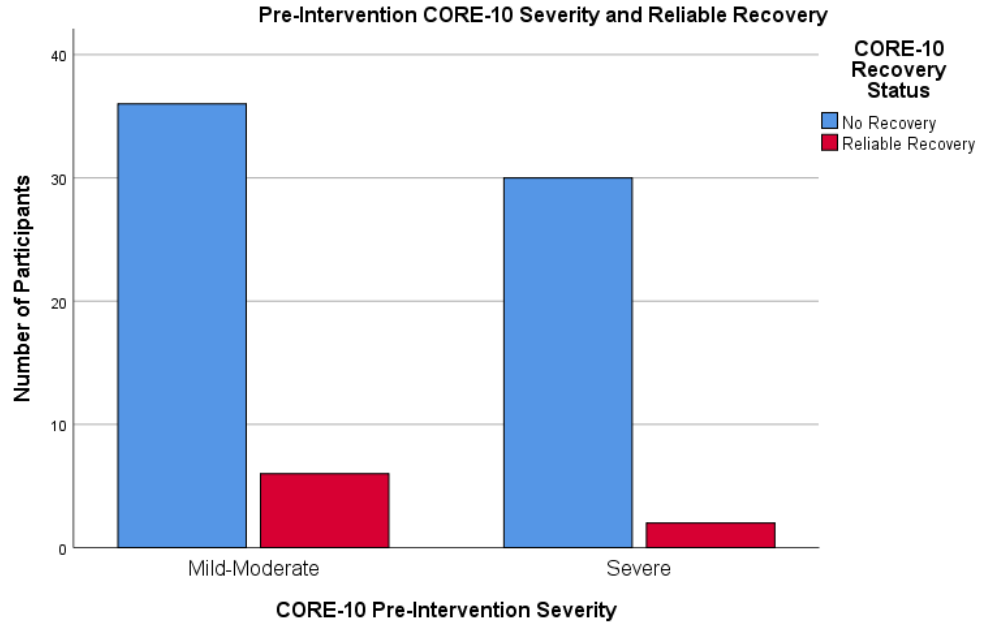

Figure 10: Number of participants who demonstrated reliable recovery as calculated by initial symptom severity on the CORE-10

Risk Outcomes

Risk was measured on a session by session basis using the risk questions on the CORE-10 and the PHQ-9. Table 7 shows the change in risk rating over time.

		Session 1	Session 2	Session 3	Session 4
CORE-10	N	82	70	59	50
	Range	0 - 4	0 – 3	0 – 2	0 – 2
	Mean (SD)	0.33 (0.832)	0.40 (0.824)	0.24 (0.567)	0.14 (0.405)
PHQ-9	N	76	70	59	51
	Range	0 - 4	0 – 4	0 – 3	0 – 3
	Mean (SD)	0.38 (0.730)	0.60 (0.984)	0.51 (0.796)	0.43 (0.728)

Table 7: Risk on session by session basis

A repeated measures ANOVA was conducted on the change in risk outcomes as measured by both the CORE-10 and the PHQ-9 over the four sessions of SC. The main effect of number of sessions was not significant: $F(3, 84) = 1.493$, $p = 0.222$. The main effect of type of questionnaire was significant: $F(1, 28) = 7.212$, $p < .05$. The interaction between number of sessions attended and type of questionnaire was not significant: $F(3, 84) = 2.157$, $p = 0.099$. Pairwise comparisons indicated that risk scores as measured by the CORE-10 decreased significantly compared to risk scores as measured by the PHQ-9. These results are displayed in Figure 11.

