



Research report

# Long-term outcome of eight clinical trials of CBT for anxiety disorders: Symptom profile of sustained recovery and treatment-resistant groups

Robert C. Durham <sup>a,\*</sup>, Cassie Higgins <sup>a</sup>, Julie A. Chambers <sup>b</sup>, John S. Swan <sup>a</sup>, Michael G.T. Dow <sup>b</sup>

<sup>a</sup> Centre for Neuroscience, University of Dundee, Scotland, United Kingdom

<sup>b</sup> Department of Psychology, University of Stirling, Scotland, United Kingdom

## ARTICLE INFO

### Article history:

Received 29 June 2011

Received in revised form 15 September 2011

Accepted 15 September 2011

Available online 19 October 2011

### Keywords:

Long-term follow-up

Anxiety disorders

Chronic anxiety

Chronic depression

## ABSTRACT

**Background:** Few clinical trials of cognitive behaviour therapy (CBT) for anxiety disorders have conducted follow-up beyond one year post-treatment. This paper summarises the long-term outcome of eight clinical trials of CBT for anxiety disorders in terms of diagnostic status, healthcare usage and symptom severity and compares the symptom profile of participants with the best and worst outcomes relative to chronic depression and the normal population. **Methods:** Follow-up at 2–14 years with 396 patients (51% of those available to contact) employed structured diagnostic interview, assessment of healthcare usage and self-report measures of symptom severity. This paper concerns 336 participants who had either no disorder or at least one anxiety disorder and information on healthcare usage over the follow-up period.

**Results:** Only 38% recovered with little or no treatment over the follow-up period while 30% had a very poor outcome despite extensive treatment for anxiety over many years. The symptom profile of this 'treatment-resistant' group was comparable to 76 patients with chronic depression and significantly worse than normative data for psychiatric outpatients. Chronic anxiety disorder with co-morbid depression has a more severe symptom profile than chronic anxiety disorder alone.

**Limitations:** The follow-up sample, although broadly representative, may have a bias towards a more favourable picture of overall outcome.

**Conclusions:** The long-term outcome of anxiety disorders, irrespective of diagnosis or active treatment, is diverse but with a tendency towards chronicity. Distinctions between acute and chronic presentations of common mental disorders are more important than distinctions between chronic anxiety and chronic depression.

© 2011 Elsevier B.V. All rights reserved.

## 1. Introduction

In 2005 we reported on the long-term outcome of participants in eight clinical trials of CBT for anxiety disorders – four for panic disorder, three for generalised anxiety disorder (GAD) and one for post-traumatic stress disorder (PTSD) (Durham et al., 2005). The primary aim of this research was to investigate the long-term effectiveness and cost

effectiveness of CBT in comparison with alternative treatments. In summary, it was found that although treatment with CBT was associated with a somewhat better long-term outcome than non-CBT in terms of overall self-reported symptom severity, it was not associated with any significant differences in terms of diagnostic status or clinical global severity. Over half the participants (52%) had at least one diagnosis at long-term follow-up, irrespective of treatment received, and few participants had none or only mild symptoms (18%). Of particular significance for this report, only 36% of participants reported receiving none or very little interim treatment for anxiety over the follow-up period (medication and/or psychological therapy) with 19% receiving almost constant treatment.

\* Corresponding author at: Centre for Neuroscience, Ninewells Hospital & Medical School, Dundee, DD1 9SY, Scotland, United Kingdom. Tel.: +44 1382 632410; fax: +44 1382 633923.

E-mail address: [rob.durham49@gmail.com](mailto:rob.durham49@gmail.com) (R.C. Durham).

Clearly the good outcomes with CBT for anxiety disorders found over the short-term – 6 to 12 months following the start of therapy in the original trials – are no guarantee of good outcomes over the longer term. This conclusion is perhaps unsurprising given the evidence that both panic disorder and GAD, the two main subjects of this research, are known to follow a fluctuating and often chronic course associated with heavy demands on health service resources (Cowley et al., 1996; Rickels and Schweizer, 1990). We know also that, where anxiety and depressive symptoms coexist, which probably occurs in the majority of cases referred for psychological therapy in an NHS context, there is a heightened probability of a chronic and relapsing course (Emmanuel et al., 1998). Notwithstanding this evidence, the ambitious claim that CBT has the power to bring about a sustained improvement in anxiety symptoms over the long-term, that is, for periods in excess of 12 months following the end of treatment, is at least implicit in recent clinical guidelines (National Institute for Clinical Excellence, 2011), reviews of CBT efficacy (Butler et al., 2006) and papers on the allocation of health service resources (Layard, 2006). The empirical foundation for this claim, however, is relatively weak. In large part it is based on extrapolating from evidence that treatment gains are generally maintained in those clinical trials of psychotherapy (Lambert and Ogles, 2004) and CBT (Butler et al., 2006) that have included follow-up over 6–12 month periods. Very few clinical trials of CBT for panic disorder, GAD or PTSD have conducted follow-up investigations beyond one year post-treatment and where this has been done, problems of attrition, intervening treatment and limited outcome assessment have made it difficult to draw firm conclusions (Roth and Fonnagy, 2005; Tyrer, 2000). It is also worth noting that follow-up studies of CBT tend to be conducted by clinical researchers with a primarily cognitive-behavioural orientation. One exception is a study reported by Seivewright et al. (1998) who conducted a 5-year follow-up of a cohort of 210 psychiatric outpatients suffering from GAD, panic disorder or dysthymic disorder and randomised to medication, CBT or self-help. 60% had a broadly favourable outcome with the remainder handicapped either intermittently or continuously throughout the follow-up period. Initial diagnosis and original treatment were found to be of no predictive value.

