



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/102719/>

Version: Accepted Version

Article:

Calvert, R., Kellett, S. and Hagan, T. (2015) Group cognitive analytic therapy for female survivors of childhood sexual abuse. *British Journal of Clinical Psychology*, 54 (4). pp. 391-413. ISSN: 0144-6657

<https://doi.org/10.1111/bjc.12085>

This is the peer reviewed version of the following article: Calvert, R., Kellett, S. and Hagan, T. (2015), Group cognitive analytic therapy for female survivors of childhood sexual abuse. *Br J Clin Psychol*, 54: 391–413, which has been published in final form at doi:10.1111/bjc.12085. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Objectives. The effectiveness of Cognitive Analytic Therapy delivered in groups has been under researched considering the popularity of the approach. This study sought to investigate the effectiveness of 24-sessions of group Cognitive Analytic Therapy (GCAT) delivered in routine practice for female survivors of childhood sexual abuse (CSA). **Methods.** In a longitudinal cohort design, N=157 patients were treated with 24-sessions of GCAT. Validated outcome measures were administered at assessment, pre and post-GCAT. This enabled rates of reliable and clinically significant change to be compared between wait-time and active group treatment. The uncontrolled treatment effect size was then benchmarked against outcomes from matched studies. **Results.** On the primary outcome measure, GCAT facilitated a moderate effect size of 0.34 with 11% of patients completing treatment meeting 'recovery' criteria. The drop-out rate was 19%. Significant improvements in interpersonal functioning, anxiety and wellbeing occurred during GCAT in comparison to wait-time on secondary outcome measures. **Conclusions.** GCAT appears a promising intervention for adult female CSA survivors, with further controlled evaluation indicated.

Practitioner points.

- GCAT appears a promising and acceptable intervention for female CSA survivors with high levels of psychological distress.
- Long-term follow up studies are required with CSA survivors to index the clinical durability of GCAT.
- A GCAT treatment fidelity measure needs to be developed and evaluated.

A cross-cultural meta-analysis found that up to 20% of women and 8% of men report being sexually abused before the age of 18 (Pereda, Guilera, Forns, & Gomez-Benito, 2009). In the UK, similar rates are reported with 21% of females and 11% of males disclosing histories of childhood sexual abuse (CSA; Cawson, Wattam, Brooker, & Kelly, 2000). There are, however, acknowledged clinical and methodological difficulties in estimating the 'true' prevalence and incidence rates of CSA. For example, varying definitions of CSA and the wide range of data collection techniques employed produce often widely differing prevalence rates (Putnam, 2003). Although 'survivors' constitute a heterogeneous clinical population with diverse trauma experiences and subsequent outcomes (Rutter, 2007), it is generally agreed that CSA appears to be a psychologically toxic early experience; Manglio (2009) studied 270,000 subjects from 587 studies to evidence that CSA survivors were at significant risk of a wide range of medical, psychological, behavioural and sexual disorders.

Whilst a variety of models have been proposed to understand the heterogeneous and often extensive and pervasive negative impact of CSA (see Freeman & Morris, 2001, for a review), there is yet no widely supported conceptualisation to guide treatment (Llewelyn, 2002) and a paucity of knowledge regarding effective interventions (for reviews see Llewelyn, 1997; Martsolf & Draucker, 2005; Price, Hilsenroth, Petretic-Jackson, & Bonge, 2001; Taylor & Harvey, 2010). There is therefore a need develop a robust and relevant evidence-base to guide services in appropriately responding to survivors' distress (Westbury & Tutty, 1999). However, there are many complex challenges in conducting outcome research for CSA, including multiple treatment targets (Trask, Walsh & DiLillo, 2011), the interaction of often co-occurring trauma experiences (including sexual, physical, emotional abuse and neglect; Briere & Runtz, 1990),

the diverse context of abuse experiences and disclosure (Ramchandani & Jones, 2003) and the lack of an agreed definition of CSA (Peters, Wyatt & Finkelhor, 1986). This study was conducted in routine practice and the service used the following definition provided by the World Health Organisation (1999): “Child sexual abuse is the involvement of a child in sexual activity that he or she does not fully comprehend, is unable to give informed consent to, or for which the child is not developmentally prepared and cannot give consent, or that violate the laws or social taboos of society. Child sexual abuse is evidenced by this activity between a child and an adult or another child who by age or development is in a relationship of responsibility, trust or power, the activity being intended to gratify or satisfy the needs of the other person.”

In terms of treatment of the difficulties arising from CSA, there is a tension between controlled *efficacy* studies (which prioritise higher degrees of internal validity through rigorously controlled methods) and *effectiveness* studies completed in routine practice settings (which tend to exemplify high external validity due to being conducted in real-world settings; Roth & Fonagy, 2005). The clinical utility of controlled trials has been challenged (Cartwright, 2007), due to use of strict inclusion/exclusion criteria leading to difficulties in applying results to routine practice. A *practice-based evidence* (PBE) paradigm has therefore been advocated (Cahill, Barkham & Stiles, 2010) that utilises routinely available ‘real world’ datasets (Holloway, 2002) to usefully complement and contextualise efficacy evidence (Bower & Gilbody, 2010). Effectiveness and efficacy of treatment is often summarised in terms of an *effect size* of the intervention, which is a standardized measure of effects of treatment (e.g. via metrics such as Cohen’s *d* and odds ratios). The acceptability of a treatment is also an important component of its effectiveness (Cavanagh et al., 2009). For example, meta-analytic evidence

suggests that dropout rates from therapy are particularly high for survivors of CSA (Taylor & Harvey, 2010). So the manner in which a therapy engages and retains CSA survivors in treatment is also an important index of its usefulness (DoH, 2007).

Several studies (e.g., Alexander, Neimeyer, Follette, Moore, & Harter, 1989; Hazzard, Rogers, & Angert, 1993) and reviews (Kessler, White, & Nelson, 2003; Taylor & Harvey, 2010) highlight the promise of group psychotherapy for CSA survivors. Whilst group cognitive analytic therapy (GCAT) has been advocated as a treatment method (Hagan & Gregory, 2001), the associated evidence base is small and contains just three small evaluations of routine practice, with insufficient reporting of results to compute effect sizes. Duignan and Mitzman (1994) delivered a 12-week GCAT intervention to seven survivors and found significant improvements in depression and wellbeing, although selection bias may have influenced results. Clarke and Llewelyn's (1994) study of seven female survivors showed that despite positive outcomes, only a small number of the women's constructs changed suggesting the persistence of the centrality of abuse despite the GCAT intervention. Finally, Ryan, Nitsun, Gilbert and Mason (2005) completed a CAT-informed psycho-educational group with 22 survivors and found statistically and clinically significant improvements to general wellbeing. All the GCAT studies had high external validity due to being conducted in routine practice, but suffered from small sample sizes and poor internal validity (e.g. lack of diagnostic validity, absence of treatment fidelity measurement and no comparison/control groups).

In summary, a wealth of evidence confirms the negative, long-term psychological impact of CSA and yet relevant outcome research remains in its

empirical infancy. GCAT is a treatment approach currently unsupported by a robust evidence base, despite the recent call to produce evidence for group CAT delivery (Ryle, Kellett, Hepple & Calvert, 2014). This study aimed to evaluate the outcomes of GCAT in routine clinical practice for female CSA survivors, through investigating differences between outcomes during treatment versus wait-time. This method supports inferences that any differences found would be attributable to introduction of treatment (Cisler, Barnes, Farnsworth, & Sifers, 2007). This study also employed benchmarking (Lueger & Barkham, 2010) to contextualise the effect size of GCAT against other group psychotherapy outcome studies for female CSA survivors. The study hypothesised that (1) significant improvements in distress and functioning would occur during GCAT, (2) more patients would recover during GCAT than during wait-time and (3) outcomes for GCAT would be equivalent to those found in published studies of group treatments in female survivor populations.

Method

Setting and design

Ethical approval was granted by the appropriate UK NHS Research Ethics Committee and registered with the participating NHS Trust's Governance department. The study used a longitudinal, cohort design. GCAT was delivered in a tertiary psychotherapy service, which offered GCAT (and other individual therapies) to adult survivors of CSA referred from Secondary Care. Consequently all patients had complex care packages that included concurrent input from other aspects of mental health services (e.g. psychiatric out-patient appointments, day care services and on-going contact with care co-ordinators). As no strict inclusion/exclusion criterion was applied, the patients were clinically

representative (Shadish, Matt, Navarro, Siegle, Crit-Christoph, & Hazelrigg et al., 1997).

Sample and Procedure

Patients receiving GCAT all had histories of CSA and were referred to the service on the criteria that the CSA played a central role in their ongoing psychological distress, disorganised engagement patterns with services, elevated risk and/or poor self-care. The service responded to appropriate referrals by sending a letter to the patient inviting them to an assessment appointment. Included with the letter was a booklet containing a range of outcome measures and a request for it to be returned to the service prior to the assessment. Initial face-to-face assessment appointments were conducted by a GCAT facilitator and typically lasted for approximately 60 minutes. No standardised diagnostic instruments were used during the assessment. Patients were not randomly allocated to interventions, but involved in collaborative discussions concerning possible treatment options. A variety of treatment choices were negotiated, with GCAT being one treatment option. Allocation thus reflected clinical decision-making rather than adherence to randomisation procedures (Buckley, Newman, Kellett & Beail, 2006). Participants who opted-in to GCAT were invited to join the next group and placed on a waiting list. At this point patients were provided with tailored CSA psychoeducational resource materials to enable them to prepare for GCAT (e.g. Ainscough & Toon, 2000). The time interval between assessment and start of the next group (pre-GCAT) represented a naturally occurring waitlist control, although this time varied according to staff availability, as only one GCAT group ran at any one time. Wait-times were a maximum of eight months.

For the purposes of the study, a clinician retrospectively coded information concerning the psychological difficulties reported at clinical assessment. Information was coded according to presence of internalising difficulties (e.g. depression/anxiety) and externalising difficulties (e.g. aggression), current substance use, self-harm behaviours, sexual difficulties (conceptualised as either an aversion or preoccupation with sexual activity) and re-victimisation experiences (e.g. domestic violence). No information was available regarding patients' CSA experiences (e.g. perpetrator(s), severity, type or duration of abuse, age at onset, etc.), as participants were never required to disclose the details of their abuse histories during the assessment.