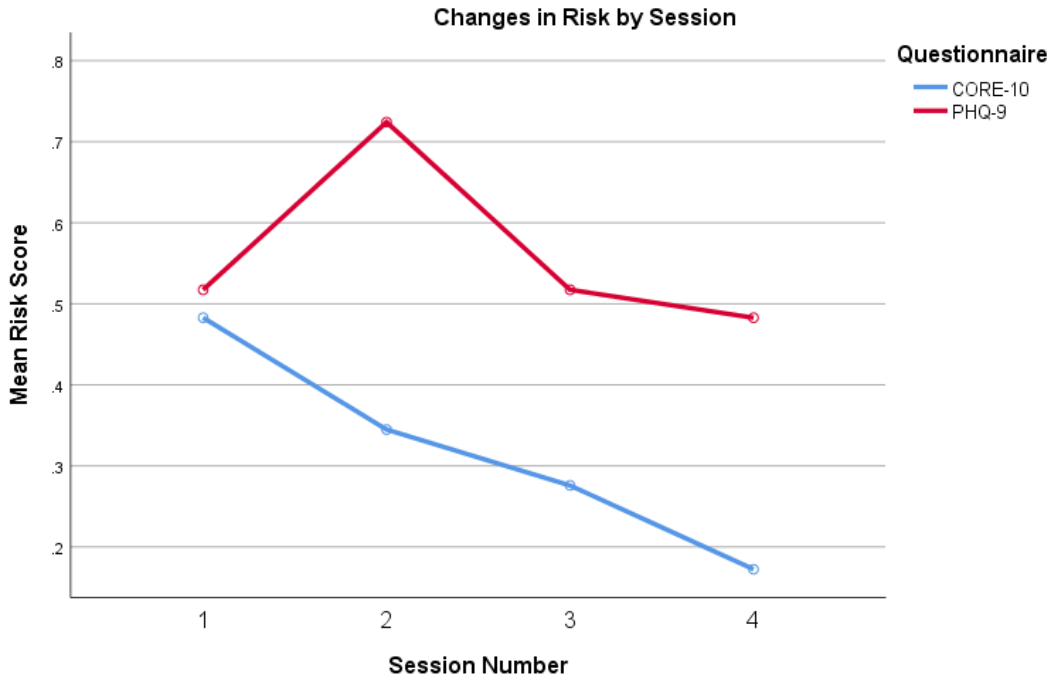

Figure 11: Changes in risk by session based on CORE-10 and PHQ-9 risk items

Feedback from Participants

100% of attendees who responded stated that they would recommend the SC course to others.

Participants were asked to rate how well they felt the course covered the main factors that SC is designed to include on a 5-point Likert scale (from "*Not at all*" to "*A lot*"). The results are shown in Table 8.

Course Focus	Not at all	Not much	Somewhat	Quite a bit	A lot
Understanding stress	0.0%	0.0%	20.4%	46.9%	32.7%
Understanding factors that maintain stress	0.0%	0.0%	16.3%	42.9%	40.8%
Developing skills to reduce stress	0.0%	2.0%	36.7%	38.8%	22.4%
Understanding how to manage panic	2.0%	4.1%	34.7%	34.7%	24.5%
Understanding how to manage low mood	0.0%	12.2%	38.8%	34.7%	14.3%
Relapse prevention	0.0%	10.2%	34.7%	38.8%	16.3%

Table 8: Participants' feedback regarding the coverage of the main aspects of the SC course

Participants were also asked to indicate how useful they found the course overall. Of those that responded, 44.9% stated that the course was "*Fairly Useful*" and 55.1% said that they found the course "*Very Useful*".

The primary positive themes that came from participant feedback on the Evaluation Questionnaire were:

- Expressed gratitude for the course and the facilitators.
    - *"Very friendly people"*
    - *"Excellent speakers, easy to listen to and explained everything well"*
    - *"Thank you. It's been so helpful"*
    - *"[Facilitators] very understanding and course well-presented"*
    - *"Great course that helps put things into perspective"*
    - *"Well presented. Made us feel at ease"*

- The positive impact of psychoeducation.
    - *"I liked that it talked a lot about panic and explained what happens in your body when you get stressed"*
    - *"It's important to understand what stress is"*
    - *"It's good to recognise stress and depression"*
    - *"I recognise that the symptoms I have experienced are stress-related and understand how to combat these through a new approach"*

- The usefulness of the handouts/CD.
    - *"Handouts have been very useful"*
    - *"Having handouts to take home with you [is good]"*
    - *"The booklets to refer back to [are good]"*

- A general positive impact of attending the course on quality of life.

    - *"It has given me a purpose to be somewhere and learn"*
    - "*Very good for confidence"*
    - *"Finding the difference between stuff I can control and stuff I can't"*
    - *"The course has been very helpful and has helped me understand stress in my own life and recognise it in others and respond to them differently - not take things personally"*
    - *"Understanding that things might not be as bad as I think"*
    - *"Just the explanation of it all made me feel better"*
    - *"The stress, anxiety and moods got less"*
    - *"A better way to look at problems"*

The primary negative themes that came from participant feedback on the Evaluation Questionnaire were:

- The lack of audience participation/interaction.

    - *"I would have liked to hear from others, just to know it isn't only me"*
    - *"Trouble concentrating on the presentation as there was no interaction during the sessions"*
    - *"Maybe let us talk more about our problems"*

- Practical issues regarding the venue.

    - *"Betters chairs!"*
    - *"Long gaps - there were tea breaks but there was no tea!"*

- Not enough practical skills being taught.

    - *"Repetitive about the feelings during stressed times. I know how I feel stressed, I needed a recommendation of what to do in the feelings and moments"*
    - *"It would be good to do some mindfulness / breathing during the session"*

- Not enough detail provided on some of the topics.