The evidence as a whole, therefore, is insufficient to support the claim that treatment gains are generally sustained beyond one year post treatment. The balance of evidence suggests a marked divergence of outcome over the longer term following active treatment with about 50% recovering in varying degrees from the markedly disabling symptoms associated with a formal diagnosis but with a significant minority (20–30%) following a chronic disorder despite varying degrees of further treatment. This divergence which, somewhat surprisingly, is only weakly related to the impact of the original treatment, at least as found in two moderately large scale studies (Durham et al., 2005; Seivewright et al., 1998), raises important questions for further investigation. Are the factors which determine chronicity located primarily in the characteristics and circumstances of the person who has the disorder, in the nature and severity of the disorder itself, in the quality and intensity of the therapy provided, or in some interaction of all three? At what point is the new learning, which is assumed to take place as a result of therapy, replaced or undermined by old

learning and what implications does this have for the provision of more effective therapy? If we are to address these challenging but important questions and develop a better understanding of the mechanisms by which anxiety disorders become chronic and treatment-resistant in some individuals, we need to examine more closely the nature of recovery and chronicity following treatment.

In this paper we first describe the overall pattern of outcome found in the long-term follow-up study (Durham et al., 2005) in terms of four groups characterised by the presence or absence of a diagnosable anxiety disorder with or without continuing treatment for anxiety over the follow-up period. We then examine the clinical characteristics of two sizeable groups of patients with markedly contrasting outcomes – a ‘sustained recovery’ group with no diagnosable anxiety disorder who had received none or very little treatment over the follow-up period and a ‘treatment-resistant’ group whose anxiety disorder had persisted despite extensive treatment for anxiety over the follow-up period with either medication or psychological therapy. The symptom profiles of these two groups are compared to clinical and non-clinical normative data and also to a sample of chronic, treatment-resistant depressed patients who were assessed for an open trial of ‘Coping with Depression’ groups (Swan et al., 2004). In subsequent papers we will report on co-morbidity, quality of life, and healthcare usage in this cohort and also on predictors of recovery and treatment-resistance using data gathered during the original trials.

## 2. Method

### 2.1. Participants

Full details of the methodology, participant characteristics and findings of the original study can be found in a Health Technology Assessment monograph (Durham et al., 2005). In summary, all participants had taken part in a study investigating the long-term clinical and cost effectiveness of CBT for anxiety disorders, relative to alternative treatments, in eight separate clinical trials undertaken across three localities in the central belt of Scotland between 1985 and 2001. A total of 861 people with diagnosed anxiety disorders entered these trials, three of which focused on generalised anxiety disorder (N = 318), four on panic disorder (N = 454) and one on post-traumatic stress disorder (N = 89). Of these, 640 (74%) completed treatment. Follow-up assessment was undertaken 2 to 14 years following the original treatment (January 1999 to December 2003) with 396 patients (46% of the original entrants, 51% of those available to contact). Of this group, 55 had no information concerning long-term follow-up diagnostic status or the degree of interim treatment. A further five participants were diagnosed with major depressive disorder, dysthymia, hypochondriasis or substance misuse but no anxiety disorder and were excluded in order to focus on the clinical features of chronic anxiety disorder. This resulted in a cohort for the present analysis of 336 participants who had either none or at least one anxiety disorder at long-term follow-up. Co-morbidity was high among those with at least one anxiety disorder (mean number of diagnoses = 2.8) with 54% having a co-morbid diagnosis of either major depressive disorder or dysthymia.

