Health Technology Assessment 2005; Vol. 9: No. 42

Long-term outcome of cognitive behaviour therapy clinical trials in central Scotland

RC Durham, JA Chambers, KG Power, DM Sharp, RR Macdonald, KA Major, MGT Dow and AI Gumley

November 2005

Health Technology Assessment
NHS R&D HTA Programme

How to obtain copies of this and other HTA Programme reports.
An electronic version of this publication, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (http://www.hta.ac.uk). A fully searchable CD-ROM is also available (see below).

Printed copies of HTA monographs cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our Despatch Agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per monograph and for the rest of the world £3 per monograph.

You can order HTA monographs from our Despatch Agents:

– fax (with **credit card** or **official purchase order**)
– post (with **credit card** or **official purchase order** or **cheque**)
– phone during office hours (**credit card** only).

Additionally the HTA website allows you **either** to pay securely by credit card **or** to print out your order and then post or fax it.

Contact details are as follows:

HTA Despatch
c/o Direct Mail Works Ltd
4 Oakwood Business Centre
Downley, HAVANT PO9 2NP, UK

Email: orders@hta.ac.uk
Tel: 02392 492 000
Fax: 02392 478 555
Fax from outside the UK: +44 2392 478 555

NHS libraries can subscribe free of charge. Public libraries can subscribe at a very reduced cost of £100 for each volume (normally comprising 30–40 titles). The commercial subscription rate is £300 per volume. Please see our website for details. Subscriptions can only be purchased for the current or forthcoming volume.

Payment methods

Paying by cheque
If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *Direct Mail Works Ltd* and drawn on a bank with a UK address.

Paying by credit card
The following cards are accepted by phone, fax, post or via the website ordering pages: Delta, Eurocard, Mastercard, Solo, Switch and Visa. We advise against sending credit card details in a plain email.

Paying by official purchase order
You can post or fax these, but they must be from public bodies (i.e. NHS or universities) within the UK. We cannot at present accept purchase orders from commercial companies or from outside the UK.

How do I get a copy of *HTA on CD*?

Please use the form on the HTA website (www.hta.ac.uk/htacd.htm). Or contact Direct Mail Works (see contact details above) by email, post, fax or phone. *HTA on CD* is currently free of charge worldwide.

The website also provides information about the HTA Programme and lists the membership of the various committees.

Long-term outcome of cognitive behaviour therapy clinical trials in central Scotland

RC Durham,[1*] JA Chambers,[1,2] KG Power,[2,3] DM Sharp,[4] RR Macdonald,[2] KA Major,[5] MGT Dow[2,6] and AI Gumley[7]

[1] Psychiatry and Behavioural Sciences, Division of Pathology and Neuroscience, University of Dundee, UK
[2] Department of Psychology, University of Stirling, UK
[3] Department of Clinical Psychology, NHS Tayside, UK
[4] Institute of Rehabilitation, University of Hull Medical School, UK
[5] Health Economics, NHS Ayrshire and Arran, UK
[6] Department of Clinical Psychology, NHS Fife, UK
[7] Department of Psychological Medicine, University of Glasgow, UK

* Corresponding author

Declared competing interests of authors: none

Published November 2005

This report should be referenced as follows:

Durham RC, Chambers JA, Power KG, Sharp DM, Macdonald RR, Major KA, *et al*. Long-term outcome of cognitive behaviour therapy clinical trials in central Scotland. *Health Technol Assess* 2005;**9**(42).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE and *Science Citation Index Expanded (SciSearch®)* and *Current Contents®*/Clinical Medicine.

NHS R&D HTA Programme

The research findings from the NHS R&D Health Technology Assessment (HTA) Programme directly influence key decision-making bodies such as the National Institute for Health and Clinical Excellence (NICE) and the National Screening Committee (NSC) who rely on HTA outputs to help raise standards of care. HTA findings also help to improve the quality of the service in the NHS indirectly in that they form a key component of the 'National Knowledge Service' that is being developed to improve the evidence of clinical practice throughout the NHS.

The HTA Programme was set up in 1993. Its role is to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined to include all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care, rather than settings of care.

The HTA Programme commissions research only on topics where it has identified key gaps in the evidence needed by the NHS. Suggestions for topics are actively sought from people working in the NHS, the public, service-users groups and professional bodies such as Royal Colleges and NHS Trusts.

Research suggestions are carefully considered by panels of independent experts (including service users) whose advice results in a ranked list of recommended research priorities. The HTA Programme then commissions the research team best suited to undertake the work, in the manner most appropriate to find the relevant answers. Some projects may take only months, others need several years to answer the research questions adequately. They may involve synthesising existing evidence or conducting a trial to produce new evidence where none currently exists.

Additionally, through its Technology Assessment Report (TAR) call-off contract, the HTA Programme is able to commission bespoke reports, principally for NICE, but also for other policy customers, such as a National Clinical Director. TARs bring together evidence on key aspects of the use of specific technologies and usually have to be completed within a short time period.

Criteria for inclusion in the HTA monograph series

Reports are published in the HTA monograph series if (1) they have resulted from work commissioned for the HTA Programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

The research reported in this monograph was commissioned by the HTA Programme as project number 96/39/18. The contractual start date was in January 1999. The draft report began editorial review in April 2004 and was accepted for publication in May 2005. As the funder, by devising a commissioning brief, the HTA Programme specified the research question and study design. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the referees for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

The views expressed in this publication are those of the authors and not necessarily those of the HTA Programme or the Department of Health.

ISSN 1366-5278

Published by Gray Publishing, Tunbridge Wells, Kent, on behalf of NCCHTA.
Printed on acid-free paper in the UK by St Edmundsbury Press Ltd, Bury St Edmunds, Suffolk.

Abstract

Long-term outcome of cognitive behaviour therapy clinical trials in central Scotland

RC Durham,[1*] JA Chambers,[1,2] KG Power,[2,3] DM Sharp,[4] RR Macdonald,[2] KA Major,[5] MGT Dow[2,6] and AI Gumley[7]

[1] Psychiatry and Behavioural Sciences, Division of Pathology and Neuroscience, University of Dundee, UK
[2] Department of Psychology, University of Stirling, UK
[3] Department of Clinical Psychology, NHS Tayside, UK
[4] Institute of Rehabilitation, University of Hull Medical School, UK
[5] Health Economics, NHS Ayrshire and Arran, UK
[6] Department of Clinical Psychology, NHS Fife, UK
[7] Department of Psychological Medicine, University of Glasgow, UK
* Corresponding author

Objectives: To establish the long-term outcome of participants in clinical trials of cognitive behaviour therapy (CBT) for anxiety disorders and psychosis, examining the effectiveness and cost-effectiveness associated with receiving CBT in comparison with alternative treatments.

Design: An attempt was made to contact and interview all of the participants in eight randomised, controlled, clinical trials of CBT for anxiety disorders and two randomised, controlled, clinical trials of CBT for schizophrenia conducted between 1985 and 2001. Case note reviews of healthcare resources used in the 2 years prior to entering the trials and the 2 years prior to follow-up interview were undertaken.

Setting: Mixed rural and urban settings in five localities in central Scotland. Anxiety disorder trials were conducted mainly in primary care and included three with generalised anxiety disorder, four with panic disorder and one with post-traumatic stress disorder (PTSD). The psychosis studies (one on relapse prevention and one with chronic disorder) were conducted in secondary care.

Participants: Of the 1071 entrants to the 10 studies, 489 agreed to participate (46% of original entrants, 52% of those available to contact).

Interventions: Follow-up interviews took place between 1999 and 2003, 2–14 years after the original treatment. Interviews for Trials 1–8 were conducted by a research psychologist blind to original treatment condition. Interviews for Trials 9 and 10 were conducted by community psychiatric nurses also blind to treatment condition. Case note reviews were completed following the interview.

Main outcome measures: For Trials 1–8 the main interview-based outcome measures were: Anxiety Disorders Interview Schedule – DSM-IV for diagnosis and co-morbidity, Clinical Global Severity (0–8) and the Hamilton Anxiety Rating Scale. The main patient-rated measures were: Brief Symptom Inventory, SF-36 II, Clinical Global Improvement (1–7), and the Positive and Negative Affect Scale. For Trials 9 and 10 the primary outcome measure was the interview-based Positive and Negative Syndrome Scale (PANSS).

Results: For the anxiety disorder studies (trials 1–8), over half of the participants (52%) had at least one diagnosis at long-term follow-up, with significant levels of co-morbidity and health status scores comparable to the lowest 10% of the general population. Only 36% reported receiving no interim treatment for anxiety over the follow-up period with 19% receiving almost constant treatment. Patients with PTSD did particularly poorly. There was a 40% real increase in healthcare costs over the two time periods, mainly due to an increase in prescribing. A close relationship was found between poor mental and physical health for those with a chronic anxiety disorder. Treatment with CBT was associated with a better long-term outcome than non-CBT in terms of overall symptom severity but not with regard to diagnostic status. The positive effects of CBT found in the original trials were eroded over longer time periods. No evidence was found for an association between more intensive therapy and more enduring effects of CBT. Long-term outcome was

found to be most strongly predicted by the complexity and severity of presenting problems at the time of referral, by completion of treatment irrespective of modality and by the amount of interim treatment during the follow-up period. The quality of the therapeutic alliance, measured in two of the studies, was not related to long-term outcome but was related to short-term outcome. The cost-effectiveness analysis showed no advantages of CBT over non-CBT. The cost of providing CBT in the original trials was only a very small proportion (6.4%) of the overall costs of healthcare for this population, which are high for both physical and mental health problems. In the psychosis studies (Trials 9 and 10), outcome was generally poor with only 10% achieving a 25% reduction in total PANSS scores from pretreatment to long-term follow-up, also cost-effectiveness analysis showed no advantages of CBT over non-CBT, although healthcare costs fell over the two time periods mainly owing to a reduction in inpatient costs.

Conclusions: Psychological therapy services need to recognise that anxiety disorders tend to follow a chronic course and that good outcomes with CBT over the short term are no guarantee of good outcomes over the longer term. Clinicians who go beyond standard treatment protocols of about 10 sessions over a 6-month period are unlikely to bring about greater improvement. Poor outcomes over the long term are related to greater complexity and severity of presenting problems at the time of referral, failure to complete treatment irrespective of modality and the amount of interim treatment during the follow-up period. The relative gains of CBT are greater in anxiety disorders than in psychosis. Longitudinal research designs over extended periods of time (2–5 years), with large numbers of participants (500+), are required to investigate the relative importance of patient characteristics, therapeutic alliance and therapist expertise in determining the cost-effectiveness of CBT in the longer term.

Contents

List of abbreviations

ADIS	Anxiety Disorders Interview Schedule
ADIS-IV	Anxiety Disorders Interview Schedule: DSM-IV version
AMT	anxiety management training
AP	analytical psychotherapy
BDI	Beck Depression Inventory
BF	befriending
BNF	British National Formulary
BPRS	Brief Psychiatric Rating Scale
BSI	Brief Symptom Inventory
BSI-GSI	Brief Symptom Inventory – General Severity Index
CAIR	an index of quality of therapeutic alliance
CANX	Composite Anxiety Scale
CAPS	Clinician-Administered PTSD Scale
CASP	an index of complexity and severity of presenting problems
CBT	cognitive behaviour therapy
CBT12	cognitive behaviour therapy over 12 1-hour sessions
CBT6	cognitive behaviour therapy over six 1-hour sessions
CBT6-CA	cognitive behaviour therapy over six 1-hour sessions plus computer-assisted instruction
CDEP	Composite Depression Scale
CEAC	cost-effectiveness acceptability curve
CFAM	composite measure of adjustment in the family domain
CGI	clinical global improvement
CGS	clinical global severity
CI	confidence interval
CMHT	Community Mental Health Team
COMCGS	composite measure of clinical global severity
CPN	community psychiatric nurse
CPRS	Comprehensive Psychiatric Rating Scale
CSOC	composite measure of adjustment in the social domain
CT	cognitive therapy
DML	direct maximum likelihood
DZ	diazepam
E+CR	exposure plus cognitive restructuring
EM	Expectation Maximisation
EMDR	eye movement desensitisation and reprocessing
FL	fluvoxamine
FQ	Fear Questionnaire
FQ-Agora	Fear Questionnaire agoraphobia subscale
GAD	generalised anxiety disorder
HADS	Hospital Anxiety and Depression Scale
HAM-A	Hamilton Anxiety Rating Scale
HAM-AD	Hamilton Anxiety Rating Scale (adjusted version)
ICER	incremental cost-effectiveness ratio
IES	Impact of Event Scale
LTOF	long-term outcome factor
MADRS	Montgomery Asberg Depression Rating Scale
NA	PANAS negative affect
PA	PANAS positive affect

continued

List of abbreviations *continued*

PANAS	Positive and Negative Affect Scale
PANSS	Positive and Negative Syndrome Scale
PAS	Psychiatric Assessment Scale
PBIQ	Personal Beliefs about Illness Questionnaire
PL	placebo
PSDI	Positive Symptom Distress Index
PTSD	post-traumatic stress disorder
PTSD-SCL	PTSD Symptom Checklist
RCT	randomised controlled trial
RT	randomised trial
SAS	Social Adjustment Scale
SD	standard deviation
SE	standard error
SES	Self-esteem Scale
SF-36	Short Form with 36 Items
SF-36 MC	SF-36 Mental Summary Component
SF-36 PC	SF-36 Physical Summary Component
SFS	Social Functioning Scale
SP	supportive psychotherapy
SRT	Kellner and Sheffield Symptom Rating Test
STAI	State–Trait Anxiety Inventory
STAI-T	State–Trait Anxiety Inventory, trait version
TAU	treatment as usual
TDM	total design method
WL	waiting list

All abbreviations that have been used in this report are listed here unless the abbreviation is well known (e.g. NHS), or it has been used only once, or it is a non-standard abbreviation used only in figures/tables/appendices in which case the abbreviation is defined in the figure legend or at the end of the table.