    - *"Some things were talked about too briefly"*

<u>Feedback from Facilitators</u>

The facilitators were positive about the idea of delivering psychoeducational courses in general, which they felt were beneficial in reducing the LPMHSS waiting list and referrals to other mental health services. However, the feedback pertaining to the SC course was that it felt *"disjointed"* and that there was *"no flow"*. They attributed this to the fact that the course had been reduced to four sessions whilst still attempting to cover all the topics from the six session version.

The facilitators felt that the slides were poor (again as a result of the shortened version of the course) and that, at times, this led to the participants having difficulties understanding the material that was being presented. This led to the facilitators *"not [always] believing in what [they] were delivering".* It was felt that this had a negative impact on staff motivation and morale and their ability to "*sell*" the content of the course to the participants. This was exacerbated by the fact that the facilitators were aware that they were providing an intervention that was not evidence-based and, therefore, they did not know whether the course was effective.

The facilitators explained that there were no scripts and that they were talking around the material that was presented on the slides. This was viewed as being problematic in light of the shortened course, as it was felt that the knowledge of what the slides meant may have been lost in the time since the course was adapted.

LPMHSS staff were very keen to return to a six session SC format. They also suggested that the first session (which is focused on introducing the course) could be used as a one-off taster session to improve participation at the full course. It was felt that this would enable people to have a better sense of what was being offered and how the course was delivered. They also suggested that the full-length course would allow time for a greater focus on signposting which could reduce the amount of people that went on to access further support from LPMHSS or other mental health services in HDUHB. It was also felt that it may be beneficial to invite guest speakers (e.g. pharmacists to talk about the advantages and disadvantages of medication).

The facilitators expressed concern about the limitations of the potential venues for psychoeducational courses because it was necessary to find venues that were free or available for a nominal fee. They explained that this had an impact on the size of the groups that could be invited to the sessions (particularly given that participants are able to bring friends/family members with them for support). The limitations in venues also presented problems in terms of accessibility with parking being limited for some settings. For example, some potential participants for the Carmarthen course (based in Glangwili General Hospital) decided not to attend as they were concerned about the lack of parking at the hospital. Staff also felt that holding the courses on general or psychiatric hospital sites created a barrier to attendance for some participants because they felt that their situation was being "*medicalised*" and made them more concerned about stigma related to being viewed as having a mental health problem. Finally, staff felt that the venues that were available made it difficult to provide adequate surroundings for participants including a lack of refreshments and chairs that were uncomfortable and not suitable for individuals with physical/mobility difficulties.

Discussion

Summary of Results

The overall results indicated that there was a statistically significant impact of the intervention on all measures; however, the current SC course in HDUHB was found to be significantly less effective than other studies against which these results were benchmarked: the developers of the SC model (White *et al*., 1992) and other controlled trials of guided self-help for anxiety symptoms (Coull & Morris, 2011).

When looking at the data on an individual level, the majority of participants did not show any reliable recovery on any of the outcome measures used, with reliable recovery rates for attenders being between 13.8% and 27.8%. This is significantly below the benchmark figure of 50% from IAPT England. In addition, the majority of participants fell in the stasis (no change) category. It should be noted that only a very small number of participants showed any clinical deterioration, suggesting that the intervention did not cause any harm to the participants.

These findings underline the importance of maintaining fidelity to the evidence base when interventions are disseminated into routine care. Meyers, Durlack and Wandersman (2012) proposed that the process of implementation of interventions requires an explicit assessment of how innovations may need to be adapted to a specific practice setting, coupled with a process of evaluation and the establishment of feedback mechanisms. Ideally, services adopting (and adapting) any evidenced-based interventions should establish a data-based feedback and clinical audit cycle as part of their implementation plans.

The results of this evaluation indicated that risk as measured by the CORE-10 was reduced over the course of attendance at the SC course, but this was not the case for risk as measured by the PHQ-9. This is likely to be due to the way in which the risk items are worded on each of the measures. The CORE-10 risk item is "*I have made plans to end my life*" and the PHQ-9 risk item is *"[How often have you had] thoughts that you would be better off dead, or of hurting yourself in some way?".* This suggests that participation in SC reduces the likelihood that an individual will make specific plans to harm themselves but does not have an impact on the degree of thoughts an individual has about suicide/self-harm.

The attrition rate of participants in the study was high, with only 30.4% of participants attending all four sessions of the course. This suggests that reducing the length of the SC course from six to four sessions has not resulted in the decrease in attrition rate that was originally hoped for. Attrition rates and missed sessions for group interventions are particularly important because sessions that are missed are essentially "*lost forever*". In contrast, individual interventions can be more flexible in the pace at which material is presented (Kellet *et al*., 2007a).