## 2.2. Representativeness of participants in relation to population demographics and original entrants

A full analysis of the representativeness of the long-term follow-up participants with respect to the clinical and demographic characteristics of the entrants to the original clinical trials can be found in the monograph (Durham et al., 2005). In general, the follow-up sample appeared to be broadly representative of the original cohort. There was, however, a tendency for non-participants to have worse pre-treatment scores on outcome measures and for participants to have been more likely to have completed treatment. Any bias in the follow-up sample is likely to represent an over- rather than under-estimation of degree of recovery.

## 2.3. Procedure

Long-term follow-up assessment comprised an interview conducted by a research psychologist, blind to treatment condition, a number of self-report measures and data from each patient's general practice case notes on healthcare resources used in the two years prior to entering the trials and the two years prior to follow-up interviews. All procedures for tracing patients, collecting data and obtaining informed consent were approved by the respective medical research ethics committees in the three localities in which the eight clinical trials were conducted: Tayside, Forth Valley and Fife Health Boards.

## 2.4. Structured clinical interview

The Anxiety Disorders Interview Schedule IV (ADIS-IV) was used to assess diagnostic status according to DSM-IV criteria (Brown et al., 2005). Clinical Global Severity (CGS) is a component of ADIS-IV and was rated using a 0–8 scale where '0' = no evidence of disorder, '4' = definitely disturbing/disabling, and '8' = very severely disturbing/disabling. A score of 4 or above equates to having at least one DSM-IV diagnosis.

The Hamilton Anxiety Rating Scale (HAM-A) (Hamilton, 1959) was used to assess overall symptom severity with clinically significant change defined as a level of two or more standard deviations below the pre-treatment mean using Jacobson criterion (a) (Jacobson and Truax, 1991).

An assessment of the amount of patient-reported interim treatment for mental health problems since the original trial, with either medication or psychological therapy, was made using a four point scale: (1) none, (2) little – treatment over one short time period, (3) moderate – treatment over several years and (4) a lot – treatment for the majority of the intervening period. Patient reports were cross-checked with medical case notes for 36% of the sample in trial one and results tallied exactly in 67% of cases. Where discrepancies existed they were primarily due to patients excluding prescribed medication for anxiety which they may have believed had been prescribed for problems other than anxiety. Thus patient-reported treatment in the interim period is likely to be under- rather than over-reported.

## 2.5. Self-report questionnaires

The general severity index (gsi) of the Brief Symptom Inventory (BSI) (Derogatis and Melisaratos, 1983) provides a

global measure of current symptomatic state. It has a number of subscales: anxiety; depression, phobic symptoms; interpersonal sensitivity; hostility; obsessive-compulsive; somatisation; paranoid ideation; and psychoticism. Normative data are available for adult non-clinical populations and adult psychiatric out-patient populations (Derogatis, 1993). Jacobson criterion (c) (Jacobson and Truax, 1991) was used to identify clinically significant change on this measure. Other measures included the trait version of the Spielberger State-Trait Anxiety Inventory – (STAI-T) (Spielberger et al., 1970) and the trait version of the Positive and Negative Affect Scale (PANAS) (Watson et al., 1988).

## 2.6. Characteristics of chronic depression cohort

Symptom severity scores on the BSI of 76 participants in an open study of the clinical effectiveness of 'Coping with Depression' groups in the management of chronic, treatment-resistant depression were used as a comparison group (Swan et al., 2004). Participants in this study were recruited from three community mental health teams in an inner city population. All had a primary diagnosis of chronic or recurrent depressive disorder, with a current depressive episode of at least moderate severity (ICD-10 F32.1–F32.2 or F33.1–F33.2) and a history of poor response to previous treatments. Participants were aged between 18 and 65, 61% were women, 44% came from relatively deprived communities, 42% were married or cohabiting, 20% were single, 33% separated, 19% were in full-time employment and 59% were on sickness benefit.

## 3. Results

### 3.1. Pattern of overall outcome for diagnostic groups and total sample

The overall pattern of outcome at long-term follow-up is presented first in terms of four distinct groups defined by diagnostic status (recovered i.e. no clinical diagnosis at long-term follow-up versus disordered i.e. at least one anxiety disorder diagnosis) and, within these two categories, by the degree of interim treatment during the follow-up period (i. e. little or none versus moderate or a lot) – see Table 1. Approximately equal proportions of the overall cohort can be described as recovered (51%) or still disordered (49%) in terms of diagnostic status. A 'sustained recovery' group ( $n = 128$ , 38%) present with no clinical diagnoses and report little or no interim treatment over the follow-up period. A much smaller group ( $n = 43$ , 13%) also present with no clinical diagnoses but report a moderate amount or a lot of interim treatment. Within the disordered category around a fifth of patients ( $n = 65$ , 19%) present with a diagnosis of at least one anxiety disorder but report receiving little or no interim treatment. Finally, a larger group, labelled 'treatment-resistant' ( $n = 100$ , 30%), present with a clinical diagnosis of at least one anxiety disorder despite receiving either a moderate amount or a lot of interim treatment (group D). There are no significant differences in the overall proportions across the three diagnostic categories ( $\chi^2(6) = 10.3$ ;  $p = .113$ ).