Executive summary

Objectives

The aim of this study was to consider the following:

- What is the long-term outcome of participants in clinical trials of cognitive behaviour therapy (CBT) for anxiety disorders and psychosis?
- Are there significant differences in effectiveness and cost-effectiveness associated with receiving CBT in comparison with alternative treatments?
- Are there significant differences in effectiveness associated with receiving different intensities of CBT?
- How well can long-term outcome be predicted from data from the original clinical trials?

Design

An attempt was made to contact and interview all of the participants in eight randomised, controlled, clinical trials of CBT for anxiety disorders and two randomised, controlled, clinical trials of CBT for schizophrenia conducted between 1985 and 2001. Case note reviews of healthcare resources used in the 2 years prior to entering the trials and the 2 years prior to follow-up interview were undertaken.

Setting

The clinical trials were conducted in mixed rural and urban settings in five localities in central Scotland. Anxiety disorder trials were conducted mainly in primary care and included three with generalised anxiety disorder, four with panic disorder and one with post-traumatic stress disorder (PTSD). The psychosis studies (one on relapse prevention and one with chronic disorder) were conducted in secondary care.

Participants

An attempt was made to follow up all 1071 entrants to the 10 studies, of whom 125 were not available to be contacted. Of the 946 who were available, 489 agreed to participate (46% of original entrants, 52% of those available to contact).

Method

Follow-up interviews took place between 1999 and 2003, 2–14 years after the original treatment. Interviews for Trials 1–8 were conducted by a research psychologist blind to original treatment condition. Interviews for Trials 9 and 10 were conducted by community psychiatric nurses also blind to treatment condition. Case note reviews were completed following the interview.

Main outcome measures

For Trials 1–8 the main interview-based outcome measures were: Anxiety Disorders Interview Schedule – DSM-IV for diagnosis and co-morbidity, Clinical Global Severity (0–8) and the Hamilton Anxiety Rating Scale. The main patient-rated measures were: Brief Symptom Inventory, SF-36 II, Clinical Global Improvement (1–7), and the Positive and Negative Affect Scale. For Trials 9 and 10 the primary outcome measure was the interview-based Positive and Negative Syndrome Scale (PANSS).

Results

Anxiety disorder studies (trials 1–8)

Over half of the participants (52%) had at least one diagnosis at long-term follow-up, with significant levels of co-morbidity and health status scores comparable to the lowest 10% of the general population. Few participants had none or only mild symptoms (18%) and a significant proportion (30%) had subthreshold symptoms of at least moderate severity. Only 36% reported receiving no interim treatment for anxiety over the follow-up period with 19% receiving almost constant treatment. Patients with PTSD did particularly poorly. There was a 40% real increase in healthcare costs over the two time periods, mainly due to an increase in prescribing. A close relationship was found between poor mental and physical health for those with a chronic anxiety disorder.

Treatment with CBT was associated with a better long-term outcome than non-CBT in terms of overall symptom severity but not with regard to diagnostic status. The positive effects of CBT

found in the original trials were eroded over longer time periods. No evidence was found for an association between more intensive therapy and more enduring effects of CBT. Long-term outcome was found to be most strongly predicted by the complexity and severity of presenting problems at the time of referral, by completion of treatment irrespective of modality and by the amount of interim treatment during the follow-up period. The quality of the therapeutic alliance, measured in two of the studies, was not related to long-term outcome but was related to short-term outcome.

The cost-effectiveness analysis showed no advantages of CBT over non-CBT. For the participants as a whole, CBT was associated with slightly higher costs than non-CBT and slightly higher benefits. For participants who completed CBT, versus all other participants, CBT was associated with somewhat lower costs and slightly higher benefits. The costs of providing CBT in the original trials was only a very small proportion (6.4%) of the overall costs of healthcare for this population, which are high for both physical and mental health problems.

Psychosis studies (Trials 9 and 10)

Outcome was generally poor and only 10% achieved a 25% reduction in total PANSS scores from pretreatment to long-term follow-up. Nearly all participants (93%) reported almost constant treatment over the follow-up period at a significantly higher level than for the anxiety disorder patients. Treatment with CBT was associated with more favourable scores on the three PANSS subscales. However, there were no significant differences between CBT and non-CBT groups in the proportions achieving clinically significant change and very few psychosis patients maintained a 25% reduction in PANSS scores from post-treatment to long-term follow-up regardless of treatment modality.

Cost-effectiveness analysis showed no advantages of CBT over non-CBT. Healthcare costs fell over the two time periods mainly owing to a reduction in inpatient costs.

Conclusions

The implications for healthcare are:

- Psychological therapy services need to recognise that anxiety disorders tend to follow a chronic course and that good outcomes with CBT over the short term are no guarantee of good outcomes over the longer term.
- Clinicians who go beyond standard treatment protocols of about 10 sessions over a 6-month period are unlikely to bring about greater improvement.
- Poor outcomes over the long term are related to greater complexity and severity of presenting problems at the time of referral, failure to complete treatment irrespective of modality and the amount of interim treatment during the follow-up period.
- The relative gains of CBT are greater in anxiety disorders than in psychosis.

Recommendations for future research

Longitudinal research designs over extended periods of time (2–5 years), with large numbers of participants (500+), are required to investigate the relative importance of patient characteristics, therapeutic alliance and therapist expertise in determining the cost-effectiveness of CBT in the longer term.

A better understanding of the mechanisms by which poor treatment responders become increasingly disabled by multiple physical and mental disorders will require close collaboration between researchers in the clinical, biological and social sciences.

Chapter 1
Introduction

Overall aims

Cognitive behaviour therapy (CBT) is a short-term, problem-oriented method of psychological treatment applicable to a wide variety of clinical problems in psychiatry and general healthcare. Distinguished from other psychotherapies by its breadth of application, its emphasis on practical coping skills and its firm foundation in the research base of psychiatry and clinical psychology, CBT has been the subject of numerous clinical trials.[1] For common mental health problems such as panic disorder and generalised anxiety these trials have found that over the short-term (6–12 months following start of therapy) CBT is significantly more effective, on average, than no treatment or other psychological therapies and is at least as effective as medication. There are also indications of a prophylactic effect that medication lacks. This evidence base has led CBT to become accepted as the treatment of choice for anxiety disorders,[2] not just as a way of resolving current distress, but also as a way of teaching coping skills that enhance a person's capacity to manage periods of heightened stress and hence reduce the probability of future episodes of disorder. There is also growing evidence for the value of CBT as an adjunct to medical and psychosocial management of much less common but potentially more severe disorders such as schizophrenia.[3] However, the evidence base here is still at an early stage of development.

The established efficacy of CBT over the short term, at least with common mental disorders, raises the central question addressed in this report: are the short-term benefits of CBT sustained? The opportunity to address this question arose from the collective experience of the authors in conducting eight clinical trials of CBT for anxiety disorders and two clinical trials of CBT for psychotic symptoms between 1985 and 2001. All of these trials were methodologically rigorous in employing research diagnostic criteria, randomisation to conditions, broadly based outcome measures, analysis of outcome in terms of both statistical and clinical significance, treatment protocols and a variety of experimental conditions including medication, brief treatment and psychoanalytic therapy. None of them, however, included provision for follow-up beyond 12 months post-treatment. Details of these trials are given in Chapter 3.

Extended follow-up data on the effects of psychological therapy are important for several reasons. First, the short-term effects of treatment, however positive, are not necessarily a reliable guide to longer term improvement. The natural history of most psychiatric disorder is variable and episodic in nature and both panic disorder and generalised anxiety disorder (GAD), the two main subjects of this report, are known to follow a fluctuating and often chronic course associated with heavy demands on health service resources.[4,5] We know also that where anxiety and depressive symptoms coexist, which is probably in the majority of cases referred for psychological therapy in the NHS, there is a heightened probability of a chronic and relapsing course.[6] The claim, therefore, that CBT leads to a sustained improvement, and confers a long-term advantage relative to other types of treatment, is an ambitious one. A very powerful treatment technology would be required to influence the overall course of a potentially chronic disorder on the basis of a limited number of hours of personal therapy and it would be unwise to assume sustained benefit in the absence of evidence. In fact, the evidence base is very weak. Only a small number of studies have followed up the outcome of clinical trials of CBT for more than 1 year and these are briefly reviewed in the next section.

A second reason for conducting extended follow-up studies is that the relative costs and also the benefits of various approaches to treatment may be more apparent over the longer term. Standard CBT is more resource intensive than alternatives such as medication or brief intervention and there are obvious increases in costs associated with increasing the dose of therapy (e.g. from 10 to 20 sessions). In routine clinical practice the intensity of psychological therapy is largely uncontrolled and this has left purchasers and decision-makers with considerable uncertainty as to whether the benefits of CBT are worth the additional costs. Clearly there is an onus on those who deliver psychological therapies to do so in the most cost-effective manner possible.[7] There is, however, little

in the way of published reports of health economic analyses of the long-term effects of CBT on which decision-making can be based and a general aim of the present research is to begin systematic investigation of this complex but important area.

Finally, extended follow-up studies can examine the various factors associated with good and poor outcomes. What are the best predictors of long-term outcome? How important is type of treatment in relation to participant characteristics such as demography, symptom severity and social adjustment? Is the quality of the therapeutic alliance, which is known to be related to short-term outcome, also significant over the long term? Does clinical status post-treatment add to the predictive power of information assessed at pretreatment? Answers to these questions are of considerable theoretical and clinical importance.

In summary, the promise of long-term follow-up research is a better understanding of (1) the scope and limitations of CBT in changing the overall trajectory of psychiatric disorder, (2) the relative costs and benefits of CBT in relation to alternative treatments and (3) the factors associated with good and poor outcomes. The key research questions addressed in this report are as follows:

- What is the overall outcome for patients in clinical trials of CBT for anxiety disorders irrespective of treatment condition?
- What proportion of patients who sustain benefit from CBT in the short term sustain this benefit over the longer term?
- Are there significant differences in effectiveness and cost-effectiveness associated with receiving CBT in comparison with alternative treatments?
- Are there significant differences in effectiveness and cost-effectiveness associated with receiving different intensities of CBT?
- How well can long-term outcome be predicted from data from the original clinical trials?

Review of long-term follow-up studies of CBT clinical trials

Methodology

In this section we review studies of the long-term outcome of clinical trials of CBT for the four diagnostic groups studied. Studies were located from searches of MEDLINE, PsycLIT and the Cochrane Controlled Trials Register over the 15-year period between 1990 and 2004. Search terms included generalised anxiety disorder, generalised anxiety disorder, panic disorder, post-traumatic stress disorder, schizophrenia, treatment outcome, clinical trials and follow-up. Searches were also made of secondary sources and by contacting researchers in the field. To be included, studies had to meet the following inclusion criteria: structured diagnostic interviews used to assign the principal diagnosis, random assignment of patients to two or more psychological treatments or control conditions and follow-up of treatment outcomes at 1 year or more post-treatment.

Panic disorder

The shorter term benefits of CBT in the treatment of panic disorder with or without agoraphobia are well established. A recent meta-analysis,[8] comparing the short-term efficacy of psychopharmacological treatments (benzodiazepines and antidepressants), CBT and combination treatments, included 106 studies, pertaining to 222 treatment conditions, 5011 participants at pretreatment and 4016 immediately post-treatment. CBT was more effective than control conditions in the reduction of panic (d = 1.25 versus 0.53), agoraphobia (d = 0.91 versus 0.32), depression (d = 0.64 versus 0.32) and anxiety (d = 1.3 versus 0.51), and as effective as antidepressants and their combination in the shorter term outcome on panic, depression, anxiety and agoraphobic avoidance. The shorter term benefits of CBT for panic disorder with or without agoraphobia appear to be sustained at 6 months,[9–15] at 12 months[16–18] and at 15 months.[19]

Generalised anxiety disorder

There is good evidence that in the shorter term CBT is an effective treatment for GAD. Building on previous work by Chambless and Gillis[20] and by Borkovec and Whisman,[21] Gould and colleagues[22] undertook a meta-analytic review of controlled trials examining CBT and pharmacotherapy for GAD. Gould and colleagues presented results from 35 studies of GAD for short-term (less than 4 months) and 'long-term' (6 months) follow-up. In total 13 studies, that included 22 treatment interventions, utilised CBT. In the short term the mean effect size for CBT on anxiety measures was 0.70. Studies combining cognitive and behavioural techniques yielded a larger effect size (n = 8; d = 0.91) than studies investigating cognitive techniques (n = 3; d = 0.59) or behavioural techniques alone (n = 3; d = 0.51). There were no differences between CBT and pharmacotherapy on measures of anxiety, but CBT was superior to pharmacotherapy on measures of depression (0.77 versus 0.46).

At 6 months, in six studies with a mean follow-up rate of 77%, effect sizes associated with CBT on measures of anxiety remained large, between 0.69[23] and 1.24.[24] The positive effects of CBT at 6 months post-treatment were confirmed by Fisher and Durham,[25] who investigated clinically significant change using the Jacobson method. Both applied relaxation and individual CBT were associated with strong recovery rates of 63 and 48%, respectively, at the end of treatment and 60 and 51% at 6-month follow-up.

Table 58 (Appendix 1) summarises the results of 10 clinical trials that provide some evidence of long-term follow-up. The evidence is of variable quality with limited outcome assessment, small numbers of trial entrants and significant attrition over the follow-up period, being the most obvious methodological weaknesses. The length of follow-up ranges from five studies at 12 months,[23,26–29] one study at 18 months,[30] three studies at 24 months,[31–33] and one study at 60 months.[34] With the exception of the final study, which was not designed to assess the maintenance of treatment gains, all the remaining studies report that treatment gains achieved by the end of the trial were maintained over the follow-up period. The proportion of patients achieving clinically significant improvements following CBT varies considerably from trial to trial but, on average, the gains are sustained.