The qualitative feedback from participants indicated that 100% would recommend the SC course to others. Furthermore, 44.9% of participants found the course "*Fairly Useful*" and 55.1% found it "*Very Useful*". Participants felt that the course provided good information regarding psychoeducation about stress and factors that maintained stress, but that it provided less information about practical skills to reduce stress, panic attacks, low mood and relapse prevention. The lack of information regarding relapse prevention is likely to have a negative impact on the longer-term outcomes of participants attending the course and on re-referral rates.

The data collected from the open-ended questions on the Evaluation Questionnaire highlighted four positive themes: expressed gratitude for the course and the facilitators; a positive impact on quality of life; a positive impact of psychoeducation; and, the usefulness of the materials that were given to participants. The four negative themes that were identified were: the lack of audience participation / interaction; problems with the facilities in the venues; a lack of practical skills being taught / practiced; and, not enough detail being provided on some topics.

The qualitative feedback from facilitators indicated that they were positive about offering psychoeducational courses in general but that they had specific concerns about the version of the SC course currently being run in HDUHB. The facilitators were concerned about offering an intervention that was not evidence-based and they indicated that they did not have confidence in the material because it felt disjointed. This had a negative impact on the motivation and morale of staff asked to present the SC course. The lack of a script meant that facilitators were expected to talk around the slides, but it was felt that the slides were not always clear as a result of the shortening of the course. Concerns were also expressed about the difficulties in finding suitable cheap or free venues for the SC course. They felt that the venues that were used were not conducive to providing an equitable service because some did not provide facilities that were suitable for individuals with physical health / mobility difficulties and were uncomfortable for all participants. The location of the venues used also prevented some people from accessing the course due to parking limitations and/or went against the ethos of the course by "*medicalising*" the participants.

<u>Limitations of the Evaluation</u>

Participants were not selected randomly and so caution in the interpretation of the results is necessary. Furthermore, the Evaluation Questionnaire was only completed by those individuals who attended the final session which, on the whole, were individuals considered to be attenders. This may

have introduced a degree of bias into the feedback obtained and provides little indication of why individuals dropped out of the course.

The design of this study did not provide an opportunity for follow-up data to be collected so it is not possible to know whether any gains were maintained over the medium to long-term.

The pre-post treatment effect sizes described in this study offer a general estimate of the "*real world*" effectiveness of SC interventions delivered in HDUHB. Results were not analysed relative to a control group, therefore, it is possible that regression to the mean (i.e. natural fluctuations in mental health symptoms due to the passage of time) may have partly accounted for some of the reported effects. Furthermore, the data do not disentangle specific SC treatment effects from effects that may be due to general contact with healthcare practitioners and other participants in a group-based setting.

Conclusions

The lack of clinically significant positive outcomes for the vast majority of participants, along with the poor performance compared to benchmarked studies, suggests that the course content and delivery need to be addressed. Outcomes are likely to be improved by returning to the original six session format and utilising the newest version of the SC course.

Analysis of the impact of attendance showed that outcomes were statistically more positive when clients attended three or more sessions. This illustrates the importance of supporting participants to attend as many of the sessions as possible.

It was found that the outcomes measured by the three psychometric questionnaires were significantly different. The largest improvement was measured on the CORE-10. The CORE-10 provides a measure of psychological distress associated with mental health difficulties; whereas the GAD-7 and PHQ-9 measure changes in symptoms. This may indicate that attendance at the SC course, as delivered in HDUHB, reduced the amount of distress that participants felt about their symptoms more than it reduced the symptoms themselves. The reduced psychological distress may be due to the normalising effect of attending a large group intervention. Kellett, Clarke and Matthews (2006) suggested that "*normalisation*" was an important component in SC. The results of the GAD-7 and PHQ-9 suggested that the SC course had a significantly bigger impact on depression symptoms than anxiety symptoms. This is despite the fact that the majority of the content of SC is focused on managing anxiety rather than low mood (Kellett *et al*., 2007b) and is in contrast to findings in published evaluations of SC (e.g. Burns *et al.,* 2016). This result may be a consequence of the shortened version of the SC course being run in HDUHB and the "*lack of detail*" on some topics reported by participants. Alternatively, it may be that the reduction in psychological distress (as indicated by the CORE-10) achieved by attending the group had a positive impact on PHQ-9 scores.

The separate analyses for the severity of symptoms and general psychological distress prior to the intervention showed that individuals in the severe groups had statistically better outcomes than those mild-moderate groups. This difference was not replicated when looking at individual outcomes where no significant differences were found between the number of participants with mild-moderate symptoms/distress and those with severe symptoms/distress showing reliable recovery. This shows the importance of understanding outcomes at both group and individual levels as well as the utility of including measures of reliable change when evaluating psychological interventions.

