

Main Section

ONE YEAR FOLLOW-UP OF COGNITIVE THERAPY, ANALYTIC PSYCHOTHERAPY AND ANXIETY MANAGEMENT TRAINING FOR GENERALIZED ANXIETY DISORDER: SYMPTOM CHANGE, MEDICATION USAGE AND ATTITUDES TO TREATMENT

Robert C. Durham, Peter L. Fisher and Linda R. Treliving

University of Dundee, Scotland

Cathryn M. Hau

University of Edinburgh, Scotland

Karen Richard

Perth and Kinross NHS Trust, Scotland

John B. Stewart

Ayrshire and Arran NHS Trust, Scotland

Abstract. A one year follow-up is reported of a randomized clinical trial with generalized anxiety disorder (GAD) in which the main comparison was between analytic psychotherapy (AP) and cognitive therapy (CT), each delivered at weekly or fortnightly intervals over a six month period. CT was found to be significantly more effective than AP. However, GAD is a chronic and relapsing condition and follow-up data are needed to assess the durability of improvement and the possibility that the benefits of analytic psychotherapy may be more apparent over the longer term. Data were collected in three areas: (1) symptomatology and overall improvement; (2) medication usage and contact with GPs; and (3) attitudes to therapy. CT was clearly superior to AP on the

Reprint requests to Rob Durham, Department of Psychiatry, Ninewells Hospital & Medical School, Dundee DD1 9SY, Scotland, U.K.

© 1999 British Association for Behavioural and Cognitive Psychotherapies

main outcome measures and only a minority of AP patients made significant improvements. CT but not AP was associated with significant reductions in medication usage, and patients receiving CT were generally more positive about treatment received. Differences between treatments were less evident in the less intensive treatment condition where overall results were relatively poor. The most positive outcomes were achieved in the more intensive CT condition in which approximately two-thirds of patients achieved clinically significant improvements.

Keywords: Generalized anxiety, follow-up, clinical trial, psychotherapy.

Introduction

This paper concerns outcome at one year follow-up of a randomized controlled trial comparing cognitive therapy (CT), analytic psychotherapy (AP) and anxiety management training (AMT) for generalized anxiety disorder (GAD) (Durham *et al.*, 1994). Patients allocated to CT and AP were in addition randomly allocated to either high or low contact and were treated by committed therapists with broadly similar levels of training and experience. Patients receiving AMT were treated by trainee psychiatrists gaining experience in cognitive behavioural therapy and were seen only in the low contact condition. CT was found to be significantly more effective than AP across a broad range of measures, with about 50% of patients markedly improved at six month follow-up. There was no significant effect for level of contact. Patients receiving AMT showed improvements that were similar to CT though less broadly-based and with lower proportions showing clinically significant change at six month follow-up. These results are broadly in line with conclusions from several reviews of outcome studies in this area (cf. Chambless & Gillis, 1993; Durham & Allan, 1993; Borkovec & Whisman, 1997): cognitive behavioural therapies are most effective, about 50% of patients achieve high endstate functioning and post treatment gains are largely maintained at six month follow-up.

Extended follow-up data from clinical trials of this kind are important for several reasons. First, despite promising results from outcome studies GAD is known to be a chronic, relapsing condition (Rickels & Schweizer, 1990) and it would be unreasonable to assume that short-term treatment gains are a reliable guide to longer-term improvement. Evidence for the maintenance of treatment gains is mixed. Barlow, Rapee and Brown (1992), for example, reported a drop in the overall proportion of treatment responders from 55% at post-treatment to 32% at two year follow-up. Both Borkovec and Costello (1993) and White (1998), on the other hand, found treatment gains largely maintained at one and two year follow-up. However, if treatment gains are sustained it is important to know whether or not this has been achieved with or without the addition of a further course of psychological therapy and/or medication. The evidence here is limited but suggests that for a significant number of patients initial treatment gains may be maintained only with continued professional help. Power, Jerrom, Simpson, Mitchell and Swanson (1989) found relatively high proportions of patients required further treatment (either psychological or medication) in the year following treatment although the proportions were significantly higher in the nonpsychological treatment conditions (30%, 70% and 55% of cognitive therapy, diazepam and pill

placebo groups respectively). Butler, Fennell, Robson and Gelder (1991) found that 42% and 79% of cognitive-behavioural and behaviour therapy patients respectively required further therapy 11–24 months following treatment.

Second, the comparative efficacy of different types of psychotherapy may be more apparent over the longer term. While it is commonly believed that there are relatively few differences in efficacy between psychological therapies, and there is some support for this in clinical trials with GAD (cf. Borkovec & Mathews, 1998; White, Keenan, & Brooks, 1992), there is also evidence that the timing of outcome assessments is of importance. Durham and Turvey (1987), for example, in a comparison of cognitive therapy and behaviour therapy with GAD, found little difference in outcome at the end of treatment but a clear advantage for cognitive therapy at six month follow-up. The timing of assessments may be of particular importance in comparisons of cognitive and analytic psychotherapy since the *primary* focus of treatment, in the former, is on acquiring skills for achieving better control over distressing symptoms, and in the latter, on understanding the origins of that distress and its expression in current relationships. Cheng and Baxter (1995) raise fundamental doubts about the validity of clinical trials such as Durham et al. (1994) on the grounds, amongst others, that the methodology assumes that different “psychological treatments will effect similar changes within the same time framework”. Certainly it is unreasonable to assume that the course of symptomatic improvement will follow the same pattern in the two types of treatment and evidence of symptomatic relief following analytic psychotherapy may be more apparent in the longer term.