**Table 1**

Outcome at long-term follow-up by original diagnosis and degree of treatment over the follow-up period.

	Recovered <sup>a</sup>		Disordered <sup>b</sup>		Total
	Without interim Rx <sup>c</sup> "Sustained recovery"	With interim Rx <sup>d</sup>	Without interim Rx <sup>c</sup>	With interim Rx <sup>d</sup> "Treatment resistant"	
GAD <sup>e</sup>	46 (40%)	14 (12%)	23 (20%)	33 (28%)	116
PD <sup>e</sup>	76 (40%)	27 (14%)	35 (19%)	51 (27%)	189
PTSD <sup>e</sup>	6 (19%)	2 (6%)	7 (23%)	16 (52%)	31
Total	128 (38%)	43 (13%)	65 (19%)	100 (30%)	336

<sup>a</sup> No diagnoses at long-term follow-up.<sup>b</sup> At least one anxiety disorder diagnosis at long-term follow-up.<sup>c</sup> None or little additional treatment for anxiety over the follow-up period.<sup>d</sup> Moderate amount or a lot of additional treatment for anxiety over the follow-up period.<sup>e</sup> GAD = generalised anxiety disorder, PD = panic disorder, PTSD = post-traumatic stress disorder.

### 3.2. Measures of clinical status across four outcome groups

Table 2 provides a summary of clinical status across the four outcome groups in terms of symptom severity scores for both assessor-rated and self-report measures and indices of clinically significant change for the primary outcome measure in each of these categories. Significant differences on symptom severity measures were assessed using ANOVA with individual interactions between the groups assessed using the Scheffe *post-hoc* test. There is an ordered, significant and consistent pattern of increasing severity across the

four groups for all measures. This pattern is also evident in the proportions of participants achieving clinically significant improvement using Jacobson methodology for the HAM-A and the BSI-gsi. For both of these measures a substantial majority (61–78%) of participants in the two recovered groups achieve criteria for clinically significant improvement whereas only a very small proportion (1–12%) do so in the two disordered groups. Within both the disordered and recovered groups, being in receipt of a moderate or a lot of interim treatment is associated with a greater degree of clinical severity on all measures. None of these differences are significant

**Table 2**

Summary of clinical status across outcome groups irrespective of original diagnosis.

	Recovered <sup>a</sup>		Disordered <sup>b</sup>		Significance and interactions using the Scheffe <i>post-hoc</i> test
	A Without Interim Rx <sup>c</sup> (N = 128) 'Sustained recovery'	B With Interim Rx <sup>d</sup> (N = 43)	C Without Interim Rx <sup>c</sup> (N = 65)	D With Interim Rx <sup>d</sup> (N = 100) 'Treatment-resistant'	
<i>Assessor-rated measures</i>					
Hamilton Anxiety Index (HAM-A)					
Mean (SD)	6.8 (3.8)	8.3 (3.1)	16.6 (3.8)	20.0 (4.8)	F(3, 277) = 197; p < .001
Clinically significant improvement (%)	76	61	2	1	A < D; B < C, C < D;
Clinical Global Severity (CGS)					
Mean (SD)	1.6 (1.0)	1.9 (0.8)	4.5 (0.6)	5.1 (0.9)	F(3, 333) = 365; p < .001
					A < D; B < C < D
<i>Self-rated measures</i>					
Brief Symptom Inventory (BSI-gsi)					
Mean (SD)	0.6 (0.5)	0.7 (0.4)	1.5 (0.7)	2.0 (0.7)	F(3, 330) = 114; p < .001
Clinically significant improvement (%)	78	77	12	7	A < D; B < C < D
Spielberger STAI-T <sup>e</sup>					
Mean (SD)	37.6 (9.2)	42.3 (9.3)	55.1 (2.4)	58.8 (10.2)	F(3, 296) = 102; p < .001
					A < D; B < C < D
PANAS (positive) <sup>e</sup>					
Mean (SD)	33.7 (7.2)	31.4 (6.6)	27.0 (7.5)	24.2 (8.1)	F(3, 319) = 32; p < .001
					A > D; B > C
PANAS (negative) <sup>e</sup>					
Mean (SD)	17.5 (6.1)	18.1 (5.1)	30.0 (7.7)	32.4 (7.7)	F(3, 319) = 111; p < .001
					A < D; B < C