This evaluation showed that participants had high levels of satisfaction with the course. Similar results have been found in previously published studies (e.g. Kellett *et al*., 2004). Attrition rates for SC in this

evaluation were higher than those reported in other evaluations of SC (e.g. Degadillo *et al*., 2016) and psychoeducational CBT groups in general (e.g. Butcher & de Clive-Lowe, 1985).

The facilitators of the SC course had significant concerns about having to provide an intervention that was not evidence-based and highlighted the negative impact that this had on their confidence to present the course material. Maintaining the motivation and morale of the staff facilitating the SC course and providing them with up to date training is vital because it has been shown that the characteristics and training of group facilitators determine their capacity to lead a group. A credible, likeable facilitator with relevant expertise, good interpersonal skills and with whom members can identify, is most likely to be effective in promoting personal change in groups (Borek & Abraham, 2018). Delgadillo *et al*. (2016) reported that differences between the ways in which groups are delivered explained up to 3.6% of variance in outcomes in SC, highlighting the importance of adhering to intervention protocols. Furthermore, it has been suggested that variability in facilitators' competence may partly explain differences in outcomes of group interventions (Burlingame, Strauss & Joyce, 2013), emphasising the need for group facilitators to be offered regular training.

<u>Recommendations</u>

- HDUHB should return to the six-session version of SC and provide the most up-to-date version of the course. These changes will ensure that the intervention provided is evidence-based.
- LPMHSS staff should be provided with training on facilitating SC as this would increase their confidence and morale and would also be in-line with the Matrics Cymru recommendation that staff should be properly trained in the interventions that they provide. Staff training will also ensure consistency of delivery across sites in the HDUHB area.
- Strategies to maintain engagement with clients at risk of dropping out of SC need to be developed and evaluated. LPMHSS staff have suggested that the first session of SC could be used as a stand-alone taster session. This could include discussing findings from previous studies which highlight the importance of session attendance.
- SC was designed for people with mild to moderate mental health problems. Kellett *et al*. (2004) have shown that providing SC to individuals with mild to moderate difficulties improves outcomes. In light of the fact that participants can self-refer to SC, it may be useful to offer an option of self-assessment or in-person assessment to ensure that the intervention is only being offered to those with mild to moderate symptoms and not those with more severe symptoms or higher levels of risk.
- SC is not currently being offered in Welsh, which goes against the Welsh Language Standards (Welsh Government, 2018). Whilst it is recognised that the current staff mix in LPMHSS means that it is difficult for SC to be delivered in Welsh, creative solutions should be developed to address this difficulty. For example, the use of interpreters; translating course materials into Welsh or providing a video-recorded version of the course in Welsh that could be provided to Welsh-speaking participants to view in their own time (subject to copyright).
- Consideration should be given to allowing more expensive community venues to be used for the delivery of SC. This would be in-line with the ethos of the course and, given the positive impact that high quality delivery of SC could have on the LPMHSS waiting list and referrals to other mental health services in HDUHB, the increase in cost is likely to be minimal or non-existent in the long-term.
- SC is not currently being run in Ceredigion due to staffing issues. This results in the services not being equitable across the HDUHB area. Creative solutions need to be found for this. For example, using video conferencing technology (e.g. Microsoft Teams) to allow courses that are being delivered in one county to be broadcast to groups of participants in other counties. In addition to addressing the current inequalities between counties, this would also increase

the reach of the intervention because participants would not be limited to the dates/times when the group is being run in a particular county/area.

- SC should be routinely evaluated on a session by session basis to ensure that the effectiveness of the intervention is maintained over time and that outcomes are in line with those expected from SC.
- SC participants should be offered the option of a telephone review following the intervention to determine whether they need any extra support. It is suggested that this is offered one to two months after the end of the SC course in order to allow participants the time to consolidate what they have learnt during the intervention before determining whether they require further interventions. It is recommended that this should be offered to all participants even if they dropped-out of the intervention. The information gathered at these reviews could be used to determine the medium-term effects of SC; gain a better understanding of reasons why participants do not attend the full SC course; and, monitor the proportion of participants that require further intervention. This review should be standardised and information recorded on a central database to enable comparisons between counties and to facilitate routine audits.

Acknowledgements

The author would like to express gratitude to all practitioners within LPMHSS services in HDUHB and to the SC participants whose engagement and participation made this evaluation possible.

References

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th Ed., Text Revision).

Bandelow, B. & Michaelis, S. (2015). Epidemiology of anxiety disorders in the 21st century. *Dialogues in Clinical Neuroscience*, *17*(3), 327 - 335.