This study drew on a clinical database of all patients referred to the service (N=378) from which a subset of those who had been offered GCAT (N=157) were identified, using the following post-hoc criteria; (1) patients had completed assessment measures, (2) attended assessment and (3) been offered and accepted GCAT regardless of whether or how much of GCAT they subsequently attended. This therefore identified an initial sample of patients who had been offered GCAT (N=157), from which a further sub-group went on to complete GCAT treatment; the 'completer sample' (N=108; see Figure 1 for patient flow through the stages of the study).

insert figure 1 here please

Outcome measures were completed at three time points (assessment, pre and post-GCAT) by 57% (N=89) of those patients offered GCAT (N=157). For participants not completing outcomes following assessment (i.e. did not complete pre-GCAT measures N=47; did not complete post-GCAT measures, N=21), last

observation carried forward was employed (Montori & Guyutt, 2001; Barkham et al, 2011). This ensured subsequent analyses were not 'weighted' and provided a conservative estimate of change (i.e. this sub-sample would show no change during GCAT).

The age of patients who were offered GCAT (N=157) ranged from 18 to 64 years (mean age=35 years, SD=11), 88% identified themselves as 'White British', 55% reported currently being in a relationship and 28% in paid employment. Difficulties with current substance use were recorded in 11% of clinical notes of those offered GCAT at assessment. Self-harming behaviour was recorded in 31% of assessment records and 10% noted current sexual difficulties. Furthermore, at assessment, 94% of the clinical assessments recorded internalising difficulties (i.e. depression, anxiety) and 29% reported current externalising difficulties (i.e. aggression). Experiences of re-victimisation were recorded in 12% of cases.

Analysis strategy

Overall the main analyses comprised of repeated measures ANOVAs of wait-time versus active treatment, calculation of rates of reliable and clinically significant change, computation of uncontrolled effect sizes and benchmarking the effect size on the primary outcome measure with the extant evidence.

Change during wait-list and GCAT was evaluated using 2 (completers/non-completers) x 3 (assessment/pre-GCAT/post-GCAT) ANOVAS. The reliable change index (RCI; Jacobson & Truax, 1991) was used to evaluate the extent to which any individual participants' change score during GCAT was beyond measurement error, and so defined *reliable change*. *Clinically significant change* required participants' scores on a measure pre/post GCAT to shift from within to outwith the scores associated with a clinical population. Where normative data was not

available, clinically significant change was operationalised by either (a) change in scores by at least two standard deviations from the mean of a clinical population towards a non-clinical population, or (b) change to within two standard deviations of the mean of a non-clinical population. This study used the more stringent index of clinically significant change using the published test-retest co-efficient and published clinical thresholds representing the cut-off between clinical and community populations where available. However, where these were not available, a clinical cut-off was derived (see Table 1 for a summary of indices and evidence used to calculate the individual change rates). Combining both individual change indices enabled rates of *reliable and clinically significant change* (RCSC) to be calculated. RCSC is often used as an index of *recovery* in practice-based evidence (Barkham, Stiles, Connell & Mellor-Clark, 2012). McNemar tests were used to compare recovery rates between (1) assessment to start of GCAT, and (2) pre-post GCAT.

Uncontrolled effect sizes (Cohen's d_+), that is the amount of change *within* GCAT from start to end of group without reference to a control/alternative treatment condition, were calculated using the pre-post change score during group therapy divided by the pre-group standard deviation (Barkham, Gilbert, Connell, Marshall, & Twigg, 2005; Westbrook & Kirk, 2005). There are no agreed conventions for classifying within-group effect sizes, however, it is reasonable to assume that within-group effect sizes would be larger than between-group effect sizes. Also given the severe, enduring and complex difficulties of the sample, effect sizes would be expected to be smaller than those achieved in other clinical populations. To reflect these issues, the Conway, Audin, Barkham, Mellor-Clark, & Russell (2003) effect size approach was followed for group-based work with clients with severe/complex difficulties. This classifies a within-group

uncontrolled effect size of $d_+ \geq 0.2$ as 'moderate improvement' and $d_+ \geq 0.5$ as 'marked improvement'. This study benchmarked the GCAT effect size on the primary outcome measure against other group therapies for female survivors of CSA that have used the Brief Symptom Inventory (Derogatis & Melisaratos, 1983) or its predecessor the Symptom Checklist-90-Revised (SCL-90-R; Derogatis, Lipman, & Covi, 1973). The results of which were summarised in a forest plot.

Measures

Patients completed a battery of outcome measures at three time points; assessment, pre-GCAT in the first GCAT session and post-GCAT at the final session. If patients provided insufficient outcome data on a measure at any time-point (according to the scoring procedure for each measure), no score was calculated for that measure.

Primary Measure

Brief Symptom Inventory (BSI; Derogatis & Melisaratos, 1983)

The BSI is a well validated and reliable short version of the SCL-R-90 (Derogatis, Lipman, & Covi, 1973). The 53-item scale measures psychiatric symptoms across nine primary symptom dimensions and three global indices: Global Severity Index (BSI-GSI), Positive Symptom Total (PST) and Positive Symptom Distress Index (PSDI). Current $\alpha = 0.97$. The BSI-GSI is a mean score, combining information about the overall number and intensity of distressing symptoms. The clinical threshold for 'caseness' is a GSI t-score of 63 or higher (Derogatis, 1993), which equates with a threshold score of 0.78. The Brief Symptom Inventory (BSI) was selected as the primary outcome measure because, (a) it is a valid and reliable index of psychological distress (Derogatis & Melisaratos, 1983), (b) given the

diverse nature of difficulties that survivors experience, a symptom specific measure would be unlikely to capture the extent of patients' distress and (c) it has been the most widely used measure across the CSA group outcome evidence base.

Secondary measures

Inventory of Interpersonal Problems-32 (IIP-32; Barkham, Hardy, & Startup, 1996).

This measure is used to identify interpersonal difficulties and is a valid and reliable short version of the original 127-item scale (Horowitz et al., 1988). The IIP-32 has eight subscales forming four bipolar factors: hard to be assertive vs. too aggressive; hard to be sociable vs. too open; hard to be supportive vs. too caring and hard to be involved vs. too dependent. Current $\alpha = 0.87$. Given that there are no published clinical thresholds for the IIP-32, a cut-off score of 1.39 was derived according to Jacobson and Truax's (1991) criterion 'C' (i.e. utilising normative data reported in Barkham et al. (1996) to enable rates of clinically significant change to be calculated).

Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1989)

The RSES is a ten item scale used to measure perceptions of global self-worth (Rosenberg, 1989). Silber and Tippet (1965) reported a test-retest co-efficient of 0.85, and Liem and Boudewyn (1999) an internal consistency co-efficient of 0.88 with survivors of CSA. Current study $\alpha = 0.81$. A clinical threshold of 25.22 was determined as two standard deviations above the assigned-to-GCAT sample's mean score at assessment to enable rates of clinically significant change to be calculated (Evans, et al., 1998).

Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983)

The HADS is a valid and reliable 14-item measure (Savard, Laberge, Gauthier, Ivers, & Bergeron, 1998) of anxiety (HAD-A) and depression (HAD-D). Crawford,

Henry, Crombie, and Taylor (2001) suggest a clinical threshold value of 11.

Current study $\alpha = 0.83$ for both HADS-A and HAD-D.

General Health Questionnaire-28 (GHQ-28; Goldberg & Hillier, 1979).

The GHQ-28 is a valid and reliable measure of non-psychotic mental health difficulties yielding a total score, and four sub-scales: somatic symptoms, anxiety/insomnia, social dysfunction and severe depression. There are two scoring methods for the GHQ-28; Likert scoring and GHQ scaled scores, with the latter employed in the current study. A WHO study of 5,438 participants from 15 different locations, derived a clinical threshold of more than or equal to seven (Goldberg et al, 1997). Current study $\alpha = 0.95$.

insert table 1 here please

Facilitators

There were 14 female GCAT facilitators; seven clinical psychologists (two of whom were CAT accredited practitioners), two mental health nurses and five trainee clinical psychologists with prior experience of delivery of individual CAT. Two staff members co-facilitated each group session and trainee psychologists were always paired with qualified clinical psychologists when delivering groups. The ACAT accredited practitioners supervised all facilitators.

Intervention: Group Cognitive Analytic Therapy

CAT integrates psychoanalytic and cognitive models to offer a transdiagnostic, time-limited (usually 16 or 24 sessions) and relational approach to facilitating therapeutic change (Ryle & Kerr, 2002). The evidence-base for CAT is made up of generally high quality studies (Calvert & Kellett, 2014), with a weighted mean

effect size across CAT outcome studies of $d_+ = 0.83$ (Ryle, Kellett, Hepple & Calvert, 2014). Theoretically, CAT draws on personal construct theory (Kelly, 1956) and object relations theory (Ryle, 1985), asserting that mental representations of self, others and the world are developmentally formed by early interactions with significant others (Ryle & Kerr, 2002). These internalised, early object relations are termed '*reciprocal roles*' and influence how individuals anticipate and react to relationships. CAT suggests that CSA survivors have learnt a repertoire of reciprocal roles and '*target problem procedures*' (TPPs; commonly referred to as traps, snags and dilemmas; Ryle & Kerr, 2002) to 'survive' the adversity experienced in childhood (Clarke & Llewelyn, 1994), but which are now maladaptive as an adult. Change in CAT is considered to arise from the process of early, collaborative narrative '*reformulation*' that develops a shared understanding to explain the developmental origins of difficulties (Ryle, 1995). A sequential diagrammatic reformulation (SDR) is constructed to identify predominant reciprocal roles and repetitive patterns that maintain difficulties and limit change (Ryle, 1997). The SDR is used to facilitate recognition of damaging patterns, both external to therapy and within the therapeutic relationship. 'Exits' are identified to actively revise maladaptive procedures, with the therapist aiming to offer a containing, non-collusive experience throughout.