Third, long term follow-up may clarify cost-effectiveness issues. Psychological therapies such as CBT are markedly more expensive than medication and it is generally agreed that specialized training is required to enable mental health professionals to implement treatment effectively, particularly with more severe and chronic problems (Roth & Fonagy, 1996). The greater effectiveness of CT over AMT in Durham et al. (1994) is consistent with this belief although it must be admitted that there has been relatively little empirical investigation of this issue (cf. Burns & Nolen-Hoeksema, 1992). The extra costs of psychological therapy may be hard to justify if therapeutic gains are not sustained, particularly in a clinical disorder such as GAD where the proportion of patients who fail to make clinically significant changes is relatively high. On the other hand, if it can be shown that more expensive treatment results in an enduring reduction in demand on healthcare resources, in terms of medication usage, consultation rates, and so forth, then the costs of psychological therapy begin to look increasingly worthwhile. The costs of therapy, of course, are proportional to the length of treatment. Although initial results from Durham et al. (1994) suggested no difference in outcome between high and low contact conditions at six month follow-up, theoretical considerations would suggest that dose effect relationships may be increasingly apparent over the longer term as the effects of more intensive learning in acquiring new coping skills prove increasingly advantageous in reducing vulnerability to adverse life events.

This study reports one year follow-up data in three areas: (1) questionnaire measures of symptomatology and overall improvement, (2) medication usage and frequency of contact with GPs and (3) attitudes to therapy.

Method

Only information necessary for understanding the analyses reported in this paper is provided here. Fuller details of the methodology of the clinical trial can be found in Durham et al. (1994).

Overview

One hundred and ten patients meeting DSM-111-R diagnostic criteria for a primary diagnosis of generalized anxiety disorder (GAD) were allocated randomly to either CT, AP or AMT and to one of several therapists delivering each treatment. Those in CT and AP were in addition randomly allocated to either “low” or “high” contact (8–10 or 16–20 hour-long sessions over six months). All patients allocated to AMT were in the “low” contact condition. Of the 110 patients offered treatment, 11 failed to attend and 19 (19% of treatment starters) subsequently dropped out (24% from AP, 10% from CT and 27% from AMT) leaving 80 completers. A variety of self-report scales were administered before and after treatment and at six month follow-up. Assessment of progress at one year follow-up comprised (1) the main self-report outcome measures, (2) data on medication usage and frequency of primary care consultation requested directly from each patient’s GP (primary care physician), and (3) a structured interview by the independent assessor who was blind to treatment allocation. The structured interview covered diagnostic status, social adjustment and attitude to treatment. However, it was only administered for the first half of patients recruited for the trial (47% of total completers) as the interviewer (KR) left for another post. Although the interview data did not provide an adequate basis for reporting reliable information on diagnostic status and social adjustment, it was felt that the limited data on attitudes to treatment were of particular interest and worth reporting.

Patient characteristics

The 110 patients recruited for the trial were judged to have a primary diagnosis of GAD as defined by DSM-111-R using the revised version of the Anxiety Disorders Interview Schedule (DiNardo & Barlow, 1988). Fifty-one per cent had received previous psychiatric treatment and 66% were taking some form of psychotropic medication. There was a significant bias in the sample towards lower socioeconomic status in comparison with the social class distribution of the adult UK population as determined by the 1981 census (1 = 5%, 2 = 23%, 3 = 48%, 4 = 18%, 5 = 6%, $\chi^2_4 = 9.8$, $p < .05$). Patterns of comorbidity were determined for axis 1 diagnoses by the independent assessor and for axis 11 diagnoses by the therapists in the study. In summary, 80% were assigned at least one additional axis 1 diagnosis and 46% were diagnosed as having co-existing personality disorders, predominantly of the dependent or avoidant kind.

Treatment

Therapies in the CT and AP conditions were delivered by therapists committed to their respective method of treatment and with equivalent years of experience. CT was delivered by two clinical psychologists and analytic psychotherapy by two consultant psychiatrists (a third consultant psychiatrist saw three cases while one of the former was

on leave). AMT was delivered by seven trainee psychiatrists supervised by the senior author. CT was manualized and based on the approach described by Beck and Emery (1985). AP was not manualized in a formal sense but the therapists met on a regular basis throughout the trial for mutual supervision. It was carried out by attempting to explore and understand the occurrence of the presenting symptom in (a) its current relationship context, (b) the individual's developmental context, and (c) in terms of the transference and resistance aspects of the therapeutic relationship. AMT was based on Clark (1989) and Meichenbaum (1985) and essentially involved a structured approach to treatment, involving education about anxiety and the acquisition of more adaptive coping strategies with the main emphasis on problem-solving and relaxation training.

Expectations of improvement and perceived suitability of treatment were each rated by patients on a 0–8 point scale at the end of the third session of treatment. One way ANOVAs indicated a significant difference across treatment groups in both expectations of improvement ($F_{2,79} = 5.3, p < .01$) and perceived suitability of treatment ($F_{2,79} = 9.7, p < .001$). Post hoc Tukey tests (HSD) indicated that CT patients had significantly higher expectations than those in AP ($p < .05$) and both CT and AMT patients perceived treatment as being significantly more suitable than those in AP ($p < .05$).

Measures

Self-report. Ratings of symptom severity consisted of the Brief Symptom Inventory (BSI), a global measure of current symptomatic state, essentially a shorter version of the SCL-90 (Derogatis & Meliseratos, 1983), the trait version of the State-Trait Anxiety Inventory (STAI-T) (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983) and 0–8 point ratings of general tension, surges of panic and irritability. Patient ratings of overall degree of improvement were made on a 6-point scale (worse, unchanged, slight, moderate, very considerable, completely better).

Medication usage and contact with general practitioner. Data on patients' use of anxiolytic medication, and number of consultations with their GP, were collected for the two six-month periods before treatment and follow-up on 54 of the 80 treatment completers (67%).

Attitude to treatment. At the end of the follow-up interview patients were asked to reflect on the therapy they had received a year previously and to give their views on the nature of the experience. Two issues were explored by means of 5-point rating scales followed by open ended questions. Patients were asked, first, "did you feel that therapy was a positive and helpful experience?" ("yes, definitely", "yes, to some extent", "yes and no", "no, not really", "no, definitely not") and, second, "how did you get on with your therapist?" ("extremely well", "reasonably well", "so so", "not very well", "not at all well"). Open-ended questions of the kind "in what ways was therapy helpful (or unhelpful)?" were then used to invite an elucidation of the rating that was made and the main comments were recorded verbatim. Whereas all patients made some comment about the helpfulness of treatment, relatively few commented on the nature of the therapeutic relationship.