Barkham, M., Bewick, B., Mullin, T., Gilbody, S., Connell, J., Cahill, J., ... & Evans, C. (2013). The CORE-10: A short measure of psychological distress for routine use in the psychological therapies. *Counselling and Psychotherapy Research*, *13*(1), 3 - 13.

Beard, C. & Björgvinsson, T. (2014). Beyond generalized anxiety disorder: Psychometric properties of the GAD-7 in a heterogeneous psychiatric sample. *Journal of Anxiety Disorders*, *28*, 547 - 552.

Beard, C., Hsu, K., Rifkin, L., Busch, A. & Björgvinsson, T. (2016)., Validation of the PHQ-9 in a psychiatric sample. *Journal of Affective Disorders*, *193*, 267 - 273.

Bennett-Levy, J., Richards, D., Farrand, P. & Christensen, H. (2010). *Oxford Guide to Low Intensity CBT Interventions*. Oxford: Oxford University Press.

Borek, A.J. & Abraham, C. (2018). How do small groups promote behaviour change? An integrative conceptual review of explanatory mechanisms. *Applied Psychological Health Well-Being*, *10*, 30 - 61.

Burlingame, G.M., Strauss, B. & Joyce, A.S. (2013). Change mechanisms and effectiveness of small group treatments. In: M.J. Lambert (Ed.), *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (6th ed.). New Jersey: John Wiley & Sons Inc.

Burns, P., Kellett, S. & Donohoe, G. (2016). "Stress Control" as a large group psychoeducational intervention at step 2 of IAPT services: Acceptability of the approach and moderators of effectiveness. *Behavioural and Cognitive Psychotherapy*, *44*, 431 - 443.

Butcher, P. & de Clive-Lowe, S. (1985). Strategies for living: Teaching psychological self-help as adult education. *British Journal of Medical Psychology*, *58*, 275 - 283.

Carter, R.M., Wittchen, H.U., Pfister, H. & Kessler, R.C. (2001). One-year prevalence of subthreshold and threshold DSM-IV generalized anxiety disorder in a nationally representative sample. *Depression and Anxiety*, *13*, 78 - 88.

Clark, L.A. & Watson, D. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology*, *100*, 316 - 336.

Connell, J. & Barkham, M. (2007). *CORE-10 User Manual, Version 1.1.* CORE System Trust & CORE Information Management Systems Ltd.

Cook, C.E. (2008). Clinimetrics Corner: The Minimal Clinically Important Change score (MCID): A necessary pretense. *Journal of Manual and Manipulative Therapy, 16*(4), E82 - E83.

Coull, G. & Morris, P.G. (2011). The clinical effectiveness of CBT-based guided self-help interventions for anxiety and depressive disorders: A systematic review. *Psychological Medicine*, *41*, 2239 - 2252.

Delgadillo, J., Kellett, S., Ali, S., McMillan, D., Barkham, M., Saxon, D., … Lucock, M. (2016). A multi-service practice research network study of large group psychoeducational cognitive behavioural therapy. *Behaviour Research and Therapy*, *87*, 155 - 161.

Delgadillo, J., McMillan, D., Leach, C., Lucock, M., Gilbody, S. & Wood, N. (2014). Benchmarking routine psychological services: a discussion of challenges and methods. *Behavioural and Cognitive Psychotherapy*, *42*(1), 16 - 30.

Evans, C., Margison, F. & Barkham, M. (1998). The contribution of reliable and clinically significant change methods to evidenced-based mental health. *Evidenced Based Mental Health*, *1*, 70 - 72.

Firth, N., Barkham, M. & Kellett, S. (2015). The clinical effectiveness of stepped care systems for depression in working age adults: A systematic review. *Journal of Affective Disorders, 170*, 119 - 130.

Firth, N., Delgadillo, J., Kellett, S. & Lucock, M. (2020). The influence of socio-demographic similarity and difference on adequate attendance of group psychoeducational cognitive behavioural therapy. *Psychotherapy Research*, *30*(3), 362 - 374.

Glover, G., Webb, M. & Evison, F. (2010). *Improving access to psychological therapies: A review of the progress made by sites in the first rollout year. Discussion Paper*. North East Public Health Observatory, Stockton on Tees.

Gyani, A., Shafran, R., Layard, R. & Clark, D.M. (2013). Enhancing recovery rates: Lessons from year one of IAPT. *Behaviour Research and Therapy*, *51*, 597 - 606.

Houghton, S. & Saxon, D. (2007). An evaluation of large group CBT psycho-education for anxiety disorders delivered in routine practice. *Patient Education and Counselling*, *68*, 107 - 110.