Post-traumatic stress disorder

CBT has been utilised to alleviate the symptoms of post-traumatic stress disorder (PTSD) and there are now a number of studies attesting to the effectiveness of CBT at end of treatment,[35–41] at 3 months,[35,36] at 6 months,[37–39] at 9 months,[40] at 12 months[36] and at 24 months.[42] These promising results are applicable for a range of traumatic events including motor vehicle accidents,[35–37] adult survivors of childhood sexual abuse,[40] natural disasters,[43] amongst refugees[39] and in sexually abused children.[42]

Table 59 (Appendix 1) provides a selection of studies that have reported the long-term efficacy of psychological interventions for PTSD patients that have met diagnostic criteria and have been treated as part of a randomised trial (RT) or randomised controlled trial (RCT). Studies that have reported long-term follow-up data beyond 12 months but have not been incorporated in *Table 59* include, for example: studies where the original sample did not meet PTSD formal diagnostic criteria,[42,44] studies where other diagnostic groups such as acute stress disorder were the primary focus of intervention;[45] studies where the focus of intervention was not primarily the alleviation of PTSD symptoms;[46] or studies where the long-term follow-up focused upon a single cohort of patients.[47] The remaining long-term follow-up studies that are here reported are diverse in nature, as illustrated in *Table 59*. With regard to patient characteristics the populations vary greatly. Of the six studies listed, three included adult civilians of both genders who had developed PTSD following trauma such as accident or assault,[48–50] one study included male and female patients who had attended accident and emergency facilities following physical injury,[51] one study included females only who had been victims of assault, primarily sexual assault,[52] and the remaining study included Vietnam veterans.[53] Among these studies the sample size varied considerably, ranging from as small as seven[48] up to 180.[53]

Most studies used some form of PTSD symptomatology outcome rating as their main measure such as the Impact of Event Scale[54] (IES) or the Clinician-Administered PTSD Scale[55] (CAPS). The high level of co-morbidity among PTSD patients resulted in most studies including the use of measures of depression, such as the Beck Depression Inventory[56] (BDI), anxiety and depression, such as the Hospital Anxiety and Depression Scale[57] (HADS), or anxiety, such as the State–Trait Anxiety Inventory[58] (STAI).

The most frequently investigated forms of treatment in these long-term follow-up studies have included imaginal exposure-based treatments such as that used by Tarrier and colleagues[49] and Power and colleagues,[50] while other studies have included combinations of imaginal plus *in vivo* techniques as used by Richards and colleagues,[48] Foa and colleagues[52] and Bisson and colleagues.[51] Other forms of exposure-based treatments have included trauma-focused group therapy which provides patients with opportunities for exposure to their own traumatic events in addition to vicarious exposure to traumatic events experienced by other group members.[53] There is a dearth of studies that have assessed the long-term efficacy of eye movement desensitisation and reprocessing (EMDR).[50]

In relation to long-term follow-up, there appears to be little evidence that favours strongly any one approach above another, the exception being Foa and colleagues' result that Prolonged Exposure was superior to Stress Inoculation Training and to Prolonged Exposure plus Stress Inoculation

Training on measures of social adjustment and anxiety.[52] The extent of long-term follow-up amongst these studies is also limited with three studies extending to 12 months, one study to 13 months, one study to 18 months and one study to 24 months. On the basis of the limited data available, it is reasonable to conclude that there is still evidence of treatment gains at long-term follow-up but the magnitude of gains appear to have lessened, probably owing to the complexity of the disorder and the high level of co-morbid symptoms of depression in particular. Of particular note amongst these studies is a lack of data regarding health service utilisation between end of treatment intervention phase and long-term follow-up assessment period. It is therefore difficult to assess to what extent long-term gains are due to initial therapies or subsequent interventions provided during the follow-up period.

Schizophrenia

There is growing evidence that CBT is effective in the alleviation of persisting positive psychotic symptoms.[59–67] All these trials have involved providing CBT in conjunction with antipsychotic medication. CBT has been found to be superior to waiting list control,[59] structured activity and informal support,[60,61] routine care/treatment as usual (TAU),[62,66,67] supportive counselling and routine care,[63] supportive therapy in combination with clozapine,[64] befriending (BF)[65] and supportive psychotherapy (SP).[66]

Table 60 (Appendix 1) shows the basic design, participants, outcomes and follow-up periods of trials of RCTs involving CBT for patients diagnosed with schizophrenia or similar condition, who had been followed up for longer than 12 months post-randomisation. A total of 23 papers describing 11 RCTs involving the comparison of CBT with at least one other treatment condition, involving a total of 1175 participants with a diagnosis of schizophrenia or similar, were identified. The studies were heterogeneous in terms of the participant populations and outcome measures used. One study[68,69] focused on participants with co-morbid substance use problems. Four studies[61,70–72] delivered cognitive therapy during the acute/recovery phase. Three studies[65,73,74] focused on those with drug-resistant symptoms. One study[75] studied group cognitive therapy. Two studies[76,77] had relapse or readmission as their primary outcome. One study[78] focused on the development of insight as the primary outcome.

Although a wide variety of outcome measures were used, there were three main primary outcome measures employed, the Brief Psychiatric Rating Scale[79] (BPRS), the Comprehensive Psychiatric Rating Scale[80] (CPRS), and the Positive and Negative Syndrome Scale[81] (PANSS). The BPRS was employed in the studies by Buchkremer and colleagues,[76] Garety and colleagues,[73] Haddock and colleagues,[70] Startup and colleagues[72] and Tarrier and colleagues.[74] The CPRS was used in the study by Sensky and colleagues.[65] The PANSS was used in the studies by Barrowclough and colleagues[68] and Lewis and colleagues.[71] Other studies[61,78] used the Psychiatric Assessment Scale[82] (PAS) and the David Insight Scale.[83]

CBT was compared with TAU in seven trials.[62,68,70–72,74,78] A total of six trials[61,65,71,74,76,77] incorporated a comparison psychological intervention into the study design.

Long-term follow-ups were conducted at 18 months,[65,69,72,84–86] 24 months,[70,87,88] 3 years[89] and 5 years.[90,91] Therefore, the best evidence for the maintenance of treatment gains post-treatment comes from the six studies that have conducted 18-month post-randomisation assessments. In the studies by Barrowclough and colleagues,[69] Garety and colleagues,[84] Kemp and colleagues[85] and Startup and colleagues,[72] immediate treatment gains are largely maintained at follow-up. All of these studies compared CBT with TAU. The study by Sensky and colleagues[65] did not find a difference between CBT and BF at post-treatment. However, at 18 months those who received CBT had continued to improve, whereas those who received BF had lost much of their gains. The study by Tarrier and colleagues[86] did not find a specific effect for CBT at 18 months, rather receipt of psychological intervention appeared to improve outcome although this positive finding did not translate to relapse or readmission outcomes. At 2 years there is no evidence for a specific effect for CBT.[87] The study by Hogarty and colleagues[89] did find an effect for personal therapy for those patients who lived with families at 3 years but treatment had been continuous over that period. Finally, there was little evidence for CBT at 5-year follow-up.[90,91] However, the study by Drury and colleagues[91] found that those who relapsed more than once during the intervening period had a very poor outcome, raising the importance of relapse prevention for this group. There was a little evidence from the study by Buchkremer and colleagues[90] that receipt of CBT plus psychoeducation plus counselling protected participants against relapse compared

to leisure activity control over the intervening 5 years.

The Cochrane Review[92] concluded that "the use of cognitive behavioural therapy has been associated with some reduction in symptoms, especially the positive symptoms of schizophrenia. However, there is considerable variability in the findings of the various studies and, at present, it is not possible to assert any substantial benefit for cognitive behavioural therapy over standard care or supportive therapies." We would argue that some of this variability in findings is due in part to a number of factors, the measures employed by trialists and the different populations investigated in trials. It does not seem reasonable to compare CBT for stable yet drug-refractory positive symptoms with CBT delivered during the acute phase. In addition, it has recently been argued by Birchwood[93] that the outcomes for CBT should not be the same as the outcomes for antipsychotic medication. In an example of this, Trower and colleagues[94] have reported positive outcomes for their trial of cognitive therapy for command hallucinations. There are also limited data pertaining to the maintenance of therapy gains following treatment; further research is needed in this regard.

Overall summary of findings

Follow-up assessment of the results of CBT trials beyond 1-year post-treatment are relatively few in number and are limited by small numbers, significant attrition over the follow-up period and often limited assessment. There are only three studies that have followed up CBT trials for more than 2 years and there are no studies that have attempted to collect systematic data on diagnosis, health status, quality of life and symptom severity. Within these limitations, the general finding with respect to anxiety disorders is that the effects of CBT are sustained. The evidence base with respect to psychosis is much weaker, with variable findings in respect of a sustained improvement following CBT.

Structure of the report

This report can be conveniently divided into three sections. Background information on the 10 clinical trials and methodology of the long-term follow-up are covered in Chapters 2–5. The results of the various investigations are described in Chapters 6–10. The concluding three chapters provide a summary of the overall findings, a description of past and future plans for dissemination of the results and a discussion of the implications of the results for healthcare together with recommendations for future research. There were some differences in methodology between the anxiety disorder studies (Trials 1–8) and the psychosis studies (Trials 9 and 10) and these are described separately within each chapter where appropriate. Similarly, the analyses of data were conducted separately for the two sets of aggregate data for the anxiety disorders studies (Trials 1–8) and the psychosis studies (Trials 9 and 10) and these are also reported separately in the later chapters. Within the anxiety disorder studies, some additional analyses were conducted on aggregate data across treatment conditions within diagnostic groupings since these data sets provided the clearest means of addressing some of the research questions (e.g. CBT versus non-CBT within GAD Trials 1 and 2, standard versus high contact CBT within GAD Trials 2 and 7, low contact versus standard CBT within panic disorder Trials 4 and 5).

Throughout the report the emphasis is on giving a clear account of how the study was conducted and what the various analyses revealed. For the sake of brevity, in what is a lengthy report, references to relevant literature have been kept to a minimum.

Chapter 2
Methods

Protocol and procedures

All of the participants in Trials 1–8 had originally been diagnosed and treated for a specific anxiety disorder, namely GAD, panic disorder or PTSD. Patients in Trials 9 and 10 had originally been diagnosed with schizophrenia, schizo-affective disorder, delusional disorder or a related disorder. All patients were followed up at a period of between 2 and 14 years after the original treatment, and long-term follow-up interviews took place between 1999 and 2003 inclusive. There was an attempt to follow up all those who were entered into the original trials since it was felt that drop-outs, non-attenders and completers were equally important in gaining a comprehensive view of the long-term course of anxiety and emotional disorders.

Tracing patients at long-term follow-up

The same overall procedure for tracing patients was followed in all 10 trials as approved by the respective medical research ethics committees in the four localities (the Tayside Committee on Medical Research Ethics for the Dundee trials; the Forth Valley Medical Research Ethics Committee for the Stirling trials; the Fife Health Board Committee on Medical Research Ethics for the Fife trials; and the Research Ethics Committees in Ayrshire and Arran and Greater Glasgow Health Boards for the West of Scotland trial). Patient details from the original trials were verified with local medical records and attempts were made to trace patients who had moved using central medical records. A number of patients (n = 60, 6% of those entered into the original trials) proved to be untraceable at long-term follow-up.

Procedure for GP and patient contact

Anxiety disorder studies (Trials 1–8)

For the eight anxiety disorder studies, permission to contact all successfully traced patients was sought from each patient's general medical practitioner and if this was refused the patient was excluded from the study. As many of the patients in Trial 2 had originally been referred by psychiatrists, permission to contact those patients from Trial 2 who were still being treated in secondary care was also sought from the relevant consultant. Medical practitioners were contacted by letter, which included details of the original trial and follow-up study, a reply slip, and a copy of the Patient Information Sheet. Both GPs and consultants were advised that, if no response was received from them within 2 weeks, it was assumed that consent was given [the exception to this was for patients outwith areas covered by the research ethics committees (see the section 'Tracing patients at long-term follow-up', above) where specific GP approval was required]. GPs were also asked in this letter to consent to the research psychologist having access to medical case notes, providing that individual patients gave consent for this access. An example of the GP letter and Patient Information Sheet is included in Appendix 2.

Patients residing in Scotland were contacted by letter and invited to complete the main outcome questionnaires and to participate in a 1-hour interview if they were within reasonable travelling distance. [Reasonable travelling distance was deemed to be within approximately 100 miles of the two main research bases (i.e. Stirling and Dundee). Only three patients available to contact (i.e. traced and GP approval given) fell outside this area, of which one returned questionnaires.] Patients living outside Scotland were sent questionnaires only. The personalised patient letter included information about the original trial and the follow-up study, a reply slip and pre-paid envelope and a copy of the Patient Information Sheet. An example of the patient letter is included in Appendix 2.