Results

Self-report measures of symptom severity

The following analysis is based on the 80 patients who completed treatment. Missing data were distributed evenly across treatment conditions and averaged 0%, 11%, 13% and 25% respectively for data collection at pre-treatment, post-treatment, six month and one year follow-up. Means and standard deviations of the self-report measures are summarized in Table 1. Overall ratings of the severity of presenting problems, made by the independent assessor at pre-treatment, are also included. An identical “intention-to-treat” analysis, based on the 99 patients who entered therapy, with missing data points for the dropouts carried forward, has also been completed. This analysis, available from the senior author, gives a very similar pattern of results.

As reported earlier (Durham et al., 1994), despite randomization there was a significant difference in pre-treatment mean scores for the completers on several of the self-report measures (e.g. BSI), with the low-contact CT condition having a lower than average mean score. It is not entirely clear why this should have occurred but it may be partly due to slightly different time lags in the various treatment conditions between initial screening and subsequent assessment by postal questionnaire. It should be noted that there were no significant pre-treatment differences between conditions on the measures derived from the screening interview and it can be seen from Table 1 that assessor ratings of overall severity were broadly equivalent across groups. In order to control for possible pre-treatment differences on the self-report measures the following analyses used repeated measures analyses of covariance with the pre-treatment score as the covariate.

(1) *CT versus AP at high and low contact.* A significant main effect for treatment was found on four of the five outcome measures, with CT showing superior outcome to AP. These results are summarized in Table 2. There were no significant main effects for level of contact and no significant interaction effects for treatment and level of contact on any of the measures. Within-subjects effects over time, from post-treatment to six month and one year follow-up, were analysed using difference contrasts in which the first contrast compared means across the three levels (two levels for 0–8 ratings), the second contrast compared post treatment and six month follow-up means (post-treatment and one year follow-up for 0–8 ratings) and the third contrast compared means at one year follow-up with the average of mean scores for post-treatment and six month follow-up. There were no significant interaction effects for either treatment or level of contact with time but there were significant interaction effects for treatment and level of contact with time on two of the five measures at the comparison of one year follow-up scores with either the mean of post-treatment and six month follow-up scores in the case of STAI-T ($F_{1,43} = 4.96, p < .05$) or the post-treatment scores as in 0–8 point tension ratings ($F_{1,37} = 5.9, p < .05$). In each case, the high contact CT condition continued to improve at one year follow-up whereas the low contact CT condition showed a deterioration.

(2) *AP versus CT versus AMT at low contact.* There was a significant overall difference between treatments on two of the five outcome measures: the STAI-T ($F_{2,31} = 4.4; p < .05$) and the 0–8 rating scale measuring the severity of panic ($F_{2,29} = 6.0, p < .01$).

Table 1. Means (SD) for outcome measures before and after treatment and at 6 and 12 months follow-up

	Contact level	AP				CT				AMT			
		Before	After	6 m	12 m	Before	After	6 m	12 m	Before	After	6 m	12 m
Overall severity (0-8)	High	6.2 (0.7) n = 19	—	—	—	6.4 (0.8) n = 19	—	—	—	—	—	—	—
	Low	5.9 (1.0) n = 19	—	—	—	6.0 (0.9) n = 20	—	—	—	6.3 (0.9) n = 22	—	—	—
BSI gsi	High	1.6 (0.4) n = 14	1.5 (0.7) n = 11	1.7 (0.8) n = 11	1.6 (1.0) n = 10	1.8 (0.7) n = 15	1.0 (0.9) n = 15	1.1 (0.9) n = 15	0.8 (0.8) n = 13	—	—	—	—
	Low	1.8 (0.6) n = 15	1.6 (0.9) n = 14	1.9 (0.9) n = 13	1.6 (0.8) n = 11	1.3 (0.6) n = 20	0.9 (0.8) n = 17	0.9 (0.7) n = 15	1.1 (0.7) n = 14	1.6 (0.5) n = 16	1.1 (0.7) n = 14	1.2 (0.8) n = 14	1.2 (0.8) n = 12
STAI-T	High	60 (6) n = 14	58 (10) n = 11	57 (10) n = 12	58 (12) n = 10	61 (6) n = 15	49 (12) n = 15	49 (13) n = 15	42 (13) n = 13	—	—	—	—
	Low	61 (7) n = 15	57 (10) n = 13	60 (10) n = 14	60 (9) n = 12	55 (9) n = 20	47 (13) n = 18	43 (15) n = 14	45 (13) n = 14	57 (11) n = 16	48 (13) n = 14	50 (13) n = 14	48 (13) n = 12
Tension (0-8)	High	5.1 (1.9) n = 14	—	4.5 (1.6) n = 11	4.9 (2.7) n = 10	5.7 (1.8) n = 15	—	3.9 (2.9) n = 15	2.2 (2.0) n = 12	—	—	—	—
	Low	5.9 (1.5) n = 15	—	5.4 (2.1) n = 11	4.6 (2.2) n = 11	5.2 (1.8) n = 20	—	3.1 (2.7) n = 14	3.5 (2.7) n = 13	5.7 (2.0) n = 16	—	3.4 (2.3) n = 12	3.7 (2.9) n = 14
Panic (0-8)	High	4.7 (2.1) n = 14	—	3.8 (2.4) n = 11	4.1 (3.1) n = 10	5.6 (2.4) n = 15	—	2.8 (2.5) n = 15	1.6 (2.4) n = 12	—	—	—	—
	Low	4.8 (1.8) n = 15	—	4.2 (1.9) n = 14	4.5 (2.3) n = 10	4.4 (2.1) n = 20	—	1.8 (2.1) n = 14	1.6 (1.8) n = 14	5.1 (2.1) n = 16	—	3.6 (2.9) n = 14	3.1 (3.5) n = 13
Irritability (0-8)	High	3.9 (2.7) n = 14	—	2.5 (2.5) n = 11	3.9 (2.7) n = 11	3.6 (2.7) n = 15	—	2.6 (1.8) n = 15	1.8 (1.9) n = 12	—	—	—	—
	Low	4.7 (2.3) n = 15	—	3.9 (2.4) n = 13	3.2 (1.5) n = 11	3.8 (2.5) n = 20	—	2.3 (2.0) n = 14	2.6 (1.5) n = 14	3.6 (2.1) n = 16	—	3.1 (2.3) n = 14	1.9 (1.6) n = 12