IBM Corp. (2017). IBM SPSS Statistics for Windows, Version 25.0 (trial version). Armonk, NY: IBM Corp.

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

Joice, A. & Mercer, S. (2010). An evaluation of the impact of a large group psychoeducational programme (Stress Control) on patient outcome: Does empathy make a difference? *The Cognitive Behaviour Therapist*, *3*, 1 - 17.

Kellett, S., Clarke, S. & Matthews, L. (2006). Session impact and outcome in group psycho-educative cognitive behavioural therapy. *Behavioural and Cognitive Psychotherapy*, *35*, 335 - 342.

Kellett, S., Clarke, S. & Matthews, L. (2007a). Delivering group psychoeducational CBT in primary care: Comparing outcomes with individual CBT and individual psychodynamic-interpersonal psychotherapy. *British Journal of Clinical Psychology*, *46*(2), 211 - 222.

Kellett, S., Clarke, S. & Matthews, L. (2007b). Session impact and outcome in group psycho-educative cognitive behavioural therapy. *Behavioural and Cognitive Psychotherapy*, *35*, 335 - 342.

Kellett, S., Newman, D.W., Matthews, L. & Swift, A. (2004). Increasing the effectiveness of large group format CBT via the application of practise-based evidence*. Behavioural and Cognitive Psychotherapy*, *32*, 231 - 234.

Kertz, S., Bigda-Peyton, J. & Bjorgvinsson, T. (2013). Validity of the Generalized Anxiety Disorder-7 scale in an acute psychiatric sample. *Clinical Psychology and Psychotherapy*, *20*(5), 456 - 464.

Kessler, R.C., Berglund, P., Demler, O., Jin, R., Merikangas, K.R. & Walters, E.E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*, 593 - 602.

Kroenke, K., Spitzer, R.L. & Williams, J.B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*(9), 606 - 613.

Löwe, B., Decker, O., Müller, S., Brähler, E., Schellberg, D., Herzog, W. & Herzberg, P.Y. (2008). Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Medical Care*, 266 - 274.

Löwe, B., Unützer, J., Callahan, C., Perkins, A.J. & Kroenke, K. (2004). Monitoring depression treatment outcomes with the Patient Health Questionnaire-9. *Medical Care*, *42*(12), 1194 - 1201.

Lucas, R.A. & Telch, M.J. (1993). *Group versus Individual Treatment of Social Phobia*. Paper presented at the annual meeting of the Association for Advancement of Behavior Therapy, Atlanta, GA.

Lueger, R.J. & Barkham, M. (2010). Using benchmarks and benchmarking to improve quality of practice and services. In: M. Barkham, G.E. Hardy & J. Mellor-Clark (Eds.), *Developing and Delivering Practice-Based Evidence: A Guide for the Psychological Therapies* (p. 223 - 256). Wiley-Blackwell.

Lukens, E.P. & McFarlane, W.R. (2004). Psychoeducation as evidence-based practice: Considerations for practice, research and policy. *Brief Treatment and Crisis Intervention*, *4*, 205 - 225.

McMillan, D., Gilbody, S. & Richards, D. (2010). Defining successful treatment outcome in depression using the PHQ-9: A comparison of methods. *Journal of Affective Disorders*, *127*, 122 - 129.

Meyers, D.C., Durlack, J.A. & Wandersman, A. (2012). The quality implementation framework: A synthesis of critical steps in the implementation process. *American Journal of Community Psychology*, *50*(3-4), 462 - 480.

Minami, T., Serlin, R.C., Wampold, B.E., Kircher, J.C. & Brown, G.S. (2008). Using clinical trials to benchmark effects produced in clinical practice. *Quality and Quantity*, *42*, 513 - 525.

National Collaborating Centre for Mental Health (2020). *The Improving Access to Psychological Therapies Manual* (version 4). Available at: https://www.england.nhs.uk/wp-content/uploads/2020/05/iapt-manual-v4.pdf [Accessed 9th July 2020].

National Institute for Health and Care Excellence (2011). *Common Mental Health Problems: Identification and Pathways to Care* (NICE Clinical Guideline CG123). Available at: https://www.nice.org.uk/guidance/cg123/chapter/1-guidance [Accessed 2nd June 2020].

National Psychological Therapies Management Committee / Public Health Wales (2017). *Matrics Cymru: Guidance for Delivering Evidence-Based Psychological Therapy in Wales.* http://www.1000livesplus.wales.nhs.uk/sitesplus/documents/1011/Matrics%20Cymru%20%28CM%20design%20%2D%20DRAFT%2015%29.pdf [Accessed 26th May 2020]

NHS Wales (2020). Hywel Dda Local Health Board. *Demography Profile*. Available at http://www.wales.nhs.uk/sitesplus/documents/888/LHB%20HYWEL%20DDA%20%2D%20Eng.pdf [Accessed 29th May 2020].