Where possible, all practical methods for improving response rates, as highlighted in previous research[95,96] [e.g. the total design method (TDM)], were employed, including the use of personalised letters which referred to each patients' own GP; letters being personally signed (in blue ink) by the principal therapist from each of the original trials (who would be known to each patient); the use of official stationery to gain trust; carefully worded letters and information sheets highlighting the benefits of the research and the usefulness of individual responses (including the opportunity for open-ended replies); stamped and addressed (i.e. not business) reply envelopes; follow-up letter at a judicious time interval, which

included the additional options of completing questionnaires (included in letter) or telephone interviews to reduce the time requirement for each patient. Unfortunately, the study design meant that some additional methods identified in the TDM described by Dillman[95] could not be employed, such as:

- reducing the length of questionnaires, printing the questionnaire in a booklet format or using graphical design to improve the readability of questionnaires (as these were determined by the outcome measures used in the original studies, some of which were copyrighted and could not be re-formatted);
- the use of a third and fourth follow-up letter, this was for ethical reasons. Although Tyrer and colleagues[97] have presented a case for cold-calling in long-term follow-up studies of patients with neurotic disorders, as a means of greatly increasing response rates, our ethical approval did not permit such procedures, and indeed, permitted no contact without GP approval. In addition, in our experience, a number of patients who had agreed to be interviewed and seemed happy to make an appointment when telephoned, later seemed to show unwillingness by withdrawing their presence (either by repeatedly failing to attend scheduled interview appointments or by failing to answer the door during arranged home visits). Two-thirds of the long-term follow-up participants were asked why they had agreed to take part in the long-term follow-up interview and, whilst the majority (71%) cited 'wanting to help' as their main reason, a minority (5%) said they only agreed because they felt under pressure to do so (either through guilt or family coercion); or
- financial inducements, as funding was not available. Reimbursement for patient travel costs was not a major issue in the current study (with only one request) as interviews were held at the most convenient location for participants, mainly in local GP surgeries or the patients' homes.

Every effort was made to interview participants at a time and place of their convenience, either at the research base, a local clinic or the patient's home. Patients were offered a choice of morning, afternoon or early evening appointments.

Psychosis studies (Trials 9 and 10)

As the majority of patients from Trials 9 and 10 were still being treated in secondary care, permission to contact these patients at long-term follow-up was first sought from the appropriate clinical psychologist or psychiatrist, and if this was refused the patient was excluded from the study. Psychologists and psychiatrists were contacted by letter in the first instance, which included details of the original trial and follow-up study, a reply slip and a copy of the Patient Information Sheet. Patients were not contacted unless permission was explicitly given by the consultant. Permission was also sought from the consultant for the researcher to have access to medical case notes held in secondary care, provided that individual patients gave consent for this access. An example of the consultant letter and Patient Information Sheet for Trials 9 and 10 is included in Appendix 2.

Once permission to contact a patient had been obtained, his or her key worker was contacted by the researcher, informed of the purposes and procedures of the long-term follow-up study and advised that their patients would soon be invited to take part. Advice on the best approach with regard to interviewing patients was also sought from the key worker, and also their help in explaining the purposes of the research to the patient. Patients were also contacted by letter, which invited them to complete the main outcome questionnaires and to participate in a 1-hour interview if they were within reasonable travelling distance. In the case of patients in Trial 10, the letter also included a date and time when the researcher would visit the patient, and asked the patient to contact the researcher if they did not wish to take part. In Trial 9, agreement to participate was received either by direct response from the patient or via the key worker. The personalised patient letter included information about the original trial and the follow-up study, a reply slip and pre-paid envelope, and a copy of the Patient Information Sheet.

Every effort was made to interview participants at a time and place of their convenience, either at the research base, a local clinic or the patient's home. Patients were offered a choice of morning or afternoon appointments.

Procedure for non-responders to first contact

Anxiety disorder studies (Trials 1–8)

Those patients who failed to respond to the initial letter within 3 weeks were sent a reminder, which included the main outcome questionnaires. This letter offered those patients not wishing to attend an interview the additional options of either a short telephone interview plus questionnaires or the completion of questionnaires only. After the first and second contact letters, 11 patients (3%) consented to be interviewed by telephone. These telephone interviews used the same diagnostic tool

as the face-to-face interviews but, because of time constraints, covered only the anxiety and depressive disorders, and were therefore less comprehensive and may have been less reliable than the interview-based assessment. Finally, 12% (n = 48) of long-term follow-up participants from Trials 1–8 were not interviewed but chose to return the self-report questionnaires [this includes a number of patients (usually repeated no-shows) who agreed to be interviewed after the first contact, but later switched to returning questionnaires only]. Diagnostic information at long-term follow-up is not available for these patients.

After ethical consideration, it was agreed that a maximum of two contacts should be made to this susceptible group of patients. This decision was supported in practice by several patients registering complaints after the second contact was made and also a number of refusals after the first contact, particularly in Trial 6 (originally treated for PTSD), where the main reason cited for refusing to take part was not wanting to be reminded of the traumatic event. In addition, the second letter resulted in only a small increase in the response rate for interviewed patients (equivalent to around 4% of the total number of contactable patients), and a slightly higher increase (less than 6%) in those completing questionnaires only. It is therefore unlikely that a third letter to non-responders would have resulted in a significant improvement in response rates, particularly of patients agreeing to be interviewed.

Psychosis studies (Trials 9 and 10)

Those patients in Trial 9 who failed to respond to the initial letter were contacted either directly by the researcher or by the key worker (depending on the key worker's advice), to ascertain whether or not they wished to take part in the study. Patients not responding in the affirmative to this second approach were assumed to have declined to take part in the long-term follow-up study. In the case of patients in Trial 10 who were not available for interview at the date and time suggested in the initial letter (e.g. they were not at home at the specified time), contact was re-established, where possible, via the key worker. If contact was not possible via this means, they were assumed to have refused consent.

Patient consent

Anxiety disorder studies (Trials 1–8)

Those patients attending interview were asked to complete a consent form before the interview commenced. Patients were given the opportunity to ask questions regarding the study at this point, and their rights regarding the study data were explained to them. Patients gave consent to take part in the study and, separately, consent to allow the research psychologist access to case notes, in order to collect data relevant to the study. A copy of the consent form is included in Appendix 2.

Patients returning questionnaires only were asked to complete a form regarding access to medical case note data. Only 7% (n = 28) of patients in Trials 1–8 either refused or failed to give permission to access medical case notes.

Psychosis studies (Trials 9 and 10)

Those patients attending interview were asked to complete a consent form before the interview commenced. Patients were given the opportunity to ask questions regarding the study at this point, and their rights regarding the study data were explained to them. Patients gave consent to take part in the study and, separately, consent to allow the researcher access to case notes, in order to collect data relevant to the study. A copy of the consent form is included in Appendix 2.

Patient interviews

Anxiety disorder studies (Trials 1–8)

Interviews for the eight anxiety disorder studies were all conducted by the same research psychologist (JAC) who was blind to treatment condition. This interview consisted of (1) the Anxiety Disorders Interview Schedule IV[98] (ADIS-IV) to assess diagnostic status according to DSM-IV criteria; (2) a structured interview to assess attitude to original treatment and current coping; and (3) an assessment of amount of treatment for mental health problems since participation in the original trial. These face-to-face semi-structured interviews were held with the majority (85%) of long-term follow-up participants for Trials 1–8. Interviews lasted an average of 1 hour (range 35–150 minutes). In addition to the semi-structured interview, patients completed a number of self-report questionnaires.

Psychosis studies (Trials 9 and 10)

Community psychiatric nurses (CPNs) were employed on a secondment basis to carry out the interviews and collect case note data in each of the three geographical regions (i.e. Ayrshire, Glasgow and Dundee) involved in the two psychosis studies. Each CPN was blind to original treatment condition. This interview consisted of (1) the PANSS,[81] a clinician-rated scale measuring positive and negative symptoms of schizophrenia and related disorders; (2) a structured interview to assess attitude to original treatment and current coping; and (3) an assessment of amount of

treatment for mental health problems since participation in the original trial. These face-to-face semi-structured interviews were held with all long-term follow-up participants for Trials 9 and 10. On average, interviews lasted between 1¼ and 1½ hours (range 60–210 minutes). In addition to the semi-structured interview, patients completed a number of self-report questionnaires.

Economic evaluation methods

Economic evaluation is the comparative analysis of two alternative courses of action in terms of both their costs and consequences. In the current study, the key question of interest relates to the economic analysis of CBT across a range of diagnostic categories over the long term. A range of methodological approaches is available for such analysis. The first of these is cost minimisation analysis. Here it is necessary that the effects across two or more interventions are identical in all respects. Clearly this could not be assumed *ex ante* in this instance and so this approach was not appropriate. With respect to the measurement of benefit, given the range of diagnoses and the length of time since intervention (and so the potential for aetiology to have changed significantly), it was crucial that the measurement of benefit was broad and as encompassing as possible. Additionally, it was also felt that the mental health diagnoses under consideration have the potential to impact substantially on physical health, especially over a chronic, long-term perspective and as such the benefit measure should also encapsulate these aspects of health. As a consequence of these considerations, it was concluded that the generic Short Form with 36 Items (SF-36) (i.e. the UK SF-36 version II[99]) offered the most comprehensive assessment of health status and therefore potential benefit in these client groups. A detailed description of this measure and its use in this study can be found in the section 'Main outcome measures' (p. 27).

As a result of these considerations the economic evaluation focused on the following primary question:

> **Is CBT a cost-effective intervention across a range of diagnoses when compared with alternative regimes over the long term?**

A slightly different approach was adopted when considering the anxiety disorder studies (Trials 1–8) to that for the psychosis studies (Trials 9 and 10) and these are elucidated below.

Identification and measurement of resource use

Anxiety disorder studies (Trials 1–8)

Where both GPs and patients gave their consent, data on resource use were collected from the patients' general practice case notes by the same research psychologist who had conducted the interviews. For each patient, data were collected for a period of 2 years prior to entry into the original trial. Additionally, data were collected for the 2 years prior to either the date of the long-term follow-up interview or the date of return of questionnaires, whichever applied to that patient. In no case did these two time periods overlap. The objective was to capture the cost of all resource use prior to the trial and prior to long-term follow-up. The key data of interest for the purpose of exploring differences in the efficiency of CBT compared with other treatments were those relating to resource use in the follow-up period. For the purposes of the analysis of cost-effectiveness it was assumed that resource use did not differ by allocated treatment group [this assumption was later confirmed by analysis whereby the mean total costs for pre-trial in the CBT group were £925 (SD = 1154) and in the non-CBT group £1202 (SD = 2505) ($t = 0.988$, $p = 0.326$)]. Pre-trial data were, however, collected to allow an exploration of changes in cost over time.

Data on **all** NHS resource use were collected (capturing both physical and psychological aspects of care) and were recorded on a standard data collection form, included in Appendix 3. The following items of resource use were identified from case notes, collated and measured.

Primary care contacts
All contacts with members of the primary care team were recorded. Additionally, out-of-hours and home visits were recorded separately, as were telephone calls.

Prescribing
All prescribing was recorded, including actual dosage and duration. It should be noted, however, that there were frequent omissions to these data, as a consequence of either omission in recording or the illegibility of records.

Direct referrals by primary care
All tests requested by primary care were recorded from case notes, such as blood tests and X-rays.

Secondary care
All visits to the secondary, and indeed tertiary, care sectors were recorded. These included inpatient

stays, outpatient attendances and the use of specialist mental health services.

Psychosis studies (Trials 9 and 10)

For Trials 9 and 10, a slightly modified approach was adopted in recognition of the significant differences in the pattern of healthcare resource use that accompanies a diagnosis of psychosis, and also the specific issues associated with research in this client group. Here, where both the patient and the appropriate consultant gave their consent, case note data were collected from the relevant **secondary** care records. This is unlikely to have had a material impact on the resource use information collected given the well documented pattern of resource use in this client group, which is overwhelmingly dominated by secondary and specialist healthcare provision rather than primary care.[100] Again, data collection was carried out by the same researcher who conducted the interview. In addition, for Trial 9, ethical approval was granted to access case notes for those patients who did not take part in the long-term follow-up process provided that the appropriate consultant had granted consent. For each patient, data were collected for a period of 2 years prior to entry into the original trial and for 2 years prior to the long-term follow-up interview and were recorded on a standard data collection form, included in Appendix 3. In the case of patients in Trial 9 who were not interviewed, the date selected was the date of consent by the consultant.

Secondary care contacts

All contacts with members of the secondary care specialist mental health team were recorded. This included inpatient stays, outpatient attendances and use of community-based specialist services.

Prescribing

All psychotropic medication was recorded, including actual dosage and duration.

Contacts with other services

This client group had significant levels of contact with other services outside the NHS that were clearly a key component of their care package, such as social work and advocacy services. As a consequence, for these two studies these costs have also been measured and valued where they were delivered alongside secondary care as a key component of that care package. Costs do not include the broader input delivered to this client group by such services. Social work salary costs were sourced from South Ayrshire Council Social Work Department. The same base year of costs was applied to these professional groups.

Valuation of resource use

Anxiety disorder studies (Trials 1–8)

Having collected data on items of resource use, it was then necessary to apply a unit cost to each piece of consumption. With respect to each item of resource use, a base year of 1998 was adopted, to be applied to resources **regardless of the year in which it occurred**. This approach ensured that any differences are genuine differences in the use of resources and cannot be attributed to changes in the prices of products and services.

Primary care contacts

All primary care contacts were costed using estimates produced by Netton and Dennett in 1997.[101] These costs were converted to 1998–99 values using an NHS inflation index[102] from the Department of Health, Leeds. The unit costs applied are detailed in *Table 1*.

Prescribing

Costs for prescribing were sourced from the BNF, 1998 (No. 36).[103] Where items of resource use were not available owing to innovation in medical technology, costs were applied from the BNF where they first appeared.

As noted earlier, there was a need to impute a significant amount of data in relation to prescribing. Where the item prescribed was determinable, but volume or dosages were unclear, the average cost for that specific item, from that period within that trial, was applied. If necessary, the other time period was utilised, that is if the data were missing from pre-long-term follow-up, they were imputed from the pre-trial period, and vice versa. As a final resort, the average from the whole cohort of trials was used. A similar approach was adopted where the item was illegible, but clearly a prescription, but the average taken was that for prescribed items generally. Imputation affected 314 individuals in pre-trial resource use and 295 individuals in pre-follow-up resource use.