Broadly, the structure of GCAT followed a reformulation, recognition and revision approach (Hagan & Gregory, 2001). Groups ran for 24 weekly sessions with each session lasting 90 minutes. Group size varied according to referral rate, with a median of eight patients in each group. Group members completed the 'Psychotherapy File' in early sessions; a CAT specific self-assessment questionnaire that helps patients to initially recognise their traps, snags and dilemmas and also various self-states (e.g. dissociated 'zombie' states). Facilitators delivering GCAT

used a resource file that defined the outlines and aims for each group session. However, GCAT was not manualised and specific interventions within groups varied over time and in response to the needs of the members and group dynamics.

Individual diagrammatic reformulations were developed with group members (Duignan & Mitzman, 1994), alongside a group reformulation based upon a conceptual tool developed by the service; 'lessons learned to survive' (Hagan & Gregory, 2001). Stowell-Smith, Gopfert and Mitzman (2001) noted that "multiple reciprocal role enactments" emerged during group dynamics, with group diagrammatic reformulations providing a framework for reflection and change. Recognition during GCAT was facilitated via patient self-monitoring (e.g. completing diaries to increase awareness and encourage reflection on TPPs) and reflecting on any enactments within the group setting (e.g. looking after the needs of other group members as a means of neglecting oneself). In the third phase of 'revision', exits were identified within groups to actively revise the maladaptive patterns (e.g. finding ways to safely expressing anger, setting interpersonal boundaries, and improving self-care). GCAT utilised a judicious approach to scaffolding 'exits' in response to the needs of group members. Exits drew on a range of change methods (Ryle, 1995) and were practiced between sessions via scheduled homework. The time-limited nature of GCAT meant that emphasis was also placed on the importance of therapeutic endings. Therefore at termination, GCAT members and facilitators exchanged goodbye letters, enabling reflection on changes achieved, potential future goals and obstacles to change.

Results

Study Group Comparisons

Table 2 contains the demographic details for patients offered GCAT (N=157) and all other patients who were referred to the service but not offered GCAT (N=211). Patients offered GCAT were more likely to be in a relationship ($\chi^2(1, N=325) = 6.88, p=0.009$). There were no other significant differences regarding age, ethnicity or employment status. Table 3 contains the outcome scores at assessment for patients offered GCAT (N=157) and patients who returned intake questionnaires but were not offered GCAT (N=81). Patients offered GCAT had significantly lower self-esteem scores ($Z=-5.34, n_1=71, n_2=149, p<0.001$). GCAT patients' mean score on the primary outcome measure at assessment (BSI-GSI; $M=2.28, SD=0.87$), was higher than UK outpatient norms ($M=1.66, SD = 0.83$; Ryan, 2007) and previous GCAT participants ($M=1.80, SD=1.13$; Clarke & Llewelyn, 1994; see Table 1). One hundred and forty patients (91%) of the GCAT sample scored within a clinical range on the BSI at assessment.

insert table 2 and 3 here please

Acceptability of GCAT

In terms of acceptability of treatment, 69% (N=108) of patients that started GCAT completed treatment. The GCAT treatment refusal rate (offered but did not attend GCAT) was 12% (N=19), with 19% (N=49) dropping out during group treatment. As shown in table 4, patients that completed GCAT scored significantly higher on the self-esteem measure at the start of the group than non-completers (RSES, $z=-3.02, p=0.003$). There were no other significant differences in terms of

demographic variables or outcome measure scores between the non-completer and completer samples.

insert table 4 here please

Primary Outcomes

Table 5 presents the pre/post GCAT mean score, change scores, 95% confidence intervals, uncontrolled effect sizes and RCI rates for the patients offered GCAT and GCAT completer sample. The effect size for the primary outcome measure (BSI-GSI $d_+ = 0.34$) suggests that completion of GCAT was associated with a moderate improvement in global psychological distress. The results of the ANOVA indicated a main effect of time on the primary outcome (BSI-GSI; $F(1.762, 260.79) = 9.93$, $p < 0.001$), but not completer/non-completer status ($F(1, 148) = 0.037$, $p = 0.85$). Simple contrasts illustrated that patients offered GCAT experienced statistically significant reductions in global distress scores from assessment to the end of GCAT ($F(1, 148) = 14.765$, $p < 0.001$), however, significant reductions from assessment to start of GCAT ($F(1, 148) = 5.08$, $p = 0.026$) and during GCAT ($F(1, 148) = 6.764$, $p = 0.01$) also occurred. Such improvement during wait-time reduces the confidence with which change during GCAT can be attributed to treatment. There was, however, a significant interaction between completer status and time ($F(1.76, 260.79) = 5.374$, $p = 0.007$), with simple contrasts suggesting that completers achieved significantly more therapeutic gains on the BSI during GCAT ($F(1, 148) = 7.43$, $p = 0.007$).

Insert table 5 here please

Figure 2 displays RCI rates for the GCAT completer sample on the BSI-GSI where scores were available at both start and end of GCAT (N=103). The dashed vertical and horizontal lines represent the clinical cut-off (0.78) at pre/post GCAT and each dot represents a GCAT patient. Eleven patients (11%) demonstrated reliable and clinically significant change and therefore met the criteria for 'recovery' and a further ten patients (10%) achieved a reliable improvement in BSI-GSI outcomes. Seven patients (7%) met the criteria for reliable deterioration, one of whom demonstrated a reliable and clinically significant deterioration during GCAT. Figure 2 shows that a 'stasis' outcome was the most common individual outcome following GCAT on the BSI-GSI.

insert figure 2 here please

Table 6 presents the benchmarking evidence from matched group treatment outcome studies of female survivors of CSA and Figure 3 presents a forest plot of associated uncontrolled effect sizes with the 95% confidence intervals. These are weighted according to sample size. The effect sizes for the group therapy studies ranged between 0.34-1.02 and had a standard deviation of 0.26. Where confidence intervals for group therapies overlap the vertical full line, it demonstrates that at the given level of confidence, the effect size did not differ from 'no effect' (the vertical full line) for that outcome study. Confidence intervals in three studies indicate detrimental therapeutic effects. However, the small sample sizes (all $N \leq 46$) of these studies resulted in much broader confidence intervals (Lueger & Barkham, 2010). The weighted mean group therapy effect size was 0.56 (the vertical dashed line) with a 95% confidence interval from 0.39 to 0.73 ($k=7$, $N=297$). GCAT had a moderate within-study uncontrolled effect size of

0.34, which contributed 37.83% to the overall between-studies effect size. There was no significant heterogeneity between studies ($p=0.08$) suggesting that it is appropriate to draw the tentative comparative conclusions.

insert table 6 and figure 3 here please

Secondary Outcomes

Table 4 also contains the pre/post GCAT secondary outcome measure means, change scores, 95% confidence intervals, uncontrolled effect sizes and RCI rates for the offered GCAT (d_+ range 0.23-0.38) and GCAT completer samples (d_+ range 0.35-0.58). Effect sizes suggest that completion of GCAT was associated with generally moderate to substantial improvements across the secondary outcome measures (excluding self-esteem).

In terms of interpersonal functioning scores, no statistically significant improvement was observed during wait-time ($F(1,145) = 1.55, p=0.22$), but a significant difference was observed during GCAT ($F(1,145) = 5.751, p=0.02$). It is possible therefore to tentatively infer that interpersonal change achieved was due to the intervention. Completer status and time interacted ($F(1.67, 242.07) = 4.219, p=0.02$) to again suggest that completers achieved more change during GCAT ($F(1,145) = 5.41, p=0.021$).

Main effects over time were observed for depression (HADS-D; $F(1.425, 212.35) = 17.721, p < 0.001$) and anxiety scores (HADS-A; $F(1.62, 243.017) = 9.39, p < 0.001$). Simple contrasts revealed both an overall improvement in depression scores during baseline ($F(1,149) = 6.86, p=0.01$) and scores during GCAT ($F(1,149) = 14.103, p < 0.001$). Anxiety scores appeared more stable during baseline ($F(1,150) = 1.284, p=0.259$), with improvements observed in anxiety scores during GCAT

($F(1,150) = 10.614, p = 0.001$). Completers' depression ($F(1,149) = 11.06, p = 0.001$) and anxiety scores ($F(1,150) = 8.79, p = 0.004$) reduced more than non-completers during GCAT. No significant interaction was observed during baseline for either depression ($F(1,149) = 0.439, p = 0.51$) or anxiety scores ($F(1,150) = 0.34, p = 0.56$).

Overall, for patients offered GCAT, there was an increase in self-esteem scores whilst waiting for treatment ($F(1,145) = 20.85, p < 0.001$) and a difference was also observed during GCAT ($F(1,145) = 12.138, p = 0.001$). The difference in self-esteem scores between assessment and end of GCAT were not significant ($F(1, 145) = 3.667, p = 0.057$). There was an interaction between time and completers status ($F(1.80, 261.54) = 6.20, p = 0.003$), with a significant difference between completers/non-completers' self-esteem scores at assessment, with completers then achieving significantly greater gains ($F(1,145) = 8.35, p = 0.004$). There was also a significant difference between completers/non-completers' self-esteem outcomes during GCAT ($F(1,145) = 13.26, p < 0.001$). However, this was not in the expected direction. During GCAT completers' self-esteem scores significantly deteriorated. No main effect for completer status was observed ($F(1,145) = 1.647, p = 0.201$).