<sup>a</sup> No diagnoses at long-term follow-up.<sup>b</sup> At least one anxiety disorder diagnosis at long-term follow-up.<sup>c</sup> None or little additional treatment for anxiety over the follow-up period.<sup>d</sup> Moderate amount or a lot of additional treatment for anxiety over the follow-up period.<sup>e</sup> STAI-T = Spielberger State-Trait Anxiety Inventory (Trait version), PANAS = Positive and Negative Affect Scale.

cant between the two recovered groups (A and B) but the differences are highly significant for four of the six measures within the disordered groups (C and D). For each of the measures the difference in scores between the recovered and disordered groups, that is, between having no diagnosis and having at least one, is much greater than the differences within these two categories. In summary, both diagnostic status and degree of treatment over the follow-up period contribute to a coherent pattern of outcome in which there are marked differences in overall symptom severity between recovered and disordered groups but also consistent differences in severity within these categories as a function of treatment over the follow-up period. At one extreme there is a group of patients who achieve a sustained recovery largely without professional help and, at the other extreme, a group who are treatment-resistant with high symptom severity despite significant professional help over an extended period of time. The remaining analyses focus on these two extreme groups.

### 3.3. Clinical global severity of 'sustained recovery' and 'treatment-resistant' groups

Fig. 1 illustrates the range of scores on the CGS for the two extreme groups. The scores represent a global judgement of the independent assessor based on symptom severity and the degree to which symptoms are disturbing and disabling. The threshold for diagnosis is a score of 4, defined as moderately severe symptoms which are definitely disturbing and/or disabling. Although 80% of the recovered group report a CGS score of 2 or less ('mild' symptoms that are slightly disturbing or disabling), only a small minority of this group (16%) are entirely free of symptoms (CGS = 0) and 20% experience symptoms just below threshold for diagnosis (CGS = 3). There is a similar range of severity in the treatment-resistant group with the majority scoring 4 or 5 ('moderately severe' symptoms) but a significant minority (36%) scoring 6 or 7 ('severe' symptoms).

### 3.4. Symptom profiles of 'sustained recovery', 'treatment-resistant' and chronic depression groups relative to non-clinical norms

Fig. 2 shows T-scores on the nine BSI subscales and global severity index for the 'sustained recovery', 'treatment-resistant' and chronic depression groups in relation to non-clinical adult normative data. Scores on all subscales for both the 'treatment-

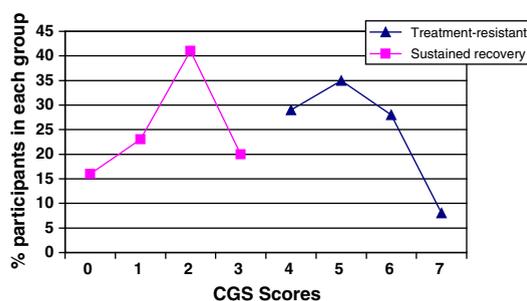


Fig. 1. Clinical Global Severity (CGS) at long-term follow-up.

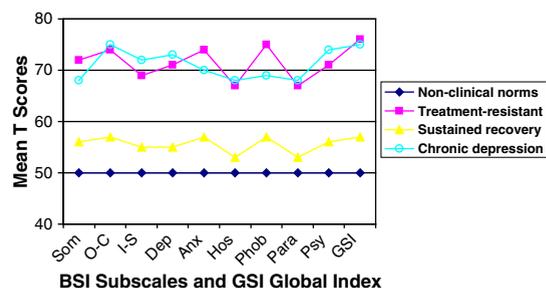


Fig. 2. T scores on BSI subscales and the GSI global index compared to non-clinical adult norms.

resistant' and chronic depression groups are significantly different (all at  $p < .001$  using univariate analysis of variance) from the non-clinical adult normative data (Derogatis, 1993), falling 1.5 to 2.5 standard deviations above the adult norms. The 'sustained recovery' group also has significantly higher subscale scores (by 1 to 1.5 standard deviations) than the adult norms (all at  $p < .001$  using univariate analysis of variance). The overall profiles of the 'treatment-resistant' and chronic depression groups are broadly similar, with virtually identical global severity indices, but there are some significant differences between them on particular subscales. The 'treatment-resistant' anxiety disorder group scores significantly higher than the chronic depression group on the somatisation subscale ( $F(1, 174) = 5.674$ ;  $p = .018$ ), the anxiety subscale ( $F(1, 174) = 9.327$ ;  $p = .003$ ) and the phobic anxiety subscale ( $F(1, 173) = 17.643$ ;  $p < .001$ ), whereas the chronic depression group scores significantly higher on the interpersonal sensitivity subscale ( $F(1, 173) = 6.486$ ;  $p = .012$ ), the depression subscale ( $F(1, 173) = 4.681$ ;  $p = .032$ ) and the psychoticism subscale ( $F(1, 173) = 6.074$ ;  $p = .015$ ).