TABLE 1 *Primary care contacts unit costs for Trials 1–8*

Primary care contact	Unit cost (£) (1998–99 prices)
GP consultation	9.52
Practice nurse consultation	6.35
GP home visit	28.57
GP out-of-hours home visit	28.57
GP telephone consultation	12.70

Direct referrals by primary care
Unit costs for tests requested by primary care were supplied by Ayrshire and Arran Acute Hospitals NHS Trust (that is, those diagnostic and laboratory tests utilised by GPs for the purposes of their own diagnosis and not those that were undertaken as part of an outpatient attendance or inpatient stay). This Trust was selected because of the difficulty of obtaining site-specific data across all of the trials. This Trust is a typical District General Hospital provider. Again, costs were converted to the base year, using the NHS index available from the Department of Health.[102] There was also some lack of specificity in this area, for example, reporting an X-ray but not the site. A small number of areas therefore required some imputation. A total of 12 items were imputed in this way.

Secondary care contacts
Costs for all outpatient attendances were obtained from Scottish Health Service Costs,[104] again for 1998–99. Where data were missing, the following rules were applied:

- If the resource use took place within the private sector, this was excluded from the analysis. This was a result of no other private sector costs, such as over-the-counter medicines, being captured by the method adopted.
- If the site was not provided, the average cost across all Scottish hospitals was applied for the relevant specialty.
- If the site was named but the specialty was absent, this was treated as general medicine.
- If the number of visits was unspecified, but known to be greater than one, the average number of visits for that site and specialty was utilised.

The incidence of data imputed in this way is detailed in *Table 2*.

Costs for inpatient bed days were obtained from Scottish Health Service Costs for 1998–99[104] and applied to length of stay data collected from patients' records. Where data were missing, the following rules were applied:

- If the resource use took place within the private sector, this was excluded from the analysis. This was a result of no other private sector costs, such as over-the-counter medicines, being captured by the method.
- If the site was not provided, the average cost per inpatient day across all Scottish hospitals was applied for the relevant specialty.
- If the number of bed days was unspecified, but known to be greater than one, the average length of stay for that site and specialty was utilised.
- In some patients, the ward number was specified rather than the clinical specialty. In these instances, sites were contacted directly to ascertain the specialty.

The incidence of data imputed in this way is detailed in *Table 3*.

TABLE 2 *Outpatient missing data imputation for Trials 1–8*

Type of missing data	Pre-trial period (Trials 1–8)	Pre-follow-up period (Trials 1–8)
Specialty, no site	16	22
Site, no specialty	8	6
No. of visits unspecified	12	12
Private care	15	10

TABLE 3 *Inpatient missing data imputation for Trials 1–8*

Type of missing data	Pre-trial period (Trials 1–8)	Pre-follow-up period (Trials 1–8)
Specialty, no site	1	0
Site, no specialty	3	1
Length of stay unspecified	0	2
Private care	1	1
Ward number rather than specialty	26	29

Psychosis studies (Trials 9 and 10)

Having collected data on items of resource use, it was then necessary to apply a unit cost to each piece of consumption. With respect to each item of resource use, a base year of 1998 was adopted to be applied to resources regardless of the year in which it occurred. This approach ensured that any differences are genuine differences in the use of resources and cannot be attributed to changes in the prices of products and services.

Prescribing

Costs for prescribing were sourced from the BNF, 1998 (No. 36).[103] Where items of resource use were not available owing to innovation in medical technology, costs were applied from the BNF where they first appeared.

Again, where the item prescribed was determinable, but volume or dosages were unclear, the average cost for that specific item, from that period within that trial, was applied. If necessary, the other time period was utilised. As a final resort, the average from the whole cohort of trials was used. No issues of illegibility existed in Trials 9 and 10. Imputation affected 22 individuals in pre-trial resource use and 31 individuals in pre-follow-up resource use. Missing prescribing data were much less of an issue for Trials 9 and 10, probably as a consequence of using secondary care data sources only.

Secondary care contacts

Published costs available for outpatient contacts are based either on the cost for a referral and the subsequent package of care or per attendance at a particular unit. It was concluded that, given the high intensity of contact by this client group both in terms of the length and number of visits, such an approach would mask considerable variations in the use of Community Mental Health Team (CMHT) resources. Consequently, salary costs for all contacts with NHS professionals were obtained from the relevant NHS circulars for 1999. Standard rates of NHS on-costs, holidays and working hours were utilised for each professional group. Each contact was assumed to be 45 minutes of contact time with 15 minutes of non-contact time per attendance.

Costs of attendance as a day case were obtained from Scottish Health Service Costs,[104] again for 1998–99. Where data were missing, the following rules were applied:

- Where the nature of the professional contact was known, but the number of visits unknown, the average number of contacts with that profession for that phase of the study was used.
- Where the profession of the contact was not specified, the average salary cost for all health professionals involved in the study was utilised.
- If the number of visits was unspecified, but known to be greater than one, the average number of visits for that site and specialty was utilised.
- If the site was unknown, the average cost for Scotland was utilised.

The incidence of data imputed in this way is detailed in *Tables 4* and *5*.

Costs for inpatient bed days were obtained from Scottish Health Service Costs for 1998–99[104] and applied to length of stay data collected from patients' records. Where data were missing, the following rules were applied:

- If the site was unknown, the average cost for Scotland was utilised.

The incidence of data imputed in this way is detailed in *Table 6*.

TABLE 4 *CMHT contact missing data imputation for Trials 9 and 10*

Type of missing data	Pre-trial period (Trials 9 and 10)	Pre-follow-up period (Trials 9 and 10)
No. of contacts unspecified	3	3
Contact designation unknown	30	11

TABLE 5 *Outpatient missing data imputation for Trials 9 and 10*

Type of missing data	Pre-trial period (Trials 9 and 10)	Pre-follow-up period (Trials 9 and 10)
No. of visits unspecified	0	0
Site unspecified	0	2

TABLE 6 *Inpatient missing data imputation for Trials 9 and 10*

Type of missing data	Pre-trial period (Trials 9 and 10)	Pre-follow-up period (Trials 9 and 10)
Site unspecified	1	0

Chapter 3

Setting

Summary of CBT clinical trials in central Scotland

This research is a long-term follow-up of 10 clinical treatment trials for anxiety and emotional disorders, which were held throughout central Scotland between 1985 and 2001. These 10 trials are reported in detail below.

Trial 1 – GAD, Stirling[24]

Trial 1 was an RCT of diazepam (DZ), placebo (PL), cognitive behavioural therapy alone (CBT), cognitive behavioural therapy plus diazepam (CBT+DZ), and cognitive behavioural therapy plus placebo (CBT+PL). The trial was based at the Anxiety and Stress Centre, University of Stirling, between 1985 and 1988. All patients were recruited from primary care health centres and had a primary diagnosis of GAD according to DSM-III criteria.[105] Treatment was carried out by two clinical psychologists, with at least 2-years post-qualification experience. The CBT used in Trial 2 was based on an abbreviated form of the approach used by Beck and Emery.[106] A full description of the trial method can be found in the original report.[24] Self-report and clinician-rated measures of outcome taken both at end of treatment and at 6-month follow-up showed preferential results for all three CBT groups over DZ or PL, and the CBT+DZ group fared best overall.

Trial 2 – GAD, Dundee[107]

Trial 2 participants took part in an RCT of cognitive therapy (CT), analytical psychotherapy (AP) or anxiety management training (AMT). The trial was based at the Department of Psychiatry, University of Dundee, between 1989 and 1991. Patients were recruited from either GPs in primary care or from psychiatrists in secondary care and all had a primary diagnosis of GAD according to DSM-III-R criteria.[108] Both the CT and AP treatments were delivered at two levels of intensity, i.e. high, with between 16 and 20 sessions over 6 months, and low, with 8–10 sessions over the same period. AMT was delivered at low contact only. The method of CT used in Trial 2 was based on the approach described by Beck and Emery.[106] CT was delivered by two clinical psychologists, AP by three consultant psychiatrists and AMT by seven trainee psychiatrists under supervision. A full description of the trial can be found in the original report.[107] At post-treatment and 6-month follow-up, outcome on both self-report and clinician-rated measures was consistently superior for CT over AP, especially in the high contact group. AMT showed a slightly better outcome than AP on some of the measures, as did CT over AMT. Overall, those in the low contact treatment groups fared worse than those in the high contact groups, particularly at 6-month follow-up.

Trial 3 – panic disorder, Stirling[109]

Trial 3 was an RCT of fluvoxamine (FL), placebo (PL), cognitive behavioural therapy alone (CBT), cognitive behavioural therapy plus fluvoxamine (CBT+FL), and cognitive behavioural therapy plus placebo (CBT+PL). Patients were GP referred and all had a primary diagnosis of panic disorder, with or without agoraphobia, as defined by DSM-III-R criteria.[108] The trial was based at the Anxiety and Stress Centre, University of Stirling, between 1989 and 1993. Treatment was carried out by one clinical psychologist, with 7-years post-qualification experience, and the treatment groups were balanced for therapist contact. Areas targeted in treatment were those outlined by Barlow.[110] A full description of the trial method can be found in the original report.[109] All active treatment groups had showed improvement at the end of treatment. Self-report and clinician-rated measures of outcome taken both at end of treatment and at 6-month follow-up showed preferential results for the three CBT groups over FL and PL, and the CBT and CBT+FL groups fared best overall.

Trial 4 – panic disorder, Stirling[111]

Trial 4 was an RCT of different levels of therapist contact in CBT for panic disorder. Patients were GP referred and all had a primary diagnosis of panic disorder, with or without agoraphobia, as defined by DSM-III-R criteria.[108] The trial was based at the Anxiety and Stress Centre, University of Stirling, between 1994 and 1995. The three treatment groups in Trial 4 were standard CBT (6 hours of therapist contact over 12 weeks), minimum contact CBT (2 hours of therapist contact over 12 weeks) and bibliotherapy (1.5 hours of therapist contact for assessment only

over 12 weeks). All treatment groups received a patient manual of CBT treatment for panic disorder. CBT treatment was similar to that of Barlow and Cerny.[112] Treatment was carried out by the same clinical psychologist as in Trial 3. A full description of the trial method can be found in the original report.[111] The standard and minimum contact CBT groups showed significant improvement on both self-rated and clinician-rated outcome measures at post-treatment, and the standard contact group fared significantly better than the bibliotherapy group. The standard contact group continued to fare best at 6-month follow-up.

Trial 5 – panic disorder, Stirling[113]

Trial 5 was an RCT of group versus individual CBT for panic disorder. Patients were GP referred and all had a primary diagnosis of panic disorder, with or without agoraphobia, as defined by DSM-IV criteria.[114] The trial was based at the Anxiety and Stress Centre, University of Stirling, between 1995 and 1997. Patients referred to Trial 5 were randomly allocated to either group CBT or a waiting list (WL) condition. At the end of the WLs patients were allowed to enter either group or individual CBT. All but one patient chose to have individual CBT. In the original report the group CBT condition was compared both with the WL condition and with the standard CBT group from Trial 4. Treatment was carried out by the same clinical psychologist as in Trials 3 and 4. A full description of the trial method can be found in the original report.[113] In the original trial both the group and standard CBT conditions had shown better outcome post-treatment than the WL, and the standard CBT group showed preferential results over the group CBT condition on a measure of clinically significant change. Dropout rates in the group CBT condition were very high (n = 18).

Trial 6 – post-traumatic stress disorder, Stirling[50]

Trial 6 was an RCT of EMDR versus exposure plus cognitive restructuring (E+CR) for PTSD. Patients were referred either by GPs or psychiatrists, and all had a diagnosis of PTSD according to DSM-IV[114] criteria. The trial was run by the Anxiety and Stress Centre, University of Stirling, and was conducted between 1997 and 1999. The three treatment groups in Trial 6 were EMDR (maximum of 10 sessions over 10 weeks), E+CR (maximum of 10 sessions over 10 weeks) and a WL condition (WL) of 10 weeks. Those allocated to WL were randomly assigned to either EMDR or E+CR at the end of the WL period. Treatment was carried out by a behavioural psychotherapist and a research registrar, both of whom had received EMDR training prior to the study. EMDR was structured according to the procedures outlined by Shapiro.[115] E+CR was structured according to a treatment manual for exposure plus cognitive restructuring devised for PTSD as used by Marks and colleagues.[38] A full description of the trial method can be found in the original report.[50] Both the EMDR and E+CR groups showed significant improvements post-treatment and at short-term follow-up, with a slight advantage for EMDR. There was no improvement for those in the WL group.

Trial 7 – GAD, Dundee, Fife and Stirling[116]

Trial 7 was a collaborative study, carried out between the Department of Psychiatry, University of Dundee, the Department of Clinical Psychology, NHS Fife, and the Anxiety and Stress Research Centre at the University of Stirling. The original trial was conducted between 1997 and 1999. Patients were GP referred and all had a diagnosis of GAD according to DSM-IV[114] criteria. Patients were allocated to one of three different intensities of CBT for GAD, i.e. standard CBT, where patients received up to 10 1-hour sessions over 6 months at fortnightly intervals; intensive CBT, consisting of up to 20 1-hour, weekly sessions over 6 months; and brief CBT, consisting of up to five 1-hour sessions over a 6-month period. At entry to the trial, the complexity and severity of presenting problems were assessed for each patient. Good prognosis patients (i.e. low severity and complexity) were allocated to the brief CBT group, whereas poor prognosis patients were randomised to either standard or intensive CBT. Treatment was carried out by experienced clinicians who were trained in CBT. CBT in all three conditions was delivered to the same treatment protocol based on well-documented standard CBT treatment for GAD.[106] In addition, all patients received the self-help treatment manual for GAD produced by Andrews and colleagues.[117] A full description of the trial method can be found in the original report.[116] In the original trial poor prognosis patients receiving intensive CBT did not fare any better than those receiving standard CBT, whereas the good prognosis patients fared best overall, despite receiving only brief therapy.