In terms of general well-being (GHQ scores; $N = 150$), there was a significant main effect of time ($F(1.736, 256.879) = 11.09, p < 0.001$). Simple contrasts revealed no significant change in well-being scores during baseline ($F(1,148) = 3.268, p = 0.073$), with a significant overall improvement during GCAT ($F(1,148) = 8.817, p = 0.003$). No main effect of completer status was found ($F(1,148) = 0.052, p = 0.821$), but completer status and time significantly interacted ($F(1.74, 256.88) = 6.04, p = 0.004$). No significant interaction was observed during baseline ($F(1, 148) = 0.001, p = 0.978$), but based on their GHQ scores completers achieved significantly more gains in terms of their general mental health during GCAT

($F(1,148) = 7.831, p = 0.006$). Pre-post individual change rates (i.e. RCI analysis) were highest on the depression (HADS-D) and general mental health (GHQ) outcome measures. The reliable and clinically significant improvement rate was 16% ($N=16$) for depression and 16% ($N=16$) for general mental health. In terms of just a reliable improvement, then 23 patients (22%) achieved a reliable reduction in anxiety scores, of whom almost half (10%) also achieved a clinically significant change. Despite the overall statistically significant deterioration in mean self-esteem scores during GCAT, 14% ($N=14$) of those completing treatment achieved reliable and clinically significant improvement and 11% ($N=11$) achieved a reliable improvement. Although mean IIP-32 scores significantly improved during GCAT, no single patient met the criteria for a reliable and clinically significant improvement in their interpersonal functioning. Reliable deterioration rates ranged from 2% to 7%, with five patients demonstrating reliable and clinically significant deterioration on the GHQ-28. McNemar tests illustrated that a significantly greater proportion of completers achieved reliable improvements during GCAT compared to wait-time: BSI-GSI ($p=0.004$); IIP-32 ($p=0.001$); RSES ($p=0.001$); HAD-A ($p<0.001$); HAD-D ($p<0.001$) and GHQ ($p<0.001$).

Discussion

This study represents the first attempt to evaluate GCAT for highly distressed female survivors of CSA in routine clinical practice. The results generally suggest that GCAT can be an effective approach for those patients completing therapy, as an adjunct to care provided in secondary mental health setting. Completing treatment was found to be beneficial, which is consistent with previous findings (e.g. Cahill et al., 2003). One in five women achieved a reliable improvement during GCAT on the primary outcome measure. It is worth noting

however that 7% of GCAT patients also experienced a reliable deterioration in their global psychological distress. A relatively small (but nontrivial) minority of patients can deteriorate during psychological treatment, with estimates ranging from 3-10% (Mohr, 1995). The low rates of change found may have been due to the relative brevity of GCAT given the severity and complexity of patients' difficulties. Previous research has suggested that patients with severe and enduring interpersonal difficulties require lengthy group treatments, which far exceed the 24-session format evaluated in the current study (Budman & Gurman, 1988; Lorentzen & Høglend, 2008).

GCAT was found to be a (statistically) moderately effective intervention with an effect size of 0.34 on the primary outcome measure. However, statistical improvements in global psychological distress were also demonstrated from assessment to start of group therapy. This improvement does undermine confidence in attributing change solely to GCAT attendance. Benchmarking the GCAT effect size found it to be lower than for other group therapies. This may have been due to the high pre-treatment distress apparent in the current sample; the GCAT effect size was comparable to analytic group therapy for CSA survivors with similarly high levels of pre-treatment distress (Lau & Kristensen, 2007). Differences may also be due to factors such as variance in group climate or specific patient/therapist variables (Ogrodniczuuk, Piper & Joyce, 2006) not measured in this study.

In terms of acceptability of treatment, 69% of patients offered GCAT completed treatment, with 12% failing to attend any sessions and 19% dropping out during treatment. This represents a relatively low dropout rate compared to other group therapy approaches used with females CSA survivors. For example, Fisher et al., (1993) reported a drop out rate of 41% from group psychodynamic

psychotherapy, Lau & Kristensen (2007) 38% from systemic group therapy and Talbot et al, (1999) 58% from a 'trauma recovery' group. Ryle et al., (2014) noted that a key feature of CAT is rapid reformulation and speculated that this may be a key factor contributing to the low dropout rates observed across CAT outcome studies. Acceptability of treatment is particularly pertinent for survivors who often endure on-going marginalisation and adversity, which can then markedly limit their capacity to effectively engage with services (Fisher, *et al*, 1993).

On the secondary outcome measures, interpersonal functioning, anxiety and wellbeing scores were stable during wait-time and improved during GCAT, indexing the impact of group treatment on these factors. Similar to the outcome pattern on the primary outcome measure, there was a trend for improvements in depression and self-esteem scores prior to GCAT during wait-time following initial screening. Previous research has highlighted the therapeutic impact of hope-inducing, collaborative assessment (Finn & Tonsager, 1997). Depression scores continued to improve during GCAT, however, the self esteem scores of completers did deteriorate during group treatment. Previous CAT research with CSA survivors has also evidenced deterioration of self-esteem scores during group treatment (Clarke & Pearson, 2000).

As this study involved the retrospective analysis of a practice-based dataset there are many aspects of internal validity that are open to criticism, such as lack of methodological control and also threats to the actual quality of the data (i.e. missing outcome data, Barkham, Stiles, Lambert & Mellor-Clark, 2010). The results therefore need to be interpreted with due caution. An obvious limitation was the lack of a contemporaneous comparison or control group to compare the GCAT outcomes against. Although the study attempted to address this by use of a within group waitlist comparison, this was not a completely adequate control as it

was not possible to account for the differing lengths of time that patients waited for GCAT. There was also no systematic recording of concurrent interventions from community mental health teams. Therefore, it is impossible to say with any level of certainty that the improvement (or deterioration) observed during GCAT could be solely attributed to group treatment. All measures were fairly generic measures of mental health and the study could have been improved by the inclusion of a CSA specific outcome measure (e.g. The Trauma-Related Guilt Inventory; Kubany et al. 1996) or CSA-related treatment targets, such as reduced incidents of self-harm and/or inappropriate sexual boundaries. No measure of therapist competence/model fidelity was used and therefore poor adherence or therapist drift (Waller, 2009) may well have occurred. Furthermore, although this study explored group CAT outcomes, it did not take into account intragroup effects that may have influenced findings, for example, correlations between group members' scores (Baldwin, Murray & Shadish, 2005). Whilst the study usefully benchmarked GCAT outcomes, the range of available benchmarks was limited and so only tentative comparative conclusions could be drawn. The lack of follow-up data represents a major study weakness.

This study does suggest some potential new avenues for future research. A pragmatic trial of group CAT for survivors is indicated from these preliminary findings. Future GCAT studies should report the intraclass correlation as this will help to identify the intragroup effects that may be influencing outcome and also ensure adequate sample sizes to ensure statistical power (Kenny, 2002). Methodologies that enable short and long-term follow-up from groups would index the clinical durability GCAT. Research is needed to determine the optimum GCAT treatment duration for survivors with significant interpersonal difficulties. As group delivery of CAT is increasingly popular (Ryle et al., 2014), a measure of

group treatment fidelity is needed to mirror the competency measure developed for use with individual CAT (CCAT; Bennett & Parry, 2004). Findings were limited to female survivors and future studies should prioritise investigating outcomes for male survivors; a much neglected clinical and research population (O’Leary & Gould, 2010).

In conclusion, this study suggests encouraging initial evidence that GCAT appears an acceptable and moderately effective treatment with highly distressed female survivors of CSA. A small proportion of female survivors achieved ‘recovery’, based on a stringent individual change criterion. Clearly, the GCAT approach is much in need of further detailed and controlled evaluation. There remains an urgent need for researchers and clinicians to coordinate strategies to improve the overall quality of psychological care offered to men and women struggling with the emotional consequences of being sexually abused as a child.

References

- Ainscough, C., & Toon, K. (2000). *Breaking free: Help for survivors of child sexual abuse* (2nd ed.). London: Sheldon Press.
- Alexander, P. C., Neimeyer, R. A., Follette, V. M., Moore, M. K., & Harter, S. (1989). A comparison of group treatments of women sexually abused as children. *Journal of Consulting and Clinical Psychology, 57*, 479-483.
- Baldwin, S.A., Murray, D.M. & Shadish, W.R. (2005). Empirically supported treatments or Type I errors?: Problems with the analysis of data from group-administered treatments. *Journal of Consulting and Clinical Psychology, 73*, 924-935. doi:
- Barkham, M., Hardy, G. E., & Startup, M. (1996). The IIP-32; a short version of the inventory of interpersonal problems. *British Journal of Clinical Psychology, 35*, 21-35. doi: 10.1111/j.2044-8260.1996.tb01159.x
- Barkham, M., Gilbert, N., Connell, J., Marshall, C., & Twigg, E. (2005). Suitability and utility of the CORE-OM and CORE-A for assessing severity of presenting problems in psychological therapy services based in primary and secondary care settings. *British Journal of Psychiatry, 186*, 239–246. doi: 10.1192/bjp.186.3.239
- Barkham, M., & Parry, G. (2008). Balancing rigour and relevance in guidelines development for depression: The case for comprehensive cohort studies. *Psychology and Psychotherapy: Theory, Research and Practice, 81*, 399-417. doi: 10.1348/147608308X322862.
- Barkham, M., Stiles, W. B., Lambert, M. J., & Mellor-Clark, J. (2010). Building a Rigorous and Relevant Knowledge Base for the Psychological Therapies. In

M. Barkham, G. Hardy & J. Mellor-Clark (Eds.), *Developing and delivering practice-based evidence*. (pp. 21-61). Chichester: John Wiley & Sons.

Barkham, M., Hardy, G., & Mellor-Clark, J. (2010). Improving practice and enhancing evidence. In M. Barkham, G. Hardy & J. Mellor-Clark (Eds.), *Developing and delivering practice-based evidence*. (pp. 329-354). Chichester: John Wiley & Sons.

Barkham, M., Stiles, W. B., Connell, J., & Mellor-Clark, J. (2011). Psychological treatment outcomes in routine NHS services: What do we mean by treatment effectiveness? *Psychology and Psychotherapy: Theory, Research and Practice*, 85, 1-16. doi: 10.1111/j.2044-8341.2011.02019.x

Bennett, D., & Parry, G. (2004). A measure of psychotherapeutic competence derived from cognitive analytic therapy. *Psychotherapy Research*, 14, 176-192. doi: 10.1093/ptr/kph016

Bower, P., & Gilbody, S. (2010). The current view of evidence and evidence-based practice. In M. Barkham, G. E. Hardy & J. Mellor-Clark (Eds.), *Developing and delivering practice-based evidence*. (pp. 3-20). Chichester: John Wiley & Sons.