It should be noted that a proportion of participants at long-term follow-up who had at least one anxiety disorder diagnosis (i.e. the 'disordered' group) also met diagnostic criteria for either major depressive disorder or dysthymia. In the case of the 'treatment-resistant' group 63% had co-morbid depression or dysthymia. A test of the difference in BSI subscale scores between the 'treatment-resistant' group with and without co-morbid depression and dysthymia found that on the general severity index and on all subscales except phobic anxiety, the presence of co-morbid depression or dysthymia is associated with significantly higher symptom severity ( $F(1, 96) = 13.67$ ;  $p < .001$  on the general severity index). The overall symptom profile of these two subgroups of the 'treatment-resistant' group remains very similar to the chronic depression group in pattern and overall severity but it is clear that the presence of an affective disorder in chronic anxiety is associated with elevated symptom severity.

The symptom profile of the 'sustained recovery' group appears to be only marginally above that of the non-clinical norms, but there is in fact a significant difference on all subscales and on the global severity index between the 'sustained recovery' group and the non-clinical norms ( $p < .005$  on the paranoia subscale and  $p < .001$  on all other measures). However, all the scores of the 'sustained recovery' group fall below the cut-off T score for caseness ( $T = 63$ ) and are within the range of scores expected for non-clinical cases.

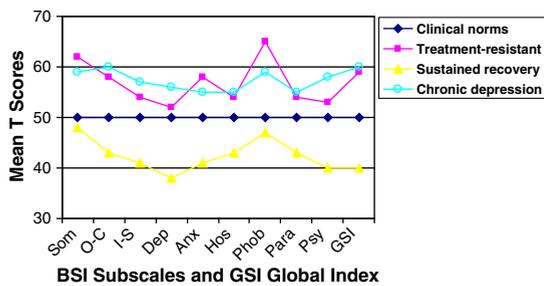


Fig. 3. T scores on BSI subscales and the GSI global index compared to adult psychiatric out-patient norms.

### 3.5. Symptom profiles of 'sustained recovery', 'treatment-resistant' and chronic depression groups relative to adult psychiatry outpatient norms

Fig. 3 repeats the above analysis using adult psychiatric out-patient normative data (Derogatis, 1993) as the comparator. Scores for the 'treatment-resistant' anxiety disorder group and the chronic depression group are, again, broadly comparable in overall severity and significantly worse than norms for psychiatric outpatients on all BSI subscales and the general severity index (all significant at  $p < .001$  except the depression subscale of the 'treatment-resistant' anxiety group at  $p < .005$ ). Scores for both groups on the global severity index is about one standard deviation above the normative data. The 'sustained recovery' group scores significantly lower ( $p < .001$ ) than psychiatric outpatients on all but two of the subscales. On the somatic and phobic anxiety subscales their scores are comparable to psychiatric outpatients.

## 4. Discussion

In this report long-term recovery from emotional disorder is analysed using a combination of diagnostic status, degree of interim treatment over the follow-up period and self-report measures of clinical severity. Three main points emerge. First, the results provide further confirmation that the long-term outcome of anxiety disorders, irrespective of diagnostic type and active treatment, is very diverse with a tendency towards chronicity (Andersch and Hetta, 2003; Roth and Fonnagy, 2005; Yonkers et al., 2003). Only about 40% of patients achieve a sustained recovery in the sense of no longer meeting diagnostic criteria for an anxiety disorder and having received little or no treatment for anxiety over the follow-up period. This group has the lowest overall levels of symptom burden and fall within the range of scores expected for non-clinical cases. Nonetheless, their overall symptom severity is significantly worse than non-clinical norms on all subscales of the BSI and only a small minority of this group is entirely free of symptoms. Most continue to experience mild symptoms that are slightly disabling and 20% of the group is just below threshold for diagnosis. Full recovery, therefore, is very unusual and a level of continuing vulnerability is common even among those who do best. The nature of this vulnerability is suggested by the finding that on the somatic and phobic anxiety subscales of the BSI their scores are comparable to psychiatric outpatient norms.