Table 2. CT vs AP at high and low contact: Summary of between subjects treatment effects from repeated measures ANACOVAs with pre-treatment scores as covariate

	Overall			Result
	<i>F</i>	<i>df</i>	<i>p</i>	
BSI	12.9	1,41	0.001	CT < AP
STAI-T	17.4	1,43	0.001	CT < AP
Tension (0–8)	7.9	1,36	0.01	CT < AP
Panic (0–8)	15.0	1,37	0.001	CT < AP
Irritability (0–8)	3.1	1,39	n.s.	

Contrasts in the ANACOVA were set to determine the source of these differences and these revealed that both CT ($p < .01$) and to a lesser degree AMT ($p = .06$) showed more improvement than AP on the STAI-T but that only CT resulted in significantly more improvement than AP on the panic ratings.

Measures of the clinical significance of outcomes

(1) *Patient ratings of overall improvements at one year follow-up.* Overall improvement ratings made by patients who completed treatment are summarized in Table 3. The significance of differences across treatment groups was examined with χ^2 with data collapsed across “worse and no change” and “very considerable and better” on account of small cell sizes in the extreme categories. In the two high contact conditions there were significantly better ratings of improvement for CT than AP ($\chi^2_1 = 5.0, p < .05$) with 64% of CT and 36% of AP patients describing themselves as very considerably improved or better and 14% of CT and 54% of AP patients describing themselves as unchanged or only slightly improved.

For patients in the three low contact conditions there were no significant differences in ratings between AP and AMT ($\chi^2_1 = 1.0, n.s.$) but significantly better ratings of improvement for CT in comparison to both AP ($\chi^2_1 = 7.8, p < .01$) and AMT ($\chi^2_1 = 3.8, p < .05$). The proportions of patients rating themselves as very considerably improved or better was 40%, 18% and 13% for CT, AMT and AP respectively, with

Table 3. Patient ratings of overall improvement at one year follow-up

	% Treatment group in each category						
	<i>N</i>	Worse	No change	Slight	Moderate	Very considerable	Better
High contact							
AP	11	18	36	0	9	36	0
CT	14	0	7	7	21	50	14
Low contact							
AP	11	7	27	40	13	13	0
CT	14	0	5	20	35	35	5
AMT	12	6	12	37	25	12	6

Table 4. Percentage of patients achieving Jacobson criterion (c) for a return to normal functioning (excluding patients with pre-treatment scores < cut off point)

Level of contact	AP		CT		AMT	
	BSI	STAI-T	BSI	STAI-T	BSI	STAI-T
High	33 (<i>n</i> = 9)	30 (<i>n</i> = 10)	73 (<i>n</i> = 11)	67 (<i>n</i> = 12)	—	—
Low	20 (<i>n</i> = 10)	0 (<i>n</i> = 11)	11 (<i>n</i> = 9)	50 (<i>n</i> = 12)	27 (<i>n</i> = 11)	10 (<i>n</i> = 10)
Overall total	26 (<i>n</i> = 19)	14 (<i>n</i> = 21)	40 (<i>n</i> = 20)	58 (<i>n</i> = 24)	27 (<i>n</i> = 11)	10 (<i>n</i> = 10)

25%, 55% and 74% of CT, AMT and AP patients respectively rating themselves as worse, unchanged or only slightly improved. Although it can be seen from Table 3 that overall ratings of improvement were rather better in the high contact than in the low contact conditions, these differences were not significant in either AP ($\chi^2_1 = 0.5$, n.s.) or CT ($\chi^2_1 = 1.3$, n.s.).

(2) *Clinically significant change on BSI and STAI-T using Jacobson criterion c.* The methodology suggested by Jacobson and colleagues (cf. Jacobson & Truax, 1991) was used to define clinically significant change. This strategy operationalizes the criteria for a return to normative functioning as (a) a reliable change between tests and (b) a level of functioning that is outside the range of the dysfunctional population or within the range of the normal population. With measures such as the BSI and STAI-T, where there is considerable overlap between the two distributions, Jacobson criterion (c) is appropriate in which the cut-off point is defined as the mid-point between two standard deviations below the pre-treatment mean and two standard deviations above the mean for the normal population. Table 4 shows the proportion of patients in each treatment condition who achieved Jacobson criterion (c) on the BSI general severity index (cut-off point = 0.95, 0.53 points required for reliable change) and STAI-T (cut-off point = 47, 8 points required for reliable change). Patients with *pre-treatment* scores below the cut-off points, predominantly those from the low contact cognitive therapy condition, were excluded from the table. In the high contact condition it can be seen that over twice as many CT as AP patients met the criterion for clinically significant change (67–73% vs 30–33%). The results are less clear in the low contact condition with a marked difference in outcome between the BSI and STAI-T. There is a markedly superior outcome for CT patients in comparison with AP and AMT patients in terms of the STAI-T (50% vs 0% & 10%) but a generally poor outcome for all three treatments in terms of the BSI (20%, 11% and 27% for AP, CT and AMT respectively).