Briere, J., & Runtz, M. (1987). Post sexual abuse trauma: Data and implications for clinical practice. *Journal of Interpersonal Violence*, 2, 367-379. doi: 10.1177/088626058700200403

Cahill, J., Barkham, M., Hardy, G., Rees, A., Shapiro, D. A., Stiles, W. B., & Macaskill, N. (2003). Outcomes of patients completing and not completing cognitive therapy for depression. *British Journal of Clinical Psychology*, 42, 133-143. doi: 10.1348/014466503321903553

- Cahill, J. L., Barkham, M., & Stiles, W. B. (2010) Systematic review of practice-based research on psychological therapies in routine clinic settings. *British Journal of Clinical Psychology*, 49, 421-453. doi: 10.1348/014466509X470789
- Cartwright, N. (2007). Are RCTs the gold standard? *Biosocieties*, 2, 11-20. doi: 10.1017/S1745855207005029
- Cavanagh, K., Shapiro, D. A., Berg, S. V. D., Swain, S., Barkham, M., & Proudfoot, J. (2009). The acceptability of computer-aided cognitive behavioural therapy: a pragmatic study. *Cognitive Behaviour Therapy*, 38, 235-246. doi: 10.1080/16506070802561256
- Cawson, P., Wattam, C., Brooker, S., & Kelly, G. (2000). *Child maltreatment in the UK: A study of the prevalence of child abuse and neglect*. London: NSPCC.
- Cisler, J. M., Barnes, A. C., Farnsworth, D., & Sifers, S. K. (2007). Reporting practices of psychological research using a wait-list control: Current state and suggestions for improvement. *International Journal of Methods in Psychiatric Research*, 16, 34-42.
- Clarke, S., & Llewelyn, S. (1994). Personal constructs of survivors of childhood sexual abuse receiving cognitive analytic therapy. *British Journal of Medical Psychology*, 67, 273-289. doi: 10.1111/j.2044-8341.1994.tb01796.x
- Clarke, S., & Pearson, C. (2000). Personal constructs of male survivors of childhood sexual abuse receiving cognitive analytic therapy. *British Journal of Medical Psychology*, 73, 169-177. doi: 10.1348/000711200160408
- Conway, S., Audin, K., Barkham, M., Mellor-Clark, J., & Russell, S. (2003). Practice-based evidence for a brief time-intensive multi-modal therapy guided by

group-analytic principles and method. *Group Analysis*, 36, 413-435. doi:
10.1177/05333164030363015

Crawford, J. R., Henry, J. D., Crombie, C., & Taylor, E. P. (2001). Brief report:
Normative data for the HADS from a large non-clinical sample. *British
Journal of Clinical Psychology*, 40, 429-434. doi:
10.1348/014466501163904

Department of Health. (2007). *Effective services for people with severe mental
illness*. London: DoH

Derogatis, L., & Melisaratos, N. (1983). The Brief Symptom Inventory: An
introductory report. *Psychological Medicine*, 13, 595-605.
doi:10.1017/S0033291700048017

Derogatis, L. (1993). *Brief Symptom Inventory: Administering, scoring and
procedures manual* (3rd ed.). Minneapolis: National Computer Systems.

Derogatis, L. R., Lipman, R. S., & Covi, L. (1973). SCL-90: an outpatient psychiatric
rating scale-preliminary report. *Psychopharmacology Bulletin*, 9, 13-28.

Duignan, I., & Mitzman, S. (1994). Measuring individual change in patients
receiving time-limited cognitive analytic group therapy. *International
Journal of Short-Term Psychotherapy*, 9, 151-160.

Evans, C., Margison, F., & Barkham, M. (1998). The contribution of reliable and
clinically significant change methods to evidence-based mental health
Evidence Based Mental Health, 1, 70-72. doi: 10.1136/ebmh.1.3.70

Finn, S. E., & Tonsager, M. E. (1997). Information-gathering and therapeutic models
of assessment: Complementary paradigms. *Psychological Assessment*, 9,
374-385. doi:10.1037/ 1040-3590.9.4.374

- Fisher, P. M., Winne, P. H., & Ley, R. G. (1993). Group therapy for adult women survivors of child sexual abuse: Differentiation of completers versus dropouts. *Psychotherapy: Theory, Research, Practice, Training*, 30, 616-624. doi:10.1037/0033-3204.30.4.616
- Francis, V. M., Rajan, P. & Turner, N. (1990) British community norms for the Brief Symptom Inventory. *British Journal of Clinical Psychology*. 29, 115-116. doi: 10.1111/j.2044-8260.1990.tb00857.x
- Freeman, K. A., & Morris, T. L. (2001). A review of conceptual models explaining the effects of child sexual abuse. *Aggression and Violent Behaviour*, 6, 357-373. doi: 10.1016/S1359-1789(99)00008-7
- Goldberg, D. P., & Hillier, V. F. (1979). A scaled version of the General Health Questionnaire. *Psychological Medicine*, 9, 139-145. doi: 10.1017/S0033291700021644
- Goldberg, D. P., Gater, R., Sartorius, N., Ustun, T. B., Piccinelli, M., Gureje, O., & Rutter, C. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine*, 27, 191-197. doi: 10.1017/S0033291796004242
- Hagan, T., & Gregory, K. (2001). Group work with survivors of childhood sexual abuse. In P. H. Pollock (Ed.), *Cognitive Analytic Therapy for Adult Survivors of Childhood Abuse*. (pp. 190-205). Chichester: Wiley.
- Hazzard, A., Rogers, J., & Angert, L. (1993). Factors affecting group therapy outcome for adult sexual abuse survivors. *International Journal of Group Psychotherapy*, 43, 453-467.

Hepple, J. (2012). Cognitive-analytic therapy in a group: Reflections on a dialogic approach. *British Journal of Psychotherapy*, 28, 474-495. doi:

10.1111/j.1752-0118.2012.01312.x

Holloway, F. (2002). Outcome measurement in mental health - welcome to the revolution. *British Journal of Psychiatry*, 181, 1-2. doi: 10.1192/bjp.181.1.1

Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureno, G., & Villesenor, V. S. (1988).

Inventory of interpersonal problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology*, 56, 885-892. doi:

10.1037/0022-006X.56.6.885

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, 12-19. doi: 10.1037/0022-

006X.59.1.12

Kellett, S. (2007). A time series evaluation of the treatment of histrionic personality disorder with cognitive analytic therapy. *Psychology and*

Psychotherapy: Theory, Research and Practice, 80, 389-405. doi:

10.1348/147608306X161421

Kenny, D. A., Mannetti, L., Pierro, A., Livi, S., & Kashy, D. A. (2002) The statistical analysis of data from small groups. *Journal of Personality and Social*

Psychology. 83,126–137. doi: 10.1037/0022-3514.83.1.126

Kessler, M. R. H., White, M. B., & Nelson, B. S. (2003). Group treatments for women sexually abused as children: A review of the literature and

recommendations for future outcome research. *Child Abuse & Neglect*, 27,

1045-1061. doi: 10.1016/S0145-2134(03)00165-0

Kubany, E. S., Haynes, S. N., Abueg, F. R., Manke, F. P., Brennan, J. M. & Stahura, C.

(1996) Development and validation of the Trauma-Related Guilt Inventory (TRGI). *Psychological Assessment*, 8, 428-444. doi: 10.1037/1040-3590.8.4.428

Lau, M., & Kristensen, E. (2007). Outcome of systemic and analytic group psychotherapy for adult women with history of intrafamilial childhood sexual abuse: A randomised controlled study. *Acta Psychiatrica Scandinavica*, 116, 96-104. doi: 10.1111/j.1600-0447.2006.00977.x

Liem, J., & Boudewyn, A. (1999). Contextualizing the effects of childhood sexual abuse on adult self and social functioning: An attachment theory perspective. *Child Abuse and Neglect*, 23, 1141-1157. doi: 10.1016/S0145-2134(99)00081-2

Llewelyn, S. (1997). Therapeutic approaches for survivors of childhood sexual abuse: A review. *Clinical Psychology and Psychotherapy*, 4, 32-41. doi: 10.1002/(SICI)1099-0879(199703)4:1<32::AID-CPP111>3.0.CO;2-9

Llewelyn, S. (2002). Therapeutic challenges in work with childhood sexual abuse survivors. *Brief Treatment and Crisis Intervention*, 2, 123-134; doi:10.1093/brief-treatment/2.2.123

Lorentzen, S. & Høglend, P. (2008). Moderators of the effect of treatment length in long term psychodynamic group psychotherapy. *Psychotherapy & Psychosomatics*, 77, 321-322. doi: 10.1159/000147946

Lubin, H., Loris, M., Burt, J., & Johnson, D. R. (1999). Efficacy of psychoeducational group therapy in reducing symptoms of posttraumatic stress disorder among multiply traumatized women. *American Journal of Psychiatry*, 155, 1172-1177.