Trial 8 – panic disorder, Fife[14]

Trial 8 was an RCT of different intensities of CBT for panic disorder, with one treatment group also receiving computer augmentation. Patients were

referred by GPs and all had a diagnosis of panic disorder according to DSM-IV[114] criteria. The trial was an international multi-centre trial, of which the Scottish patients were followed up for the current study. This arm of the original trial was run by the Department of Clinical Psychology, NHS Fife and the University of St. Andrews. It was conducted between 1997 and 2000. The three treatment groups in Trial 8, all with varying degrees of therapist contact, involved 12 1-hour sessions of CBT (CBT12), six 1-hour sessions of CBT (CBT6) or six 1-hour sessions of CBT supplemented by computer-assisted instruction (CBT6-CA). There was also a delayed (i.e. WL) condition. Treatment was carried out by chartered clinical psychologists with extensive experience of CBT. The CBT approach was based on panic control treatment[118] which incorporated the cognitive and behavioural theories of panic[118,119] and standard cognitive and behavioural techniques. All active treatment conditions had the same content and the same supplementary handouts. In the CBT6-CA condition, patients carried a palmtop computer for a further 6 weeks after therapist contact. The computer was programmed to signal the subject five times daily to prompt the practice of the therapy components. A full description of the trial method can be found in the original report.[14] In the original trial, results showed that at post-treatment CBT12 was more effective than CBT6 but not different from CBT6-CA, but also that CBT6 did not differ from CBT6-CA. The active treatments did not differ statistically at 6-month follow-up of the original trial.

Trial 9 – schizophrenia, schizo-affective disorder or delusional disorder, Tayside and Fife[66]

Trial 9 was a collaborative study, carried out by two adjacent mental health services in Tayside and Fife. The original trial was conducted between 1997 and 2001. The three treatment groups in Trial 9 were CBT (plus TAU), SP (plus TAU) and TAU. The treatment period was 9 months and both CBT and SP patients received up to 20 half-hour therapy sessions over this time. CBT was carried out by clinical nurse specialists who were registered as cognitive behaviour therapists with the British Association of Behavioural and Cognitive Psychotherapy. The treatment protocol for CBT was based on best practice, as exemplified by the treatment manuals of Tarrier[120] and Kingdon and Turkington.[121] SP was carried out by mental health professions with no training in CBT. A full description of the trial method can be found in the original report.[66] In the original trial patients receiving CBT had shown greater improvement in overall symptom severity than those receiving either SP or TAU, and the CBT and SP groups had shown greater improvement in severity of delusions over the TAU group.

Trial 10 – schizophrenia or related disorder, West of Scotland[67]

Trial 10 was run jointly by the NHS trusts in Glasgow and Ayrshire and Arran, and was conducted between 1997 and 2000. Patients fulfilled DSM-IV criteria for schizophrenia or a related disorder, were on antipsychotic medicine and were considered relapse prone. Patients were randomised to either CBT plus TAU (CBT) or TAU. CBT treatment was carried out by one clinical psychologist, according to a treatment protocol detailed in Gumley and Power.[122] CBT was divided into two phases, an engagement phase which consisted of five sessions over 12 weeks, and a targeted phase delivered at the appearance of early signs of relapse. A full description of the trial method can be found in the original report.[67] In the original trial, CBT was associated with significant reductions in relapse rate, admission rate and duration of relapse over a 12-month period in comparison with TAU.

Treatment groups

Anxiety disorder studies (Trials 1–8)

The original treatment conditions were allocated to combined treatment groups as follows (see *Table 7*).

Non-CBT (*n* = 238)

None of the patients in this group received any CBT treatment during the original trial. Some patients received alternative therapies, and all patients received some therapist contact. The mean number of treatment sessions in this group was 5.8 (4.2) with an average total contact time of 4.8 (SD = 4.0) therapist hours (including time spent on assessments). The non-CBT group includes the DZ and PL groups from Trial 1 (which was balanced for therapist contact), the AP treatment group from Trial 2 (delivered at both low and high frequency), the FL and PL groups from Trial 3 (which was balanced for therapist contact), the EMDR group from Trial 6 and the WL only conditions from Trial 6 (those in the WL condition who did not go on to enter trial treatment) and from Trial 8 (i.e. the delayed group, who did not go on to enter trial treatment).

With the exception of the WL conditions from Trials 6 and 8, all of these groups had therapist

TABLE 7 *Original trial treatment groups by combined CBT group*

	Trial 1	**Trial 2**	**Trial 3**	**Trial 4**	**Trial 5**	**Trial 6**	**Trial 7**	**Trial 8**	**Trial 9**	**Trial 10**
Average time to long-term follow-up (years)	12.9 (0.7)	9.2 (0.8)	8.9 (0.8)	5.6 (0.4)	4.7 (0.6)	3.8 (0.5)	3.8 (0.6)	3.9 (0.6)	5.0 (0.7)	5.3 (0.7)
Original diagnosis	GAD DSM-III	GAD DSM-III-R	Panic disorder DSM-III-R	Panic disorder DSM-III-R	Panic disorder DSM-III-R	PTSD DSM-IV	GAD DSM-IV	Panic disorder DSM-IV	Schizophrenia and related DSM-IV	Schizophrenia and related DSM-IV
Non-CBT	DZ PL	AP	FL PL			EMDR WL only		Delayed CBT	TAU SP	TAU
CBT: low contact				Minimum contact CBT Bibliotherapy	Group CBT		Brief CBT			
CBT: standard contact	CBT CBT+DZ CBT+PL	AMT Low frequency CT	CBT CBT+FL CBT+PL	Standard contact CBT	Individual CBT	E+CR	Standard CBT	CBT6 CBT6-CA	CBT	CBT
CBT: high contact		High frequency CT					Intensive CBT	CBT12		

AMT, anxiety management training; AP, analytical psychotherapy; CBT, cognitive behaviour therapy; CBT6, CBT over 6 sessions; CBT6-CA, CBT over 6 sessions with computer assistance; CBT12, CBT over 12 sessions; CT, cognitive therapy; DZ, diazepam; E+CR, exposure plus cognitive restructuring; EMDR, eye movement desensitisation and reprocessing; FL, fluvoxamine; PL, placebo; SP, supportive psychotherapy; TAU, treatment as usual; WL, waiting list.
With the exception of the WL only condition in Trial 6 and the delayed condition in Trial 8, all of the non-CBT treatment conditions had a degree of therapist contact during the trial in addition to any contact during assessment sessions.

contact over and above the assessment sessions. Although the two WL conditions did not receive any treatment as part of the original trial, they were still included in the analysis as, because of the length of time to follow-up, the majority of them had received some form of additional treatment in the interim period.

CBT group (*n* = 620)

All of the patients in this group received CBT treatment according to standard protocol for the anxiety disorder (i.e. GAD, panic disorder or PTSD) being treated during the original trial. The mean number of treatment sessions in this group was 7.0 (3.7) with an average total contact time of 5.6 [standard deviation (SD) = 3.9] therapist hours (including time spent on assessments). The CBT group includes the CBT, CBT+PL, and CBT+drug (i.e. FL or DZ) conditions from Trials 1 and 3, the low and high frequency CT groups and the AMT group from Trial 2, the standard, minimum and bibliotherapy groups from Trial 4, the group CBT group and those who went on to have individual CBT following the WL condition from Trial 5, the E+CR group from Trial 6, the brief, standard and intensive CBT groups from Trial 7 and the CBT6, CBT6-CA and CBT12 groups from Trial 8.

In addition, in the analysis in Chapter 8, the CBT group has been broken down by intensity of treatment as follows.

Low contact CBT

All treatment conditions included in this group had delivered CBT to standard protocol over fewer than six sessions (mean 4.3, SD = 1.9) with an average total contact time of 2.6 (SD = 1.9) therapist hours (including time spent on assessments). The low contact CBT group includes the minimum and bibliotherapy treatment groups from Trial 4 and the group CBT treatment group from Trial 5. (Note: the brief CBT group from Trial 7 has been excluded from any analysis on intensity of CBT as patients in this group were not randomised but were allocated to treatment on the basis of good prognosis.)

Standard CBT

All treatments in this condition followed a standard protocol for CBT for the relevant diagnosis (i.e. GAD, panic disorder or PTSD). Treatment was delivered over an average of 8.0 (SD = 1.4) sessions, with an average total therapist contact time of 6.6 (SD = 1.5) hours including assessments. This standard CBT group includes the CBT, CBT+PL, and CBT+drug (i.e. FL or DZ) conditions from Trials 1 and 3, the low intensity CT and AMT groups from Trial 2, the standard contact CBT group from Trial 4, those who went on to have individual CBT following the WL condition from Trial 5, the E+CR group from Trial 6, the standard CBT group from Trial 7 and the CBT6 and CBT6-A groups from Trial 8.

High contact CBT

All treatments in this condition followed a standard protocol for CBT for the relevant diagnosis (i.e. GAD or panic disorder), which was conducted over an extended time period. Treatment was delivered over an average of 13.7 (SD = 4.6) sessions, with an average total therapist contact time of 13.6 (SD = 4.7) hours including assessments. This high contact CBT group includes the high frequency CT group from Trial 2, the intensive CBT group from Trial 7 and the CBT12 group from Trial 8.

Psychosis studies (Trials 9 and 10)

The CBT condition in each of the two psychosis studies followed the same standard protocol as for schizophrenia and related disorders, and included TAU. Other treatment conditions were TAU (alone) (in both studies), and SP+TAU (SP) (in Trial 9 only). For the purposes of the analysis, the treatment groups were amalgamated across the two studies, divided into standard CBT (n = 44) (i.e. the two CBT groups) and non-CBT (n = 49) [i.e. the SP group from Trial 9 and the TAU (alone) groups from both studies].

Chapter 4

Participants

Summary of long-term follow-up participants

Anxiety disorder studies (Trials 1–8)

Trial 1 – GAD, Stirling[24]

Long-term follow-up for Trial 1 was between 11 and 14 years after entry to the original trial. Of the 111 patients entered into Trial 1, 10 had failed to attend for treatment, 18 had dropped out of treatment before the mid-point and 83 had completed treatment. Seventeen of these 111 patients were either untraceable at long-term follow-up or had died in the interim period, and one GP refused permission to contact a patient. Of the 93 patients who were contacted and invited to participate in the long-term follow-up study, 33 (30%) took part, with 28 of these attending for full diagnostic interview. A flow chart of recruitment from entry into the original trial to long-term follow-up is shown in Appendix 4.

Trial 2 – GAD, Dundee[107]

Long-term follow-up for Trial 2 was between 8 and 11 years after entry to the original trial. There were 110 participants entered into the original trial, of whom 11 had failed to attend for treatment, 19 had dropped out of treatment and 80 had completed treatment. Sixteen of these 110 patients were either untraceable at long-term follow-up or had died in the interim period, and one GP refused permission to contact a patient. Of the 93 patients who were invited to participate, 61 (66%) took part in the long-term follow-up study, with 51 of these attending for full diagnostic interview. A flow chart of recruitment from entry into the original trial to long-term follow-up is shown in Appendix 4.

Trial 3 – panic disorder, Stirling[109]

Long-term follow-up for Trial 3 was between 7 and 11 years after entry to the original trial. Of the 190 patients entered into Trial 3, 41 had dropped out of treatment before the mid-point and 149 had completed treatment. Nineteen of these 190 patients were either untraceable at long-term follow-up or had died in the interim period, and GPs refused permission to contact three patients. Of the 168 patients who were contacted, 90 (54%) took part in the long-term follow-up study, with 75 of these attending for full diagnostic interview. A flow chart of recruitment from entry into the original trial to long-term follow-up is shown in Appendix 4.

Trial 4 – panic disorder, Stirling[111]

Long-term follow-up for Trial 4 was between 5 and 7 years after entry to the original trial. Of the 104 patients entered into Trial 4, 13 had dropped out of treatment before the mid-point and 91 had completed treatment. Nine of these 104 patients were untraceable at long-term follow-up or had died in the interim period, and GPs refused permission to contact two patients. Of the 93 patients who were contacted, 50 (54%) took part in the long-term follow-up study, with 40 of these attending for full diagnostic interview. A flow chart of recruitment from entry into the original trial to long-term follow-up is shown in Appendix 4.

Trial 5 – panic disorder, Stirling[113]

Long-term follow-up for Trial 5 was between 4 and 6 years after entry to the original trial. Of the 65 additional patients entered into Trial 5, five failed to attend for treatment, 21 had dropped out of treatment before the mid-point and 39 had completed treatment. One of these 65 patients was untraceable at long-term follow-up, and GPs refused permission to contact three patients. Of the 61 patients who were therefore contacted, 38 (62%) took part in the long-term follow-up study, with 32 of these attending for full diagnostic interview. Although those patients allocated to the WL condition in Trial 5, who then chose individual CBT, were not randomly allocated to treatment and hence excluded from the original analysis, they have been included at long-term follow-up. This was felt to be an acceptable approach for the current analysis, as the majority of patients in Trials 1–10 inclusive had received some form of additional treatment (beyond the original treatment trial) in the interim period to long-term follow-up. A flow chart of recruitment from entry into the original trial to long-term follow-up is shown in Appendix 4.

Trial 6 – post-traumatic stress disorder, Stirling[50]

Long-term follow-up for Trial 6 was between 3 and 4 years after entry to the original trial. Of the 89 patients entered into Trial 6, 33 had dropped out of treatment before the mid-point and 56 had

completed treatment. Two of these 89 patients were untraceable at long-term follow-up and GPs refused permission to contact three of the remaining patients. Of the 84 patients who were contacted, 34 (40%) took part in the long-term follow-up study, with 31 of these attending for full diagnostic interview. The low response rate at long-term follow-up for Trial 6 was felt to be due to the nature of PTSD, with participants (both recovered and non-recovered) finding it distressing to talk about their traumatic experiences. As no treatment was being offered at long-term follow-up, it was therefore not surprising that many chose not to take part (three patients actually indicated their refusal for this reason). A flow chart of recruitment from entry into the original trial to long-term follow-up is shown in Appendix 4.