Medication usage and contact with general practitioner

At the start of treatment 70% of the sample were taking psychotropic medication for anxiety with no significant differences between treatment groups in the proportions taking medication (65% in AP, 66% in CT and 87% in AMT; $\chi^2_2 = 2.9$, n.s.). At one year follow-up data on medication usage were obtained on 54 of the 80 cases (67%) of whom 65% were taking medication. There was a significant difference between treatment groups at follow-up with 82% of AP and AMT patients on medication and only

Table 5. Frequency (%) of patient responses to the question “did you feel that therapy was a positive and helpful experience?”

	Yes, definitely	To some extent	Yes and no	Not really	Not at all
AP (<i>n</i> = 15)	2 (13)	4 (27)	4 (27)	2 (13)	3 (20)
CT (<i>n</i> = 16)	13 (81)	2 (12)	1 (6)	—	—

38% of CT patients ($\chi^2 = 10.7, p < .01$). There were no significant pre-treatment differences in medication usage between treatment groups for the follow-up sample (59% in AP, 57% in CT and 82% in AMT; $\chi^2 = 2.1, n.s.$) suggesting that missing data were distributed evenly in respect of medication usage between treatments.

The mean number of primary care consultations for the total sample during the six month period before treatment was 7.9 (*SD* = 3.9). One way ANACOVA revealed no significant differences between treatment groups (AP, mean 8.5, *SD* 4.2; CT mean 7.7, *SD* 4.0; AMT, mean 7.4, *SD* 3.0; $F_{2,77} = 0.5, n.s.$). Data at one year follow-up were available for 61% of the total sample, at which point there was a small but significant drop in the mean number of consultations to 6.0 (*SD* = 4.8), ($t_{48} = 2.3, p < .05$). There were no significant pre-treatment mean differences in consultations between treatment groups for the follow-up sample (AP, mean 7.7, *SD* 3.9; CT mean 7.6, *SD* 4.1; AMT, mean 7.2, *SD* 2.6; $F_{2,48} = 0.6, n.s.$) suggesting that missing data were distributed evenly between treatments in respect of this variable. Differences between treatment groups in the mean number of consultations from pre-treatment to follow-up were analysed using a one way ANOVA on difference scores and no significant differences emerged (AP, mean difference 0.9, *SD* 5.2, CT mean difference 2.3, *SD* 5.1, AMT, mean difference 1.2, *SD* 2.9; $F_{2,48} = 0.45, n.s.$).

Attitude to treatment

Approximately half of the patients in the two main treatment groups (*n* = 15, 16 & 7 in AP, CT and AMT respectively) were interviewed regarding their attitudes to treatment. There were no pre-treatment differences in severity of symptomatology between these three groups on any of the main outcome measures. Pre-treatment mean scores (and standard deviations) on the Hamilton Anxiety Rating Scale, for example, were 35.4 (5.8), 34.8 (5.9) and 37.0 (8.6) for AP, CT and AMT respectively (one way ANOVA; $F_{2,34} = 0.26; n.s.$). Due to the small amount of data from the AMT condition (three patients with one therapist and no more than one patient for four other therapists), it was decided to restrict the analysis to AP and CT. Ratings of the helpfulness of therapy are summarized in Table 5. A Mann–Whitney U test revealed that CT was rated as being significantly more helpful than AP ($U = 29, n_1 = 15, n_2 = 16, p < .001$). It can be seen from Table 5 that over 80% of CT patients, compared to less than 15% of AP patients, rated therapy as definitely helpful. Moreover, whereas all CT patients rated therapy as being helpful, at least to some degree, one third of AP patients rated therapy as being not helpful. Ratings in response to the question “how well did you

Table 6. Frequency (%) of patient responses to the question “how did you get on with your therapist?”

	Extremely well	Reasonably well	So-so	Not very well	Not at all well
AP ($n = 15$)	3 (20)	8 (53)	1 (7)	2 (13)	1 (7)
CT ($n = 16$)	11 (69)	4 (25)	1 (6)	—	—

get on with your therapist?” are summarized in Table 6. A Mann–Whitney U test revealed that CT patients rated the therapeutic relationship significantly more positively than did AP patients ($U = 56$, $n_1 = 15$, $n_2 = 16$, $p < .05$). It can be seen that nearly 70% of CT patients reported getting on extremely well with their therapist and none reported a negative relationship. In contrast, only 20% of AP patients reported getting on extremely well with their therapist and 20% reported a negative relationship.

Invited comments on the two ratings were categorized according to the principal theme expressed and whether the comment was positive or negative. They are summarized in Table 7 in terms of the frequency of the comments made by patients in each of the two main treatment conditions. It can be seen that comments by CT patients were predominantly positive and were mainly concerned with what had been learnt from the experience in terms of new perspectives or particular coping strategies. A few patients clearly felt that treatment had been too brief. Positive comments by AP patients were fewer than in CT and more general in character, mainly focusing on the value of being able to talk about personal problems with an interested therapist. Negative comments by AP patients were more frequent than in CT and focused on both the absence of practical help in learning how to cope and general difficulties in relating to the therapist.

There was a clear indication in the AP condition that the therapeutic relationship was perceived positively by some patients but by others as a rather anxiety-provoking experience. This finding suggested the possibility of therapist differences within AP but not CT and this was investigated in terms of the two rating scales. Mann–Whitney U tests revealed significant differences in terms of the helpfulness of the therapy between the two AP therapists ($U = 8.5$, $n_1 = 10$, $n_2 = 5$, $p < .05$) but not between the two CT therapists ($U = 25.5$, $n_1 = 7$, $n_2 = 9$, n.s.). The same test for ratings of the relationship with the therapist revealed no significant differences between the two therapists for either AP ($U = 25.5$, $n_1 = 10$, $n_2 = 5$, n.s.) or CT ($U = 16.5$, $n_1 = 7$, $n_2 = 9$, n.s.) although in both cases p was very close to significance. Inspection of the raw data suggested that one of the AP therapists was rated rather negatively and, conversely, one of the CT therapists was rated very positively, with one therapist in each of the treatments rated in broadly similar and generally positive terms.