- Lueger, R. J., & Barkham, M. (2010). Using benchmarks and benchmarking to improve quality of practice and service. In M. Barkham, G. Hardy & J. Mellor-Clark (Eds.), *Developing and delivering practice-based evidence*. (pp.223-256). Chichester: John Wiley & Sons.
- Lundqvist, G., & Ojehagen, A. (2001). Childhood sexual abuse: An evaluation of a two-year group therapy in adult women. *European Psychiatry*, 16, 64-67. doi: 10.1016/S0924-9338(00)00537-X
- Manglio, R. (2009). The impact of child sexual abuse on health: A systematic review or reviews. *Clinical Psychology Review*, 29, 647–57. doi: 10.1016/j.cpr.2009.08.003
- Maple, N. & Simpson, I. (1995). CAT in groups. In A. Ryle (Ed.), *Cognitive analytic therapy: Developments in theory and practice*. (pp. 77-102). Chichester: John Wiley & Sons.
- Martsof, D. S., & Draucker, C. B. (2005). Psychotherapy approaches for adult survivors of childhood sexual abuse: An integrative review of outcomes research. *Issues in Mental Health Nursing*, 26, 801-825. doi: 10.1080/01612840500184012
- Mohr, D. C. (1995). Negative outcome in psychotherapy: A critical review. *Clinical Psychology: Science and Practice*, 2, 1–27. doi:10.1111/j.1468-2850.1995.tb00022.x
- Montori, V. M., & Guyatt, G. H. (2001). Intention-to-treat principle. *Canadian Medical Association Journal*, 165, 1339–1341.
- O’Leary, P. & Gould, N. (2010). Exploring coping factors amongst men who were sexually abused in childhood. *British Journal of Social Work*, 40, 2669-2686. doi:10.1093/bjsw/bcq098

- Pereda, N., Guilera, G., Forns, M., & Gomez-Benito, J. (2009). The prevalence of child sexual abuse in community and student samples: A meta-analysis. *Clinical Psychology Review, 29*, 328-338. doi: 10.1016/j.cpr.2009.02.007
- Price, J. L., Hilsenroth, M. J., Petretic-Jackson, P. A., & Bonge, D. (2001). A review of individual psychotherapy outcomes for adult survivors of childhood sexual abuse. *Clinical Psychology Review, 21*, 1095-1121. doi: 10.1016/S0272-7358(00)00086-6
- Putnam, F. W. (2003). Ten-year research update review: Child sexual abuse. *Journal of the American Academy of Child & Adolescent Psychiatry, 42*, 269-276.
- Ramchandani, P. and Jones, D. P. H. (2003). Treating psychological symptoms in sexually abused children: From research findings to service provision. *The British Journal of Psychiatry, 183*, 484-490. doi: 10.1192/03-99
- Rosenberg, M. (1989). *Society and the adolescent self-image*. Revised edition. Middletown, CT: Wesleyan University Press.
- Roth, A., & Fonagy, P. (2005). *What works for whom? A critical review of psychotherapy research* (2nd ed.). New York: Guildford Press.
- Rutter, M. (2007). Commentary: Resilience, competence, and coping. *Child Abuse & Neglect, 31*, 205-209. doi: 10.1016/j.chiabu.2007.02.001
- Ryan, M., Nitsun, M., Gilbert, L., & Mason, H. (2005). A prospective study of the effectiveness of group and individual psychotherapy for women CSA survivors. *Psychology and Psychotherapy: Theory, Research and Practice, 78*, 465-479. doi: 10.1348/147608305X42226
- Ryle, A. (1985). Cognitive theory, object relations and the self. *British Journal of Medical Psychology, 58*, 1-7. doi: 10.1111/j.2044-8341.1985.tb02608.x

Ryle, A. (1995). *Cognitive analytic therapy: Developments in theory and practice*.
Chichester: John Wiley.

Ryle, A. (1997). *Cognitive Analytic Therapy and Borderline Personality Disorder: The Model and the Method*. Chichester: Wiley & Sons.

Ryle, A., & Kerr, I. (2002). *Introducing cognitive analytic therapy: Principles and practice*. Chichester: John Wiley & Sons.

Ryle, A., Kellett, S., Hepple, J., & Calvert, R. (2014). Cognitive analytic therapy at 30. *Advances in Psychiatric Treatment*, 20, 258-268. doi:
10.1192/apt.bp.113.011817

Savard, J., Laberge, B., Gauthier, J. G., Ivers, H., & Bergeron, M. G. (1998). Evaluating anxiety and depression in HIV-infected patients. *Journal of Personality Assessment*, 7, 349-367.

Saxe, B. J., & Johnson, S. M. (1999). An empirical investigation of group treatment for a clinical population of adult female incest survivors. *Journal of Child Sexual Abuse*, 8, 67-88.

Shadish, W. R., Matt, G. E., Navarro, A. M., Siegle, G., Crits-Christoph, P., Hazelrigg, M. D...Weiss, B. (1997). Evidence that therapy works in clinically representative conditions. *Journal of Consulting and Clinical Psychology*, 65, 355-365. doi: doi:10.1037/0022-006X.65.3.355

Silber, E., & Tippet, J. (1965). Self-esteem: Clinical assessment and measurement validation. *Psychological Reports*, 16, 1017-1071.

Stiles, W., Barkham, M., Mellor-Clark, J., & Connell, J. (2008). Effectiveness of cognitive-behavioural, person-centered, and psychodynamic therapies in UK primary care routine practice: Replication with a larger sample. *Psychological Medicine*, 36, 677-688. doi: 10.1017/S0033291707001511

- Stowell-Smith, M., Gopfert, M., & Mitzman, S. (2001). Group CAT for people with a history of sexual victimisation. In P. Pollock (Ed.), *Cognitive analytic therapy for adult survivors of childhood abuse*. (pp. 206-230) Chichester: Wiley.
- Talbot, N. L., Houghtalen, R. P., Duberstein, P. R., Cox, C., Giles, D. E., & Wynne, L. C. (1999). Effects of group treatment for women with a history of childhood sexual abuse. *Psychiatric Services*, 50, 686-692.
- Taylor, J. E., & Harvey, S. T. (2010). A meta-analysis of the effects of psychotherapy with adults sexually abused in childhood. *Clinical Psychology Review*, 30, 749-767. doi: 10.1016/j.cpr.2010.05.008
- Trask, E. V., Walsh, K., & DiLillo, D. (2011). Treatment effects for common outcomes of child sexual abuse: A current meta-analysis. *Aggression and Violent Behavior*, 16, 6-19. doi: 10.1016/j.avb.2010.10.001
- Waller, G. (2009). Evidence-based treatment and therapist drift. *Behaviour Research and Therapy*, 47, 119-127. doi: 10.1016/j.brat.2008.10.018
- Westbrook, D., & Kirk, J. (2005). The clinical effectiveness of cognitive behaviour therapy: Outcome for a large sample of adults treated in routine practice. *Behaviour, Research and Therapy*, 43, 1243-1261. doi: 10.1016/j.brat.2004.09.006
- Westbury, E., & Tutty, L. M. (1999). The efficacy of group treatment for survivors of childhood abuse. *Child Abuse & Neglect*, 23, 31-44. doi: 10.1016/S0145-2134(98)00109-4
- Willmott, S. A., Boardman, J. A. P., Henshaw, C. A. & Jones, P. W. (2004). Understanding General Health Questionnaire (GHQ-28) score and its threshold. *Social Psychiatry and Psychiatric Epidemiology*. 39, 613-617. doi: 10.1007/s00127-004-0801-1

World Health Organisation. (1999). *Report of the Consultation on Child Abuse Prevention*. Geneva.

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, 67, 361-370. doi: 10.1111/j.1600-0447.1983.tb09716.x

Figure 1; diagram to illustrate flow of patients through the study

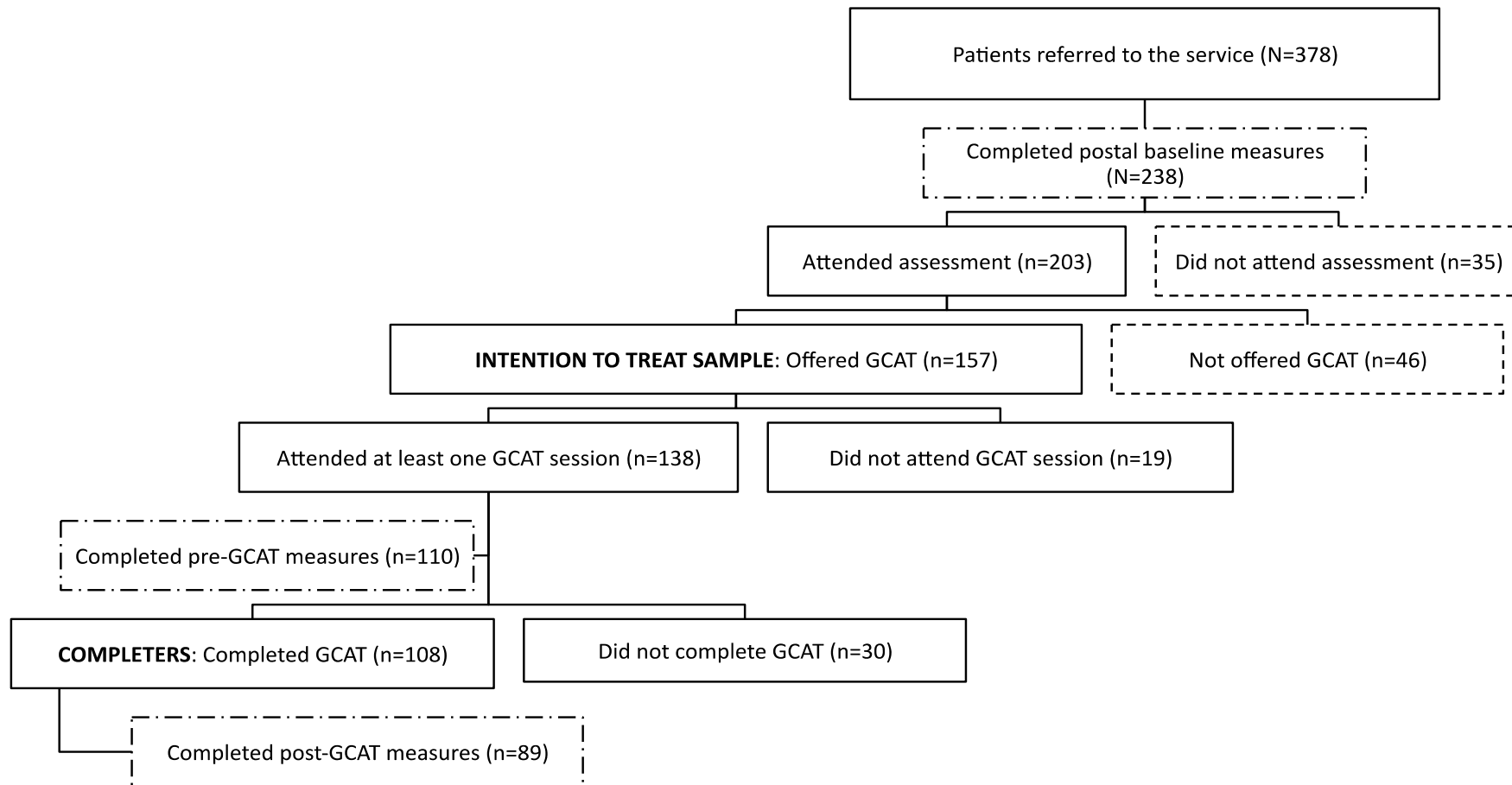


Table 1; summary of evidence used to calculate reliable and clinically significant change