Oosterbaan, D.B., Verbraak, M.J.P.M., Terluin, B., Hoogendoorn, A.W., Peyrot, W.J., Muntingh, A. & van Balkom, A.J.L.M. (2013). Collaborative stepped care v. care as usual for common mental disorders: 8-month, cluster randomised controlled trial. *British Journal of Psychiatry*, *203*, 132 - 139.

Rapaport, M.H., Clary, C., Fayyad, R. & Endicott, J. (2005). Quality-of-life impairment in depressive and anxiety disorders. *American Journal of Psychiatry*, *162*, 1171 - 1178.

Roth, A. & Fonagy, P. (2004). *What Works for Whom? A Critical Review of Psychotherapy Research.* (2nd ed.) Guilford Publications: New York.

Siev, J., Huppert, J. & Chambless, D. (2009). The dodobird, treatment technique and disseminating empirically supported treatments. *The Behavior Therapist*, *32*, 69 - 76.

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

Toussaint, A., Hüsing, P., Gumz, A., Wingenfeld, K., Härter, M., Schramm, E. & Löwe, B. (2020). Sensitivity to change and minimal clinically important difference of the 7-item Generalized Anxiety Disorder Questionnaire (GAD-7). *Journal of Affective Disorders, 265*, 395 - 401.

Turpin, G., Richards, D., Hope, R. & Duffy, R. (2008). Improving access to psychological therapies: A national initiative to ensure the delivery of high quality evidence-based therapies. *Papeles del Psicólogo*, *29*(3), 271 - 280.

Van Daele, T., Van Audenhove, C., Vansteenwegen, D., Hermans, D. & Van den Bergh, O. (2013). Effectiveness of a six session stress reduction program for groups. *Mental Health and Prevention*, *1*, 19 - 25.

Welsh Government (2012). *Psychological Therapies in Wales. Policy Implementation Guidance* http://wales.gov.uk/docs/dhss/publications/120326psychologicalen.pdf [Accessed 2nd June 2020]

Welsh Government (2018). The Welsh Language (no. 7) Standards. Welsh Statutory Instruments No. 441 (W. 77) http://www.legislation.gov.uk/wsi/2018/441/made [Accessed 8th July 2020]

White, J. (1995). Stresspac: A controlled trial of a self-help package for the anxiety disorders. *Behavioural and Cognitive Psychotherapy*, *23*, 89 - 108.

White, J. (1997). "Stress Control" large group therapy: Implications for managed care systems. *Depression and Anxiety*, *5*, 43 - 45.

White, J. (1998). "Stress Control" large group therapy for generalized anxiety disorder: Two year follow-up. *Behavioural and Cognitive Psychotherapy, 26(3)*, 237 - 245.

White, J. (2000). *Treating Anxiety and Stress: A Group Psycho-educational Approach Using Brief CBT*. Chichester: Wiley.

White, J. (2008). CBT and the challenge of primary care: Developing effective, efficient, equitable, acceptable and accessible services for common mental health problems. *Journal of Public Mental Health*, *7*, 32 - 41.

White, J., Keenan, M. & Brooks, N. (1992). "Stress Control": A controlled comparative investigation of large group therapy for generalized anxiety disorder. *Behavioural Psychotherapy*, *20*(2), 97 - 114.

White, J. & Keenan-Ross, M. (1997). Stress Control large group didactic therapy for anxiety: An approach for managed care systems. *The Behavior Therapist*, *20*, 192 - 196.

Wise, E.A. (2004). Methods for analysing psychotherapy outcomes: A review of clinical significance, reliable change, and recommendations for future directions. *Journal of Personality Assessment*, *82*, 50 - 59.

Wittchen, H.U., Jacobi, F., Rehm, J., Gustavsson, A., Svensson, M., Jönsson, B., … Steinhausen, H.C. (2011). The size and burden of mental disorders and other disorders of the brain in Europe 2010. *European Neuropsychopharmacology*, *21*, 655 - 679.

Wittchen, H.U., Kessler, R.C., Beesdo, K., Krause, P., Höfler, M. & Hoyer, J. (2002). Generalized anxiety and depression in primary care: Prevalence, recognition and management. *Journal of Clinical Psychiatry*, *63*, 24 - 34.

Wood, S., Kitchiner, N.J. & Bisson, J.I. (2005). Experiencing and implementing an adult educational approach to treating anxiety disorder. *Journal of Psychiatric and Mental Health Nursing*, *12*, 95 - 99.