Trial 7 – GAD, Dundee, Fife and Stirling[116]

Long-term follow-up for Trial 7 was between 2.5 and 5 years after entry to the original trial. Of the 97 patients entered into Trial 7, 34 had either failed to attend treatment or dropped out of treatment before the mid-point and 63 had completed treatment. Five of these 97 patients were untraceable at long-term follow-up or had died in the interim period. Of the 92 patients who were contacted, 45 (49%) took part in the long-term follow-up study, with 41 of these attending for full diagnostic interview. A flow chart of recruitment from entry into the original trial to long-term follow-up is shown in Appendix 4.

Trial 8 – panic disorder, Fife[14]

Long-term follow-up for Trial 8 was between 2 and 5 years after entry to the original trial. Of the 95 patients entered into Trial 8, 15 had failed to start treatment or dropped out before the mid-point and 80 were defined as having completed treatment. Three of these 95 patients were untraceable at long-term follow-up. Of the 92 patients who were contacted, 45 (49%) took part in the long-term follow-up study, with 39 of these attending for full diagnostic interview. A flow chart of recruitment from entry into the original trial to long-term follow-up is shown in Appendix 4.

Psychosis studies (Trials 9 and10)

Trial 9 – schizophrenia, schizo-affective disorder or delusional disorder, Tayside and Fife[66]

Long-term follow-up for Trial 9 was between 3 and 6 years after entry to the original trial. Of the 66 patients entered into Trial 9, seven had dropped out of treatment, one had died of natural causes during treatment and 58 had completed treatment. Eight of these patients were untraceable at long-term follow-up or had died in the interim period, and consultant permission was refused for one individual. Of the 57 patients who were contacted, 30 (53%) took part in the long-term follow-up study, with all of these attending for full diagnostic interview. Case note data were collected on eight patients who did not attend interview. A flow chart of recruitment from entry into the original trial to long-term follow-up is shown in Appendix 4.

Trial 10 – schizophrenia or related disorder, West of Scotland[67]

Long-term follow-up for Trial 10 was between 3 and 6 years after entry to the original trial. Of the 144 patients entered into Trial 10, 11 had withdrawn from treatment and 133 had completed treatment. Ten of these 144 patients were untraceable at long-term follow-up or had died in the interim period, and consultants refused permission to contact another 21 patients. Of the 113 patients who were contacted, 63 (56%) took part in the long-term follow-up study, with all of these attending for full diagnostic interview. A flow chart of recruitment from entry into the original trial to long-term follow-up is shown in Appendix 4.

Summary of participants at long-term follow-up

Table 8 summarises the above information. A small number of patients (n = 7) in the studies based at the Anxiety and Stress Research Centre in Stirling took part in more than one of the original treatment trials, as they had not recovered following previous treatment. Numbers of those who had already participated in an earlier study are also displayed in *Table 8*.

Where the trials are combined in the current analysis, duplicated patients are only included once, that is, with their data from the first of the treatment trials (chronologically) in which they participated. However, where treatment trials are examined individually in the current analysis, and hence would not be duplicated, all of the participants from the relevant treatment trials have been included. This accounts for the apparent discrepancies in numbers of participants reported in different sections of this report.

Representativeness of participants in relation to population demographics

Anxiety disorder studies (Trials 1–8)

Demographic data on a number of variables [i.e. age, gender, marital status, employment,

TABLE 8 *Trial and long-term follow-up participants by original treatment trial*

	Original trial number										
Participants	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Total**
In original trial											
Entered	111	110	190	104	65	89	97	95	66	144	1071
Completed	83	80	149	91	39	56	62	80	58	133	831
Dropped out/FTA	28	30	41	13	26	33	35	15	8	11	240
At long-term follow-up											
Unavailable to contact[a]	18	17	22	11	4	5	5	3	9	31	125
Took part	33	61	90	50	38	34	45	45	30	63	489
Duplicated in previous studies	0	0	4	2	1	0	0	0	0	0	7

FTA, failed to attend.
[a] These patients were either untraced, had died in the interim period or their medical practitioner had refused permission for contact.

education, socio-economic status and deprivation category[123] (see section 'Additional data collected', p. 31)] were collected both pretreatment and at long-term follow-up in each of the trials. A summary of these data at long-term follow-up for the anxiety disorder studies is given in *Tables 9* and *10*. Also shown in *Tables 9* and *10* are the nearest equivalent data, based on adults aged 16–74 years, from the 2001 census statistics for Scotland.[124]

It can be seen from *Table 9* that, in comparison with the census data, a higher proportion of the clinical sample is female (61 versus 52%), married (61 versus 52%) and not working owing to ill-health (26 versus 7%). Similarly, fewer of the clinical sample are single (12 versus 26%) and in gainful employment (47 versus 58%). *Table 10* shows that smaller proportions of the clinical sample are from the professional socio-economic group (6 versus 8%), and that fewer are educated to degree level (9 versus 20%).

Psychosis studies (Trials 9 and 10)

Demographic data at long-term follow-up for patients in the psychosis studies are shown in *Table 11*.

It can be seen from *Table 11* that a larger percentage of the psychosis patients were male (67%), and they were also more likely to be single (66%) and less likely to be married or cohabiting (18 and 3%) as compared with the census data (49, 26, 52 and 9%, respectively).

Power of study to achieve aims

The power of a significance test is the probability of obtaining a significant result at some level of significance when a particular effect size is present. In the present study the primary aim was to see whether those patients who had received CBT were any better at long-term follow-up than those patients who had not received CBT. In order to obtain a measure of the power of the study to test this comparison, it is necessary to specify the sample size of the CBT and non-CBT groups.

Anxiety disorder studies (Trials 1–8)

Determining the sample size of the CBT and non-CBT groups was difficult for the anxiety studies as the numbers on each long-term follow-up variable differ depending on circumstances such as whether a patient attended diagnostic interview. The following power analysis assumes the sample sizes are based on the number of patients having a score on the composite long-term outcome factor [see the section 'Composite measures', p. 29] which equates to having scores on measures of clinical global severity and improvement (both recorded at interview), and also on the main outcome measures used across all eight studies. On this criterion there are 81 non-CBT scores and 230 CBT scores. For these numbers, the powers of a two-sample *t*-test to detect what Cohen[125] defined as large, medium and small effects (0.8, 0.5 and 0.2 of a standard deviation) using a significance level of 0.05 is virtually 1, 0.971 and 0.338, respectively. Hence a result which did not

TABLE 9 *Summary of long-term follow-up data for Trials 1–8 by original trial: age, gender, marital status and employment*

Maximum *n*	**Trial 1 33**	**Trial 2 61**	**Trial 3 90**	**Trial 4 50**	**Trial 5 38**	**Trial 6 34**	**Trial 7 45**	**Trial 8 45**	**Overall 396**	**2001 census[a] 3,731,079**
Age at long-term follow-up: mean (SD) (years)	52.67 (11.05)	48.18 (11.32)	44.24 (11.18)	46.12 (12.56)	46.84 (12.20)	44.76 (11.73)	44.64 (10.40)	42.71 (9.91)	45.84 (11.49)	–
Gender (%)										
Male	42.4	34.4	27.8	30.0	44.7	64.7	57.8	26.7	38.6	48.5
Female	57.6	65.6	72.2	70.0	55.3	35.3	42.2	73.3	61.4	51.5
Marital status (%)										
Married	67.7	61.1	64.9	65.1	70.6	58.1	47.6	57.5	61.1	51.8
Cohabiting	9.7	5.6	10.4	4.7	5.9	25.8	7.1	10.0	9.5	8.8
Single	3.2	11.1	13.0	9.3	5.9	9.7	26.2	12.5	12.1	26.0
Separated	6.5	1.9	1.3	0.0	2.9	3.2	7.1	2.5	2.9	3.2
Divorced	3.2	11.1	6.5	14.0	8.8	3.2	11.9	15.0	9.5	5.6
Widowed	9.7	9.3	3.9	7.0	5.9	0.0	0.0	2.5	4.5	4.6
Employment status (%)										
Working	43.8	41.5	51.3	44.2	61.1	38.7	50.0	48.7	47.6	58.0
Unemployed	6.3	5.7	10.3	9.3	5.6	16.1	11.9	17.9	10.3	4.0
Retired (age)	28.1	9.4	7.7	11.6	11.1	3.2	4.8	0.0	9.2	13.9
Not working – health[b]	18.8	35.8	21.8	32.6	19.4	35.5	28.6	15.4	25.8	7.4
Housewife	3.1	7.5	3.8	2.3	0.0	6.5	2.4	10.3	4.6	5.5
Student	0.0	0.0	5.1	0.0	2.8	0.0	2.4	7.7	2.6	7.3

[a] 2001 census data:[124] based on available data for adults aged between 16 and 74 years (inclusive) living in Scotland.
[b] On incapacity or disability benefit or taken early retirement on health grounds.

TABLE 10 *Summary of long-term follow-up data for Trials 1–8 by original trial: deprivation category, socio-economic status and education*

Maximum *n*	**Trial 1 33**	**Trial 2 61**	**Trial 3 90**	**Trial 4 50**	**Trial 5 38**	**Trial 6 34**	**Trial 7 45**	**Trial 8 45**	**Overall 396**	**2001 census[a] 3,731,079**
Deprivation category[b] (%)										
1	12.5	19.7	1.1	0.0	0.0	0.0	4.4	0.0	5.0	
2	9.4	6.6	3.4	22.4	31.6	9.1	33.3	2.4	13.4	
3	28.1	14.8	18.0	30.6	26.3	9.1	6.7	9.8	17.8	
4	40.6	11.5	46.1	26.5	28.9	54.5	33.3	73.2	37.8	
5	3.1	8.2	27.0	16.3	13.2	24.5	6.7	9.8	15.2	
6	6.3	39.3	4.5	4.1	0.0	3.0	15.6	4.9	10.8	
Socio-economic status (%)										
Professional	14.3	3.9	3.0	5.0	0.0	13.8	9.8	0.0	5.6	8.4
Managerial/technical	32.1	19.6	17.9	30.0	31.3	3.4	26.8	12.8	21.3	21.4
Skilled	28.6	29.4	38.8	30.0	34.4	41.4	41.5	33.3	35.2	27.8
Part-skilled	7.1	21.6	20.9	25.0	6.3	20.7	2.4	17.9	15.7	15.5
Unskilled/ unemployed	17.9	25.5	17.9	10.0	25.0	20.7	19.5	35.9	21.6	18.0
Student	0.0	0.0	1.5	0.0	3.1	0.0	0.0	0.0	0.6	8.8
Education level (%)										
Degree or higher	20.0	5.8	5.4	10.8	10.0	6.9	22.0	0.0	9.1	19.5
HND or equivalent	20.0	7.7	4.1	10.8	10.0	10.3	9.8	17.9	10.3	7.0
A levels/highers	0.0	13.5	20.3	10.8	16.7	27.9	17.1	23.1	16.7	16.7
O, std grades, SVQ 1 or 2	26.6	23.0	29.7	24.4	23.3	34.5	32.0	28.2	27.7	24.7
None	33.3	50.0	40.5	43.2	40.0	20.7	19.5	30.8	36.1	33.2

[a] 2001 census data:[124] based on available data for adults aged between 16 and 74 years (inclusive) living in Scotland.
[b] There were no long-term follow-up participants in deprivation category 7. Comparative data on deprivation category from the 2001 census were not available at the time of report.

TABLE 11 *Summary of long-term follow-up data for Trials 9 and 10 by original trial: age, gender, marital status and deprivation category*

Maximum *n*	**Trial 9 30**	**Trial 10 63**	**Overall 93**	**2001 census[a] 3,731,079**
Age at long-term follow-up: mean (SD) (year)	41.67 (9.85)	40.60 (11.13)	41.05 (10.55)	
Gender (%)				
Male	53.3	73.0	66.7	48.5
Female	46.7	27.0	33.3	51.5
Marital status (%)				
Married	26.7	13.1	17.6	51.8
Cohabiting	0.0	4.9	3.3	8.8
Single	53.3	72.1	65.9	26.0
Separated	10.0	1.6	4.4	3.2
Divorced	10.0	8.2	8.8	5.6
Widowed	0.0	0.0	0.0	4.6
Deprivation category[b] (%)				
1	3.3	0.7	1.0	
2	10.0	7.4	7.7	
3	20.0	8.1	11.9	
4	26.7	27.9	28.4	
5	20.0	23.5	22.7	
6	20.0	25.0	23.2	
7	0.0	7.4	5.2	

[a] 2001 census data:[124] based on available data for adults aged between 16 and 74 years (inclusive) living in Scotland.
[b] Comparative data on deprivation category from the 2001 census were not available at the time of report.

reach the 0.05 level of significance is inconsistent with large and medium CBT effects but it does not rule out the possibility that there is a small undetected effect of CBT on the long-term outcome measures.

Psychosis studies (Trials 9 and 10)

Determining the sample size of the CBT and non-CBT groups was easier for the psychosis studies as they had used consistent measures at long-term follow-up. The following power analysis is based on 49 non-CBT scores and 44 CBT scores.

The power of a *t*-test between CBT and non-CBT conditions with these numbers (total $n = 91$) to detect large (0.8), medium (0.5) and small (0.2) effect sizes, using a significance level of 0.05 (two-tailed), is 0.967, 0.663 and 0.156, respectively. Hence the data set for the psychosis studies would give 80% power of detecting an effect size of 0.588 with a two-tailed significance of 0.05. The estimated effect size of the difference between the mean change scores from pretreatment to short-term follow-up for the CBT and non-CBT conditions was 0.50 for Trial 9 and 0.64 for Trial 10. The psychosis data set therefore has a reasonable chance of detecting similar CBT versus non-CBT treatment effects if they were present at long-term follow-up.