Measure	Norm N	Norm Mean	Norm SD	Reliability co-efficient	Reliable Change Index value	Clinical significance cut-off score	Clinical significance criterion source
BSI-GSI ¹	252 ^a 8 ^b 376 ^c	1.66 1.80 0.44	0.83 1.13 0.47	0.90*	0.79	0.78	Externally derived (Derogatis, 1993)
IIP-32 ²	76 ^d 45 ^e	1.47 0.95	0.65 0.52	0.70*	0.99	1.39	Jacobson & Truax (1991), criteria 'C'
RSES ³	Not available	Not available	Not available	0.85*	5.86	25.22	Jacobson & Truax (1991), criteria 'A'
HADS ⁴							
<i>Anxiety</i>	1792 ^f	3.68	3.07	0.89**	4.27	11	Externally derived (Crawford, Henry, Crombie & Taylor, 2001)
<i>Depression</i>	1792 ^f	6.14	3.76	0.92**	3.66	11	
GHQ-28 ⁵	1670 ^g	5.68	6.15	0.90*	7.67	7	Externally derived (Goldberg et al., 1997)

¹Brief Symptom Inventory – Global Severity Index ²Inventory of Interpersonal Problems-32 ³Rosenberg Self-Esteem Scale ⁴Hospital Anxiety and Depression Scale ⁵General Health Questionnaire

^aRyan, 2007; clinical sample ^bClarke & Llewelyn, 1994; clinical sample ^cFrancis, Rajan, & Turner, 1990; non-clinical sample ^dBarkham, Hardy, & Startup, 1996; clinical sample ^eBarkham, Hardy, & Startup, 1996; non-clinical sample ^fCrawford, Henry, Crombie, & Taylor, 2001 non-clinical sample ^gWillmott, Boardman, Henshaw, & Jones, 2004; non-clinical sample

*Published test-retest co-efficient **No published test-retest reliability co-efficients, therefore published Cronbach α are used to derive the RCI.

Table 2; demographic characteristics for patients not offered GCAT and offered GCAT

Characteristic	Patients referred to the service and not offered GCAT (n=221)	Patients referred to the service and offered GCAT (n=157)	Mann Whitney U Test / Chi-squared test
Age (years)	34.51 ± 11.74 (17-87 years)	34.65 ± 10.67 (18-64 years)	Z=-0.35, p=0.72
Ethnicity			
<i>White</i>	164 (74%)	138 (88%)	χ^2 (1, N=317) =0.005, p=0.94
<i>Non-White</i>	8 (4%)	7 (4%)	
<i>Unknown</i>	49 (22%)	12 (8%)	
Relationship status			
<i>In a relationship</i>	77(35%)	86 (55%)	χ^2 (1, N=325) =6.88, p=0.009*
<i>Not in a relationship</i>	100 (45%)	62 (39%)	
<i>Unknown</i>	44 (20%)	9 (6%)	
Employment status			
<i>Paid employment</i>	39 (18%)	44 (28%)	χ^2 (2, N=271) =2.53, p=0.28
<i>Unemployed</i>	84 (38%)	66 (42%)	
<i>Other (e.g. studying, retired)</i>	23 (10%)	15 (10%)	
<i>Unknown</i>	75 (34%)	32 (20%)	

*p < .01 significant

Table 3; *outcome scores at assessment for patients who returned assessment outcome booklets and were not offered GCAT and patients who returned assessment outcome booklets and were offered GCAT*

Measure	Patients who returned baseline measures and not offered GCAT (n=81)			Patients who returned baseline measures and offered GCAT (n=157)			Mann Whitney <i>U</i> test
	N	Mean	SD	N	Mean	SD	
BSI-GSI ¹	73	2.21	0.90	154	2.28	0.87	<i>Z</i> =-0.98, <i>p</i> =0.43
IIP-32 ²	72	1.99	0.65	148	1.97	0.62	<i>Z</i> =-0.39, <i>p</i> =0.69
RSES ³	71	16.57	5.47	149	15.30	4.96	<i>Z</i> =-5.35, <i>p</i> <0.001*
HADS ⁴							
<i>Anxiety</i>	76	14.06	4.45	153	14.01	4.09	<i>Z</i> =-0.04, <i>p</i> =0.97
<i>Depression</i>	76	11.21	3.91	153	11.29	4.64	<i>Z</i> =-0.54, <i>p</i> =0.59
GHQ-28 ⁵	70	16.27	8.10	152	16.58	8.21	<i>Z</i> =-0.94, <i>p</i> =0.35

Note. n ranged due to missing data on some measures

¹Brief Symptom Inventory – Global Severity Index ²Inventory of Interpersonal Problems-32 ³Rosenberg Self-Esteem Scale ⁴Hospital Anxiety and Depression Scale ⁵General Health Questionnaire

*p < .001 significant

Table 4; *demographic characteristics, presenting problems and pre-GCAT scores for the non-completer and completer samples*

	Non-completers (N=49)	Completers (N=108)	Mann Whitney U test / Chi Squared test / Independent samples t test
Age (years)	33.01 ±11.23 (18-64 years)	35.39 ±10.36 (19-60 years)	$z = -1.48, p = 0.14$
Ethnicity			$\chi^2 (1, N=145) = 0.034, p = 0.85$
White	44 (90%)	94 (87%)	
Non-White	2 (4%)	5 (5%)	
Unknown	3 (6%)	9 (8%)	
Relationship status			$\chi^2 (1, N=148) = 0.001, p = 0.97$
In a relationship	28 (57%)	58 (54%)	
Not in a relationship	20 (41%)	42 (39%)	
Unknown	1 (2%)	8 (7%)	
Employment status			$\chi^2 (2, N=125) = 0.401, p = 0.82$
Paid employment	14 (29%)	30 (38%)	
Unemployed	21 (43%)	45 (42%)	
Other (e.g. studying, retired)	6 (12%)	9 (8%)	
Unknown	8 (16%)	42 (22%)	
Internalising difficulties			$\chi^2 (1, N=157) = 0.625, p = 0.73$
Recorded at assessment	47 (96%)	100 (93%)	
Not recorded at assessment	2 (4%)	8 (7%)	
Externalising difficulties			$\chi^2 (1, N=157) < 0.001, p = 1.00$
Recorded at assessment	14 (29%)	31 (29%)	
Not recorded at assessment	35 (71%)	77 (71%)	
Substance use			$\chi^2 (1, N=157) = 0.043, p = 0.79$
Recorded at assessment	6 (12%)	12 (11%)	
Not recorded at assessment	43 (88%)	96 (89%)	
Self-harm behaviours			$\chi^2 (1, N=157) = 0.57, p = 0.46$
Recorded at assessment	17 (35%)	31 (29%)	

<i>Not recorded at assessment</i>	32 (63%)	77 (71%)	
Sexual difficulties			
<i>Recorded at assessment</i>	3 (6%)	13 (12%)	$\chi^2 (1, N=157)=1.288, p=0.39$
<i>Not recorded at assessment</i>	46 (94%)	95 (88%)	
Re-victimisation experiences			
<i>Recorded at assessment</i>	9 (18%)	10 (9%)	$\chi^2 (1, N=157)=2.63, p=0.12$
<i>Not recorded at assessment</i>	40 (82%)	98 (91%)	

	N	Mean	SD	N	Mean	SD	
Measure							
BSI-GSI ¹	47	2.11	0.89	103	2.18	0.92	$t(148)=0.282, p=0.78$
IIP-32 ²	46	1.96	0.70	101	1.91	0.62	$t(145)=-0.52, p=0.60$
RSES ³	45	16.36	5.48	102	19.42	5.21	$z=-3.02, p=0.003^*$
HADS ⁴							
<i>Anxiety</i>	48	13.33	4.85	104	14.03	4.10	$z=-0.611, p=0.54$
<i>Depression</i>	48	10.08	4.76	103	11.18	4.61	$t(149)=1.345, p=0.18$
GHQ-28 ⁵	49	14.48	8.20	101	16.11	9.05	$z=-1.314, p=0.19$

Note. *n* ranged due to missing data on some measures

¹Brief Symptom Inventory – Global Severity Index ²Inventory of Interpersonal Problems-32 ³Rosenberg Self-Esteem Scale ⁴Hospital Anxiety and Depression Scale ⁵General Health Questionnaire

* $p < .01$

Table 5; Pre/Post GCAT outcomes, effect sizes, and reliable and clinically significant change for the assigned-to-GCAT and completer samples

Measure	Assigned-to-GCAT sample (N=157)						Completer sample (N=108)									
	N	pre-GCAT mean (sd)	post-GCAT mean (sd)	pre-post change score	95% C. I.	effect size	N	pre-GCAT mean (sd)	post-GCAT mean (sd)	pre-post change score	95% C. I.	effect size	RCSI ¹	RI ²	RD ³	RCSD ⁴
BSI-GSI	154	2.10 (0.90)	1.93 (1.00)	0.17	0.00- 0.46	0.19	103	2.18 (0.92)	1.87 (1.08)	0.31	0.06- 0.61	0.34	11 (11%)	10 (10%)	6 (6%)	1 (1%)
IIP-32	148	1.93 (0.65)	1.78 (0.69)	0.15	0.00- 0.46	0.23	101	1.91 (0.62)	1.69 (0.68)	0.22	0.07- 0.63	0.35	0	11 (11%)	1 (1%)	1 (1%)
RSES	149	18.49 (5.47)	16.42 (5.76)	2.07	0.15- 0.61	0.38	102	19.42 (5.21)	16.41 (5.91)	3.01	0.29- 0.86	0.58	14 (14%)	11 (11%)	3 (3%)	0
HAD-A	153	13.81 (4.35)	12.53 (4.60)	1.28	0.06- 0.51	0.29	104	14.03 (4.10)	12.26 (4.47)	1.77	0.15- 0.71	0.43	10 (10%)	13 (13%)	2 (2%)	2 (2%)
HAD-D	152	10.83 (4.67)	9.39 (5.18)	1.44	0.08- 0.54	0.31	103	11.18 (4.61)	9.13 (5.37)	2.05	0.17- 0.72	0.44	16 (16%)	3 (3%)	2 (2%)	0
GHQ-28	154	15.57 (8.78)	12.75 (9.29)	2.82	0.09- 0.55	0.32	101	16.11 (9.05)	11.97 (9.72)	4.14	0.17- 0.74	0.46	16 (16%)	7 (7%)	2 (2%)	5 (5%)