Discussion

The most striking finding of this follow-up study is that mean scores on the two main outcome measures (BSI and STAI-T), at high contact, remained largely unchanged from pre-treatment to one year follow-up in AP but showed an obvious and sustained improvement in CT. Indeed, about twice as many patients in CT as AP, at high contact,

Table 7. Summary (and frequency) of comments by patients on the helpfulness of treatment and relationship with therapist

Helpfulness of treatment		
	Positive comments	Negative comments
CT (<i>n</i> = 16)	Learnt beneficial approaches to tackling difficult situations (6) Reassuring to talk about problems (3) New ways to look at problems (2) Learnt to stop the spiral of anxiety (2) Relaxation techniques (2) Understand the nature of anxiety (1) Learnt that was able to cope (1) Breathing exercises helped (1)	Treatment too brief (3) Treatment too focused, needed to address other issues (1)
AP (<i>n</i> = 15)	Opportunity to talk (6) Understand the cause of the problem (2) New ways of looking at the problem (1)	Needed more practical suggestions on how to cope (5) Treatment too brief (3) Difficult to know what to talk about (2) Anxiety about having to attend (1)
Relationships with therapist		
	Positive comments	Negative comments
CT (<i>n</i> = 16)	Able to talk about anything (3) Felt therapist was trying to help (1)	There were no negative comments
AP (<i>n</i> = 15)	Therapist interested in problems (2) Felt able to discuss anything (1) Sympathetic (1) Liked it that words were not put into my mouth (1)	Difficult to talk to therapist about sensitive issues (3) Lack of confidence with the therapist's thoughts as to the root of the problem (2) Difficult that the therapist was the opposite sex (2)

achieved clinically significant change in terms of self-rated improvement (64% vs 36%) and Jacobson criterion c (67–73% vs 21–33%). The differences in outcome between the three treatments at low contact were less clearcut and less marked than at post treatment and six month follow-up, although CT was still significantly better than AP in terms of STAI-T scores, self-reported panic and patient ratings of overall improvement. It should be noted, however, that overall outcome for all three treatments at low contact was generally very poor in terms of the BSI, a self-report measure of overall psychiatric symptomatology. Only a small minority of patients (11–27%) were defined as being “recovered” on this measure in terms of Jacobson criterion c. Similarly, although there was evidence of a significant reduction in medication usage following CT (66% to 38%), in contrast to an increase in medication usage following AP (65% to 82%), the overall reduction in consultation rates for the sample was modest and nonsignificant.

These results, therefore, confirm the potential value of cognitive-behavioural therapy with generalized anxiety disorder and provide no support for the use of a dynamic,

insight-oriented approach with this condition. They run counter to the more usual finding of relatively small differences in outcome between different types of treatment. Frank (1985), for example, includes in his summary of the psychotherapy outcome literature the following generalization: "... when two therapies yield differences in outcome at the close of treatment, with rare exceptions these differences disappear over time, and the closing of the gap seems to depend more on patients who receive the less successful therapy catching up than on both groups regressing equally towards the mean" (p. 55). There is certainly no evidence from the present study of a delayed improvement effect in patients receiving AP; at best, no more than one third of AP patients showed any evidence of a sustained improvement at one year follow-up. The present data, therefore, are not consistent with the general notion that "... the main beneficial effect of psychotherapy with many patients may be to accelerate improvement that would have occurred eventually in any case" (Frank, 1985, p. 55). Unless it is assumed that AP actually hindered improvement that would have occurred without therapy, and there is no evidence that this was the case, the data suggest that the provision of psychological therapy per se is no guarantee that improvement will take place with GAD patients, even over the longer-term. On the other hand, if improvement does take place it is most likely to be sustained with a cognitive therapy approach that is reasonably intensive (i.e. 16–20 sessions over six months).

Our discussion of these results must begin with an acknowledgement of some methodological limitations of the clinical trial on which this follow-up study is based. Borkovec and Newman (in press) list seven "methodological flaws" in the study that they believe "render the conclusions questionable", namely, "brief duration of GAD, therapist-by-treatment confound, lower credibility for the psychodynamic condition, no therapy manuals, large drop-out rate, higher severity for the psychodynamic condition, and no therapy integrity checks". We have already discussed the pre-treatment differences on the two main self-report measures in which the mean scores on the low contact CT condition were significantly lower than in the other groups. Borkovec and Newman (in press) rather misleadingly describe this as "higher severity for the psychodynamic condition" when in fact there is no evidence of significant difference between the severity of pre-treatment scores in the high contact condition of CT and AP as can be seen from an inspection of Table 1. The problem lies in the lower severity of the low contact cognitive therapy condition on the two main self-report measures. We have tried to manage this with covariance analysis and, in the case of Jacobson criteria for clinically significant change, by excluding cases in which pre-treatment scores were already lower than the cut-off point for change. The overall conclusion of a superior outcome for CT, especially in the high contact condition where there were clearly no pre-treatment differences, seems robust. However, it is unsatisfactory that a "randomization failure" was evident in the low contact condition and the results here should be treated with caution. Shapiro (1995) describes a similar failure in the Second Sheffield Psychotherapy Project, an occurrence attributed to research using samples in the tens rather than the hundreds.

The "therapist-by-treatment confound" might reasonably be considered as not so much a methodological flaw as an inevitable consequence of choosing a design that emphasized external rather than internal validity considerations in respect of comparing two very contrasting therapies that are typically available in most clinical settings

in the U.K. Psychotherapy research, as Aveline, Shapiro, Parry and Freeman (1995) concluded in the final chapter of a lengthy and critical review of the field, should be “representative of real practice, i.e. patients, practitioners and therapy”. Randomized controlled trials comparing clearly different therapies, delivered by committed and experienced therapists, with well-defined clinical populations, are few in number, especially in the case of the psychoanalytic therapies (Chambless & Gillis, 1993), and it was our intention to remedy this state of affairs. In practice, of course, it is very difficult to know either how representative the therapists were or, in the absence of therapy integrity checks, how competent they were. This latter omission was certainly regrettable; audiotapes were collected but we were unable to obtain resources to make the necessary ratings. We were careful to balance training and experience levels across the analytic and cognitive therapists but we cannot, of course, be sure that the very marked differences in outcome between AP and CT, especially at high contact, were due to treatment and not therapist differences.