Second, within both the recovered and disordered groups, a higher level of symptom severity is associated with receiving additional treatment over the follow-up period. Within the recovered group it might be argued that the relatively small percentage of patients (13%) who have continued to receive treatment might still be disordered were it not for this additional treatment. This argument is less convincing with the 'treatment-resistant' group since their overall level of severity could hardly be much worse. In their case the additional treatment appears to have been at best ineffective and it may well be that the complexity and severity of this group require a different or more intensive treatment regime. The similarity in overall outcome between panic disorder and GAD is worth noting in this connection. The differences in clinical features between these two diagnostic groups may be of importance in the treatment of acute presentations but, as found by Seivewright et al. (1998), there is no indication that their long-term outcome is distinct. The present findings do, however, support the validity of defining the diagnostic threshold, irrespective of diagnostic type, in terms of moderately severe symptoms that are definitely disturbing or distressing. Thus, although there is a step-like increase in symptom severity over the four outcome groups, the largest and most significant step occurs across the threshold for diagnosis, that is, between groups B and C. This is a consistent pattern for both self-report and assessor-rated measures and is strongly confirmed in the proportions meeting criteria for clinically significant change on the two main outcome measures.

Third, a sizeable minority of the cohort (30%), the 'treatment-resistant' group, have a very poor outcome indeed despite extensive treatment over many years. Their overall level of symptom severity is significantly worse than psychiatric outpatients, and they have a global severity index on the BSI about one standard deviation above this group. High scores on the phobic anxiety and somatisation subscales suggest that avoidant coping (Friedman and Silver, 2006) and anxiety sensitivity (Taylor, 1995) may play a significant role in the psychopathology of chronicity in anxiety disorders. Their overall symptom profile is broadly similar to, and of comparable severity, to the chronic depression cohort and both show similar levels of severity on the obsessive-compulsive, hostility and paranoid subscales. The chronic depression group has higher levels of severity on the depression, interpersonal sensitivity and psychoticism subscales but the significant differences in symptom profile between chronic depression and anxiety are less striking than their overall similarity. However, the presence of co-morbid depression or dysthymia among the 'treatment-resistant' anxiety disorder group is associated with significantly higher levels of symptom severity. It is the combination of chronic anxiety and chronic depression that is most disabling.

### 4.1. Methodological strengths and limitations

How reliable are these findings? They are based on a sample of less than half of the original cohort of entrants to the eight clinical trials and although broadly representative of the original entrants in respect of demography, treatment conditions and short-term treatment outcome there was a consistent finding that those who completed treatment were more likely to participate in long-term follow-up. Non-participants

tended to have worse pre-treatment scores on outcome measures and it is likely that the follow-up sample may have a bias towards a more favourable picture of overall outcome. It should also be borne in mind that the assessment of interim treatment over the follow-up period is based on self-report and, although consistent with healthcare usage data from a small group of participants, it would have been strengthened by a more detailed investigation. Set against these uncertainties the sample is relatively large ( $n = 336$ ) and is drawn from eight separate clinical trials in which all entrants met diagnostic criteria for an anxiety disorder and received a comprehensive assessment of demographic status and clinical outcome using well established measures. All these trials were conducted as part of routine clinical services within a nationalised health service context in which treatment was provided by experienced clinicians and therapists. There was no attempt to limit participation to people with pure diagnoses; co-morbidity was common and the sample as a whole, at least in respect of GAD and panic disorder, can be considered as broadly representative of people suffering from common anxiety disorders and referred to psychological therapy services. The PTSD sample is small and comes from a single trial and some caution needs to be exercised in interpreting the apparently poorer outcome of this group.

#### 4.2. Implications for service delivery

Manualised and relatively brief treatment of specific anxiety disorders may well turn out to produce lasting and effective treatment for a proportion of people with acute disorders but the present findings suggest that there will be a substantial proportion of patients, at least 30%, whose long-term course is chronic and unresponsive to current treatment. Evidence of a short and medium term positive response to therapy, as found in all the clinical trials included in this cohort, is no guarantee of long-term gains. It may reflect in large measure the powerful therapeutic benefits of support, understanding and hope that are present to some degree in most therapies. On the other hand, the deep seated learning required to bring about fundamental changes in coping strategies and interpersonal style in emotionally vulnerable people may require more intensive therapy of a rather different kind than is envisaged in current mental health policy. It will also require a greater recognition of the importance of treating anxiety disorders as well as depression, in both acute and chronic presentations, and of the similar levels of disability and distress in these two groups. Of crucial importance will be early identification of vulnerability to chronicity.

#### Role of funding source

The preparation of this article was made possible by a grant from the NHS Executive Health Technology Assessment Programme (NHS R&D HTA Programme Grant 96/39/18) whose support is gratefully acknowledged. The views and opinions expressed in this article do not necessarily reflect those of the NHS Executive. The HTA Programme had no further role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

#### Conflict of interest

All the authors declare that they have no conflicts of interest.

#### Acknowledgements

The authors are grateful to Jen Petrie for her invaluable assistance in coordinating all phases of this project and to Kevin Power, Donald Sharp, Randal Macdonald, Kirsten Major and Andrew Gumley for their contribution as members of the Long-term Follow-up Research Team.