Note. *n* ranged due to missing data on some measures

¹Reliable and Clinically Significant Improvement ²Reliable Improvement ³Reliable Deterioration ⁴Reliable and Clinically Significant Deterioration

Figure 2; scatterplot of individual BSI outcomes for GCAT completers completing pre/post outcomes (n=103)

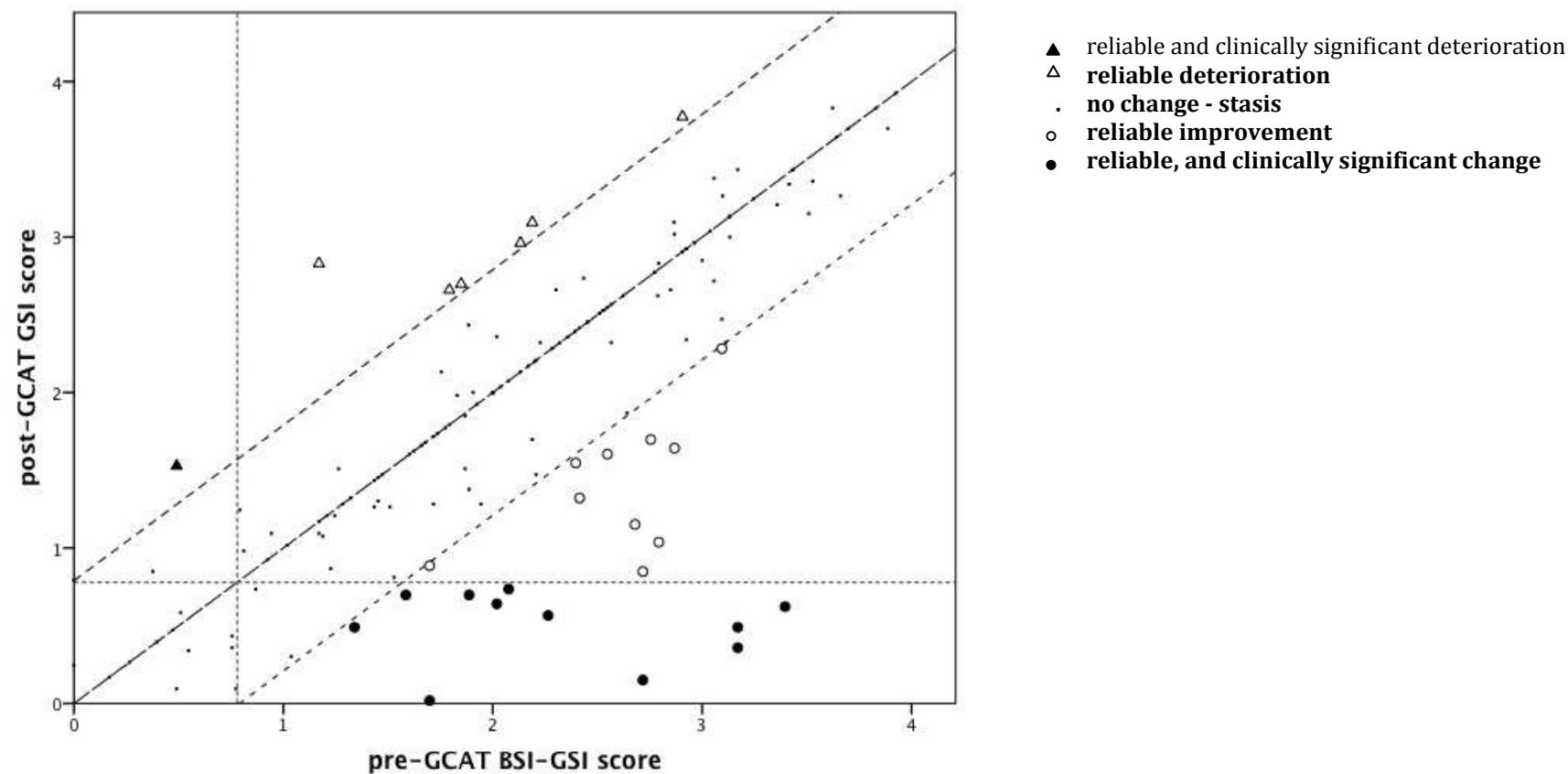


Figure 3; forest plot of group interventions for CSA survivors

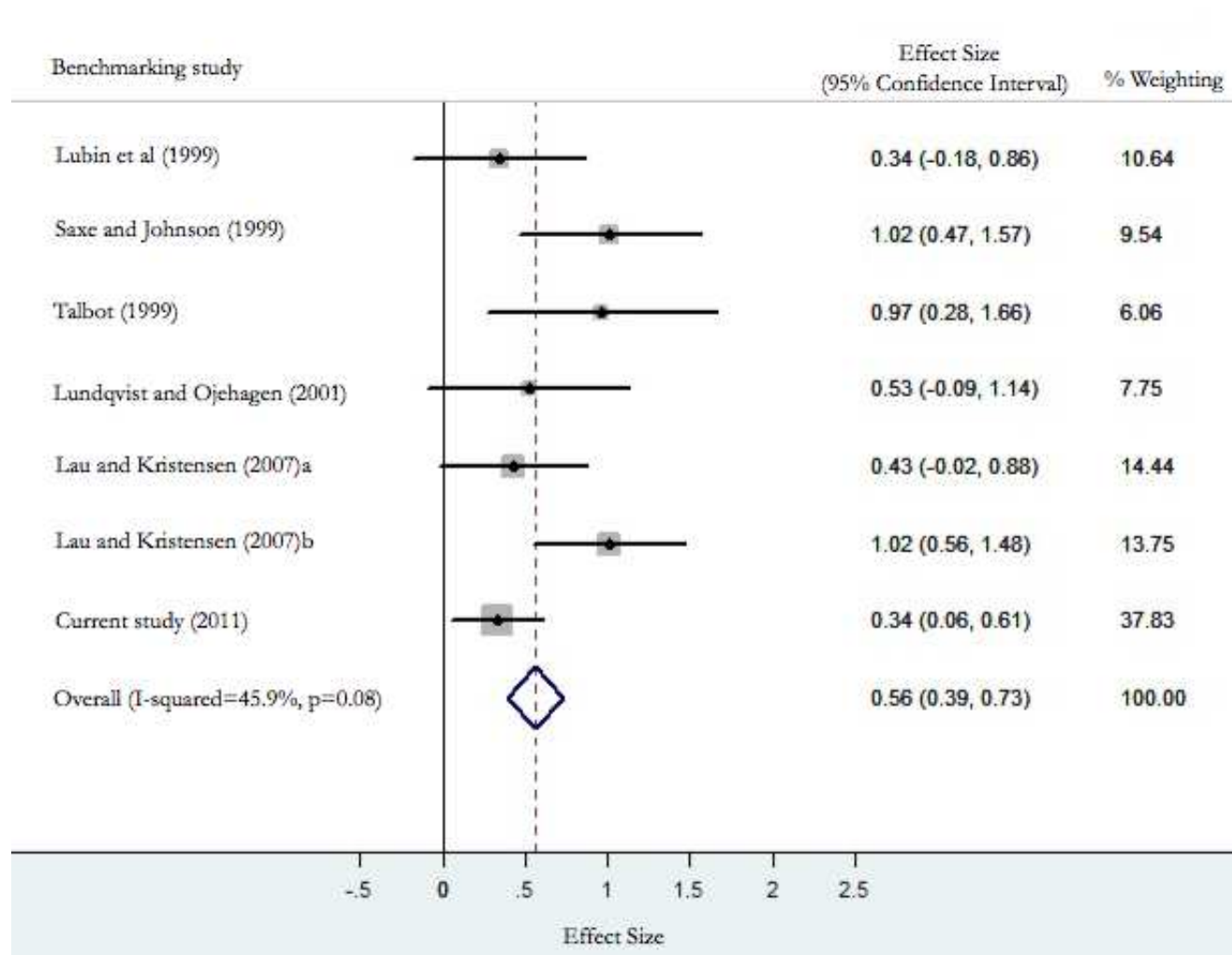


Table 6; pre- and post-therapy scores and effect sizes for group interventions for CSA survivors

Author(s)	Group Treatment (Duration); Sample Size	Setting	Pre-therapy M (SD)	Post-therapy M (SD)	Effect size
Lubin (1998) [‡]	Trauma focused CBT (16 sessions); n=29	Community	113.31 (78.14)	86.69 (75.32)	0.34
Saxe and Johnson (1999) [†]	'Recovery' therapy (20 sessions); n=32	Clinical outpatient	1.62 (0.58)	1.03 (0.65)	1.02
Talbot (1999) [‡]	Trauma recovery therapy (10 sessions); n=20	Inpatient	57.45 (7.79)	49.90 (9.41)	0.97
Lundqvist and Ojehagen (2001) [‡]	Long-term psychodynamic therapy (2 years); n=22	Clinical outpatient	1.38 (0.80)	0.96 (0.80)	0.53
Lau and Kristensen (2007) [‡]	Analytic therapy ^a (46 sessions); n=40	Clinical outpatient	1.95 (0.75)	1.63 (0.77)	0.43
	Systemic therapy ^b (17 sessions); n=46		1.61 (0.61)	0.99 (0.69)	1.02
Current sample [†]	Cognitive Analytic Therapy (24 sessions); n=108	Clinical outptient	2.18 (0.92)	1.87 (1.08)	0.34

[‡]SCL-90, [†]BSI