Some therapists are undoubtedly more effective than others (Lambert, 1989) and there is an indication in the present analysis of possible therapist differences, with one of the AP therapists being perceived in significantly more negative terms than the others. There is also evidence, as Borkovec and Newman rightly point out, that there were overall significant differences between the treatments in their perceived credibility and there were also, probably as a consequence of this, quite large although nonsignificant differences in the drop-out rates. Patients in CT rated both their expectations of improvement and the suitability of their treatment more highly than patients in AP and twice as many AP as CT patients dropped out of treatment. This may reflect therapist differences in their provision of the so-called nonspecific ingredients of therapy (a warm empathic relationship, an explanation of symptoms, and so forth) or it may reflect the inherent unsuitability of psychodynamic psychotherapy for the majority of patients recruited to the study, bearing in mind that there was a significant bias in the sample towards lower socioeconomic status. The results presented here suggest that both factors are probably of importance and that, for patients suffering from GAD, a combination of a negatively perceived therapist with an inappropriate treatment may result in a particularly poor outcome. Extensive data were in fact collected on the quality of the therapeutic relationship, from both patient and therapist perspectives, using the Penn Helping Alliance Scales (Alexander & Luborsky, 1986) and a report on this is in preparation.

We would agree with Borkovec and Newman, therefore, that a different methodology in which therapists were crossed with treatments as, for example, in the Second Sheffield Psychotherapy Project (Shapiro, Barkham, Hardy, & Morrison, 1990), would have overcome the therapist-by-treatment confound, and thereby enhanced internal validity, but it is far from clear that the resulting “analytic psychotherapy” would have been representative of this type of therapy as typically practised. Our main interest was in testing the relative efficacy of cognitive-behavioural therapy and short-term analytic psychotherapy with a group of patients who were commonly referred to the psychiatric outpatient clinic in which we were practising at the time. This clinical setting, typical surely of many psychiatric services, is one in which resources do not permit lengthy treatment or careful matching of patient and therapist, and where the majority of patients are likely to be more interested in obtaining rapid symptomatic relief than in

acquiring psychological insights into the nature of their problems (cf. Durham, 1995).

The context in which outcome studies are conducted will, of course, inevitably have some influence on the results and it should be noted that the problem of “investigator allegiance” (Shapiro, 1995) was not wholly overcome. Although the project was jointly conceived by two clinicians whose respective clinical practice was in cognitive-behavioural and analytic psychotherapy, the clinical context and research methodology would generally be considered as more congenial to the cognitive-behavioural approach. Clinical trials of this kind have only rarely been attempted within a public health service setting and very few analytic psychotherapists have been bold enough to submit their clinical practice to the rigours of randomization, short-term treatment, objective ratings, and so forth. Thus, while the relatively clearcut nature of the present results constitute a strong argument for a replication to test their generalizability, the relatively unambiguous outcome viewed from a psychoanalytic perspective may well encourage greater caution concerning projects of this kind.

Finally, the superior outcome of more intensive CT may have significant implications for clinical practice and merits further research. The proportion of CT patients showing clinically significant improvements was substantially higher in the more intensive therapy condition and, in general, therapy at lower contact was less powerful in all three conditions. This finding should also be considered in the light of an analysis of outcome predictors for this same study reported by Durham, Allan and Hackett (1997). They found that outcome over the one year follow-up period could be predicted with a relatively high degree of accuracy on the basis of variables reflecting the complexity and severity of the presenting disorder. In particular, axis one comorbidity, social isolation and marital tension, were strongly associated with a poor outcome. Clinical experience and evidence from clinical trials (cf. Shapiro et al., 1994; Elkin et al., 1995) suggests that more intensive treatment is likely to be more effective, particularly for those patients with more complex and severe problems; a multi-centre study, based in Dundee, Fife and Stirling, is currently underway to investigate this possibility.

Acknowledgements

The preparation of this article was made possible by a grant to R. C. Durham (and Dr T. J. C. Murphy, Consultant Psychotherapist, Royal Edinburgh Hospital) by the Scottish Hospital Endowments Research Trust whose support is gratefully acknowledged.

References

- ALEXANDER, L. B., & LUBORSKY, L. (1986). The Penn Helping Alliance Scales. In L. S. Greenberg & W. M. Pinsof (Eds.), *The psychotherapeutic process: A research handbook*. New York: Guilford.
- AVELINE, M., SHAPIRO, D. A., PARRY, G., & FREEMAN, C. (1995). Building research foundations for psychotherapy practice. In M. Aveline & D. A. Shapiro (Eds.), *Research foundations for psychotherapy practice*. New York: John Wiley & Sons.
- BARLOW, D. H., RAPEE, R. M., & BROWN, T. A. (1992). Behavioural treatment of generalised anxiety disorder. *Behaviour Therapy*, 23, 551–570.
- BECK, A. T., & EMERY, G. (1985). *Anxiety disorder and phobias: A cognitive perspective*. New York: Basic Books.