#### References

- Andersch, S., Hetta, J., 2003. A 15-year follow-up study of patients with panic disorder. *European Psychiatry* 18, 401–408.
- Brown, T.A., DiNardo, P., Barlow, D.H., 2005. *Anxiety Disorder Interview Schedule ADIS-IV*. Oxford University Press, New York City, NY.
- Butler, A.C., Chapman, J.E., Forman, E.M., Beck, A.T., 2006. The empirical status of cognitive-behavioural therapy: a review of meta-analyses. *Clinical Psychology Review* 26, 17–31.
- Cowley, D., Flick, S., Roy-Byrne, P., 1996. Long-term course and outcome in panic disorder: a naturalistic follow-up study. *Anxiety* 2, 13–21.
- Derogatis, L.R., 1993. *Brief Symptom Inventory (BSI) Administration, Scoring and Procedures Manual*, 3rd ed. National Computer Systems, Minneapolis, MN.
- Derogatis, L.R., Melisaratos, N., 1983. The brief symptom inventory: an introductory report. *Psychological Medicine* 13, 595–605.
- Durham, R.C., Chambers, J.A., Power, K.G., Sharp, D.M., Macdonald, R.R., Major, K.A., Dow, M.G.T., Gumley, A., 2005. Long-term outcome of cognitive behaviour therapy clinical trials in central Scotland. *Health Technology Assessment* 9, 42. ([www.hta.ac.uk/1081](http://www.hta.ac.uk/1081)) Accession date: 3/2/11.
- Emmanuel, J., Simmonds, S., Tyrer, P., 1998. Systematic review of the outcome of anxiety and depressive disorders. *The British Journal of Psychiatry* 173 (Suppl. 34), 35–41.
- Friedman, H.S., Silver, R.C., 2006. *Foundations of Health Psychology*. Oxford University Press, USA, p. 124.
- Hamilton, M.C., 1959. The assessment of anxiety states by rating. *The British Journal of Medical Psychology* 32, 50–55.
- 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.
- Lambert, M.J., Ogles, B., 2004. The efficacy and effectiveness of psychotherapy. In: Bergin, A.E., Garfield, S. (Eds.), *Handbook of Psychotherapy and Behaviour Change*, 5th ed. John Wiley, New York, pp. 139–193.
- Layard, R., 2006. The case for psychological treatment centres. *British Medical Journal* 332, 1030–1032.
- National Institute for Clinical Excellence, 2011. Clinical guideline 113 – generalised anxiety disorder and panic disorder (with or without agoraphobia) in adults. National Institute for Clinical Excellence, London. ([www.nice.org.uk/guidance/CG113/FullGuidance](http://www.nice.org.uk/guidance/CG113/FullGuidance)) Accession date: 3/2/11.
- Rickels, K., Schweizer, E., 1990. The clinical course and long-term management of generalized anxiety disorder. *Journal of Clinical Psychopharmacology* 10 (Suppl. 3), 101–110.
- Roth, A., Fonnagy, P., 2005. *What Works for Whom? A Critical Review of Psychotherapy Research*, 2nd edn. Guilford Press, New York.
- Seivewright, H., Tyrer, P., Johnson, T., 1998. Prediction of outcome in neurotic disorder: a 5-year prospective study. *Psychological Medicine* 28, 1149–1157.
- Spielberger, C.D., Gorsuch, R.L., Lushene, R., Vagg, P.R., Jacobs, G.A., 1970. *State-Trait Anxiety Inventory*. Mind Garden Inc, Menlo Park, CA.
- Swan, J.S., Sorrell, E., MacVicar, R., Durham, R.C., Matthews, K., 2004. 'Coping with depression': an open study of the efficacy of a group psychoeducational intervention in chronic, treatment-refractory depression. *Journal of Affective Disorders* 82, 125–129.
- Taylor, S., 1995. Anxiety sensitivity: theoretical perspectives and recent findings. *Behaviour Research and Therapy* 33, 243–258.
- Tyrer, P., 2000. *Anxiety: A Multidisciplinary Review*. Imperial College Press, London.
- Watson, D., Clark, L.A., Tellegen, A., 1988. Development and validation of brief measures of positive and negative affect: The PANAS scale. *Journal of Personality and Social Psychology* 54, 1063–1070.
- Yonkers, K.A., Bruce, S.E., Dyck, I.R., Keller, M.B., 2003. Chronicity, relapse, and illness—course of panic disorder, social phobia, and generalized anxiety disorder: findings in men and women from 8 years of follow-up. *Depression and Anxiety (Special Issue: Anxiety Disorders in Women)* 17 (3), 173–179.