- BORKOVEC, T. D., & COSTELLO, E. (1993). Efficacy of applied relaxation and cognitive-behavioural therapy in the treatment of generalised anxiety disorder. *Journal of Consulting and Clinical Psychology*, *61*, 611–619.
- BORKOVEC, T. D., & MATHEWS, A. (1988). Treatment of non-phobic anxiety disorders: A comparison of nondirective, cognitive and coping desensitisation therapy. *Journal of Consulting and Clinical Psychology*, *56*, 877–884.
- BORKOVEC, T. D., & NEWMAN, M. G. (in press). Worry and generalised anxiety disorder. In P. Salkovskis (Ed.), *Comprehensive clinical psychology*. Oxford: Elsevier Science.
- BORKOVEC, T. D., & WHISMAN, M. A. (1997). Psychosocial treatment of generalised anxiety disorder. In M. Mavissakalian & R. Prien (Eds.), *Anxiety disorders: Psychological and pharmacological treatments*. Washington, DC: American Psychiatric Association.
- BURNS, D. D., & HOLEN-HOEKSEMA, S. (1992). Therapeutic empathy and recovery from depression in cognitive-behavioural therapy. *Journal of Consulting and Clinical Psychology*, *60*, 441–449.
- BUTLER, G., FENNELL, M., ROBSON, P., & GELDER, M. (1991). Comparison of behaviour therapy and cognitive-behaviour therapy in the treatment of generalised anxiety disorder. *Journal of Consulting and Clinical Psychology*, *59*, 167–175.
- CHAMBLESS, D. L., & GILLIS, M. M. (1993). Cognitive therapy of anxiety disorders. *Journal of Consulting and Clinical Psychology*, *61*, 248–260.
- CHENG, L. Y., & BAXTER, H. (1995). Comparing treatments for generalised anxiety disorder. *British Journal of Psychiatry (correspondence)*, *166*, 265–266.
- CLARK, D. M. (1989). Anxiety states: Panic and generalised anxiety. In K. Hawton, P. Salkovskis, J. Kirk, & D. M. Clark (Eds.), *Cognitive behaviour therapy for psychiatric problems: A practical guide*. Oxford: Oxford Medical Publications.
- DEROGATIS, L. R., & MELISARATOS, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine*, *13*, 595–605.
- DI NARDO, P. A., & BARLOW, D. H. (1988). *Anxiety Disorders Interview Schedule-Revised (ADIS-R)*. Albany, NY: Graywind Publications.
- DURHAM, R. C. (1995). Comparing treatments for generalised anxiety disorder: Author's reply. *British Journal of Psychiatry (correspondence)*, *166*, 266–267.
- DURHAM, R. C., & ALLAN, T. (1993). Psychological treatment of generalised anxiety disorder: A review of the clinical significance of results in outcome studies since 1980. *British Journal of Psychiatry*, *163*, 19–26.
- DURHAM, R. C., ALLAN, T., & HACKETT, C. (1997). On predicting improvement and relapse in generalised anxiety disorder following psychotherapy. *British Journal of Clinical Psychology*, *36*, 101–119.
- DURHAM, R. C., MURPHY, T. J., ALLAN, T., RICHARD, K., TRELIVING, L., & FENTON, G. W. (1994). A comparison of cognitive therapy, analytic psychotherapy and anxiety management training in the treatment of generalised anxiety disorder. *British Journal of Psychiatry*, *165*, 315–323.
- DURHAM, R. C., & TURVEY, A. A. (1987). Cognitive therapy vs behaviour therapy in the treatment of chronic general anxiety. *Behaviour Research and Therapy*, *25*, 229–234.
- ELKIN, I., GIBBONS, R. D., SHEA, M. T., SOTSKY, S. M., WATKINS, J. T., PILKONIS, P. A., & HEDEKER, D. (1995). Initial severity and differential treatment outcome in the NIMH treatment of depression collaborative research program. *Journal of Consulting and Clinical Psychology*, *63*, 841–847.
- FRANK, J. D. (1985). Therapeutic components shared by all psychotherapies. In M. J. Mahoney & A. Freeman (Eds.), *Cognition and psychotherapy*. New York: Plenum Press.
- JACOBSON, N. S., & TRAU, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*, 12–19.

- LAMBERT, M. (1989). The individual therapist's contribution to psychotherapy process and outcome. *Clinical Psychology Review*, 9, 469–486.
- MEICHENBAUM, D. (1985). *Stress inoculation training*. New York: Pergamon Press.
- OFFICE OF POPULATION CENSUSES AND SURVEYS. (OPCS) (1980). *Classification of occupations*. London: HMSO.
- POWER, K. G., JERROM, D. W. A., SIMPSON, R. J., MITCHELL, M. J., & SWANSON, V. (1989). A controlled comparison of cognitive-behaviour therapy, diazepam and placebo in the management of generalised anxiety. *Behavioural Psychotherapy*, 17, 1–14.
- RICKELS, K., & SCHWEIZER, E. (1990). The clinical course and long-term management of generalised anxiety disorder. *Journal of Clinical Psychopharmacology*, 10, 101–110.
- ROTH, A., & FONAGY, P. (1996). *What works for whom? A critical review of psychotherapy research*. London: Guilford Press.
- SHAPIRO, D. A. (1995). Decisions, decisions, decisions: Determining the effects of treatment method and duration on the outcome of psychotherapy for depression. In M. Aveline & D. A. Shapiro (Eds.), *Research foundations for psychotherapy practice*. Chichester: Wiley.
- SHAPIRO, D. A., BARKHAM, M., HARDY, G. E., & MORRISON, L. A. (1990). The Second Sheffield Psychotherapy Project: Rationale, design and preliminary outcome data. *British Journal of Medical Psychology*, 63, 97–108.
- SHAPIRO, D. A., BARKHAM, M., REES, A., HARDY, G. E., REYNOLDS, S., & STARTUP, M. (1994). Effects of treatment duration and severity of depression on the effectiveness of cognitive/behavioural and psychodynamic/interpersonal psychotherapy. *Journal of Consulting and Clinical Psychology*, 62, 522–534.
- SPIELBERGER, C. D., GORSUCH, R. L., LUSHENE, R., VAGG, P. R., & JACOBS, G. A. (1983). *Manual for the State-Trait Anxiety Inventory (Form Y Self-evaluation Questionnaire)*. Palo Alto, CA: Consulting Psychologists Press.
- WHITE, J. (1998). 'Stress control' large group therapy for generalised anxiety disorder: Two year follow-up. *Behavioural and Cognitive Psychotherapy*, 26, 237–245.
- WHITE, J., KEENAN, M., & BROOKS, N. (1992). Stress control: A controlled investigation of large group therapy for generalised anxiety disorder. *Behavioural Psychotherapy*, 20, 97–113.