Chapter 5

Measures and missing data considerations

Main outcome measures

A wide range of outcome measures were used in each of the original trials. In order to handle better the data at long-term follow-up, two merged data files were constructed, one for the eight anxiety studies and one for the two psychosis studies. These two data files include the main outcome measures used within each study, plus any measures which are common across studies. The measures used fall into two categories: clinician-rated measures and patient-rated (self-report) scales. Alphabetical listings of the measures in each category are given below. In addition, the main outcome measures used at long-term follow-up within each study are shown in *Table 12*.

Clinician-rated measures

Anxiety Disorders Interview Schedule: DSM-IV version[98] (ADIS-IV)

The ADIS was developed by Di Nardo and colleagues to assist the reliable diagnosis of anxiety disorders. It records clinical diagnoses for all of the anxiety disorders and selected mood disorders (depression and dysthymia) and also acts as a screening tool for bipolar and psychotic disorders. In a large reliability study[126] the DSM-III-R version of the ADIS[127] showed moderate to excellent reliability for the anxiety disorders, whether given as a primary or secondary diagnosis, and excellent reliability for major depression. Dysthymia showed good reliability only when diagnosed as a secondary disorder. In the present study, only diagnoses for current episode were used, and no distinction was made between primary and secondary diagnoses.

Clinical Global Severity[128] (CGS)

CGS was measured at long-term follow-up on a nine-point scale (0–8), where 0 = no evidence of disorder, 4 = definitely disturbing/disabling and 8 = very severely disturbing/disabling. The same convention as that used by Brown and Barlow[129] was adopted in the current study, that is, a CGS

TABLE 12 *Main outcome measures used at long-term follow-up by original trial across all studies*

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8	Trial 9	Trial 10
Clinician-rated measures										
ADIS-IV	✓	✓[a,b]	✓	✓	✓	✓	✓[a]	✓		
CGS (0–8)	✓[a,b]	✓	✓[a,c]	✓[a,c]	✓[a,c]	✓[a,d]	✓[a]	✓[a]	✓	✓
HAM-A	✓[a]	[a]	✓[a]	✓[a]	✓[a]	✓[a]	✓[a]	✓[a]		
PANSS									✓[a]	✓[a]
Patient-rated measures										
BSI	✓	✓[a]	✓	✓	✓	✓	✓[a]	✓	✓[a]	✓[a]
SF-36	✓	✓	✓	✓	✓	✓	✓	✓[a]	✓	✓
CGI (1–7)	✓[a]	✓[a,e]	✓[a]	✓[a]	✓[a]	✓[a,c]	✓[a]	✓[a]		
SRT	✓[a]		✓[a]	✓[a]	✓[a]					
FQ-Agora			✓[a]	✓[a]	✓[a]			✓[a]		
PANAS	✓	✓	✓	✓	✓	✓[a]	✓	✓		
STAI-T		✓[a]	✓	✓	✓	✓	✓[a]	✓[a]		
PBIQ									✓[a]	✓[a]

ADIS-IV, Anxiety Disorders Interview Schedule for DSM-IV; BSI, Brief Symptom Inventory; CGI, Clinical Global Improvement; CGS, Clinical Global Severity; FQ-Agora, Fear Questionnaire, agoraphobia subscale; HAM-A, Hamilton Anxiety Scale; MADRS, Montgomery Asberg Depression Rating Scale; PANAS, Positive and Negative Affect Scale; PANSS, The Positive and Negative Syndrome Scale; PBIQ, Personal Beliefs about Illness Questionnaire; SF-36, SF-36 Health Survey Scale; SRT, Kellner and Sheffield Symptom Rating Test; STAI-T, Test–Trait Anxiety Inventory – trait version.

[a] Also used pretreatment.
[b] Pretreatment DSM version III-R.
[c] Pretreatment scale = 1–7.
[d] Pretreatment scale = 0-4.
[e] Pretreatment scale = 0–5.

score of 4 or above equates to having at least one ADIS-IV diagnosis. It should be noted that the CGS used in the original trial for Trials 1, 3, 4 and 5 was measured on a seven-point scale (1–7) where 5 represented a definite degree of impairment.

The Hamilton Anxiety Rating Scale[130] (HAM-A)

The HAM-A is completed by a clinician and based on information elicited from the patient during an interview. This scale incorporates 14 anxiety-related symptoms each rated 0–4. The Hamilton Anxiety Glossary[131] was used for scoring in the majority of studies pretreatment and in all studies at long-term follow-up, as it enables a more reliable and valid assessment based on the frequency, duration and severity of the components of the HAM-A.

The Montgomery Asberg Depression Rating Scale[132] (MADRS)

The MADRS is completed by a clinician and based on information elicited from the patient during an interview. It assesses 10 symptoms of depression each rated on a 0–6 scale. Items are then totalled to provide an overall MADRS score.

The Positive and Negative Syndrome Scale[81] (PANSS)

The PANSS is a 30-item clinician-rated scale measuring positive and negative symptoms of schizophrenia and related disorders. Each item is rated on a severity scale ranging from 1 (absence of psychopathology) to 7 (extremely severe). The sum of the first seven items consisting of delusions, conceptual disorganisation, hallucinatory behaviour, excitement, grandiosity, suspiciousness/persecution and hostility are totalled to obtain the positive scale score (range 7–49). Items 8–14 (e.g. blunted affect, emotional withdrawal) are totalled to obtain the negative scale score, and items 15–30 are summed to obtain the global psychopathology score (e.g. somatic concern, anxiety).

The Social Adjustment Scale (SAS)[133]

This clinician-rated scale has four subscales related to adjustment in the following areas: family relationships, social relationships, marital relationships and leisure and work. Only pretreatment scores were used in the current analysis.

The Social Functioning Scale (SFS)[134]

This clinician-rated scale has seven subscales related to functioning in the following areas: social engagement/withdrawal, interpersonal communication, independence – performance, independence – competence, recreation, prosocial behaviour and occupation/employment. Raw scores are converted to scaled scores calculated in relation to normative data for unemployed schizophrenics (mean = 100, SD = 10).

Patient-rated measures

The Beck Depression Inventory[56] (BDI)

The BDI is a self-report measure of severity of depression, consisting of 21 items each scored on a four-point scale.

The Brief Symptom Inventory[135] (BSI)

This is a shortened version of the SCL-90[136] and is a global measure of current symptomatic state, consisting of 53 items each measured on a five-point scale (range 0–4). The BSI has a number of subscales, namely anxiety, depression, phobic symptoms, interpersonal sensitivity, hostility, obsessive–compulsive, somatisation, paranoid ideation and psychoticism, plus an overall global severity index (BSI-GSI), which represents a mean score for all completed items.

Clinical Global Improvement[128] (CGI)

This is a patient-rated seven-point scale (range 1–7) measuring change in condition since the beginning of the original trial using the following seven categories: marked, moderate or mild improvement, no change, mild, moderate or marked deterioration.

The Fear Questionnaire[137] (FQ)

This self-rated scale provides a rating of agoraphobic, social and blood-injury avoidance and also ratings of mood disruption and global phobic distress. The scale consists of 21 items scored on a scale of 0–8. The agoraphobia subscale (FQ-Agora), with possible scores ranging from 0 to 40, was the main measure used in the current research.

The Hospital Anxiety and Depression Scale[57] (HADS)

This 14-item measure assesses the presence and frequency of symptoms related to anxiety and depressive symptoms each on a four-point scale. Items are totalled to provide two subscale scores for anxiety and depression (both seven items).

The Kellner and Sheffield Symptom Rating Test[138] (SRT)

This is a 30-item self-report questionnaire designed to measure changes in symptoms of distress in neurotic patients. Each item is scored on a four-point scale. The SRT also has four subscales; the two subscales representing anxiety

and depression, and also the total SRT score, were used in this report.

The Personal Beliefs about Illness Questionnaire[139] (PBIQ)

This self-report measure assesses key beliefs appraising self and psychosis, including perceived control or entrapment in psychotic illness; loss of autonomy and social role; humiliation both in terms of lowered rank or status and devaluation in relation to self or others; and attribution of causality of illness. It consists of 16 items, each scored on a four-point scale.

The Positive and Negative Affect Scale[140] (PANAS)

The PANAS is a 20-item scale representing two broad mood states, namely positive affect (PA) and negative affect (NA). NA reflects negative feelings and emotions such as fear, distress, hostility and shame, and PA reflects positive feelings and emotions such as interest, determination, enthusiasm and pride. There are 10 items for both PA and NA, each scored on a five-point scale (1–5). By changing the time frame of the questioning, the PANAS can be used as either a trait (i.e. how you feel generally) or state (e.g. how you feel at this moment) measure. The trait measure was used in the current study.

Self-esteem Scale[141] (SES)

This is a 10-item Likert scale measuring attitude towards oneself. Items are generally scored on a 1–4 scale ('strongly agree', 'agree', 'disagree', 'strongly disagree'). In Trial 7 a variant of the Rosenberg scale was used,[142] consisting of eight of the original 10 items scored on a seven-point scale ('strongly agree', 'agree', 'slightly agree' 'neither agree nor disagree', 'slightly disagree' 'disagree', 'strongly disagree').

The Sheehan Disability Scale[143]

This is a simple measure of social functioning. It assesses disruption to daily lifestyle and comprises three 10-point subscales on which patients self-rate disruption to work, social life and family/home life. This measure was only used at the time of the trial.

The State–Trait Anxiety Inventory[58] (STAI)

The trait version of the State–Trait Anxiety Inventory (STAI-T) was used in the current research. The STAI-T has 20 items, each scored on a scale of 1 ('almost never') to 4 ('almost always'), which relate to how patients feel generally. Items are totalled to provide an overall trait anxiety score.

The UK SF-36 version II[99]

The SF-36 Health Survey was developed by Ware and colleagues[144] as a measure of health status comprising eight subscales: physical role, physical functioning, emotional role, vitality, bodily pain, mental health, social role and general health. Jenkinson and colleagues[99] developed norms for a UK population, using a version of the SF-36 which has minor wording differences from the original measure, in order to make it more acceptable to UK populations. In addition, it is possible to calculate physical (SF-36 PC) and mental (SF-36 MC) components of the SF-36, which summarise the eight original subscales, based on factor analyses. These were also used in the current research.

Composite measures

Anxiety disorder studies (Trials 1–8)

The eight anxiety disorder studies had each used a variety of different measures of clinical severity, depression, and anxiety symptoms at the time of the original trial. Although a few measures were used consistently across up to four studies, there was only one measure (the clinician-rated HAM-A) used across all eight trials at pretreatment. However, there was at least one scale in each of the original trials which had measured anxiety symptoms (e.g. Beck Anxiety Index, Symptom Rating Test – anxiety subscale, Hospital Anxiety and Depression Scale – anxiety scale, STAI-T), at least one scale which had measured depressive symptoms (e.g. Symptom Rating Test – depression subscale, MADRS, Brief Symptom Inventory – depression subscale), plus a measure of overall clinical severity.

It was therefore possible to devise composite scales, which meant that the pre- and post-treatment severity of patients' symptoms could be included in the analysis for all eight trials. The ideal approach would have been to convert raw scores to *z*-scores as compared with population norms for all measures, which would have preserved any differences between the clinical samples, while providing a degree of consistency between the various measures used. Unfortunately, appropriate normative data were not available for the majority of measures [due to there being no normative data available, the sample size for the normal population being very small (e.g. <100), or the normative data being inappropriate (e.g. from self-report symptom severity data collected more than 30 years ago, where levels of reporting may be expected to be much lower than in current

populations], so this method proved unviable. The adopted alternative was to calculate *z*-scores, using the greatest number of available data for each measure, and then create the composite scales for pre- and post-treatment symptoms from these *z*-scores. In general, the measures which had the maximum *n* across all eight studies were those chosen to be included in the composite scales. The following composite scales were constructed.

Composite Anxiety Scale (CANX)

The *z*-scores were calculated based on all available data for the following self-report measures of anxiety symptoms:

- Symptom Rating Test – anxiety subscale
  Trials 1, 3, 4, 5
- State–Trait Anxiety Index – trait version
  Trials 2, 4 (data available for a subset of participants only), 7, 8
- Hospital Anxiety and Depression Scale (Anxiety)
  Trial 6

These values were then combined to form a Composite Anxiety Scale. In Trial 4 the Symptom Rating test scores were used in preference to the STAI-T owing to the larger number of patients completing this measure pretreatment ($n = 103$ versus 78). In addition, the STAI-T was not used post-treatment in Trial 4.

Composite Depression Scale (CDEP)

The *z*-scores were calculated based on all available data for the following measures of depression:

- Symptom Rating Test – depression subscale
  Trials 1, 3, 4, 5
- Beck Depression Inventory
  Trials 2, 8
- Brief Symptom Inventory – depression subscale
  Trials 2, 7
- Montgomery Asberg Depression Scale
  Trial 6

These values were then combined to form a Composite Depression Scale. For Trial 2, the BDI was used in preference to the BSI depression subscale owing to the larger volume of available data pretreatment ($n = 108$ versus 97).

In addition, the following combined measures were used.

Hamilton Anxiety Scale (HAM-A)

The HAM-A had been used across all eight trials pretreatment (although not post-treatment for Trial 8). Unfortunately, a different scoring mechanism had been used on the HAM-A for Trial 2, which meant that the resulting scores were noticeably inflated over the other studies, even taking into account any differences on other measures of anxiety symptoms. An adjusted measure (HAM-AD) was therefore created which forced the mean of the data from Trial 2 to be equal to the mean for the whole sample (i.e. by reducing the HAM-A scores for Trial 2 by 12.12 points for the pretreatment data and 6.39 points for the post-treatment data). The original HAM-A values were used for the remaining trials.

Clinical Global Severity (CGS)

All eight trials used some measure of overall clinical severity, although these were based on different scales (i.e. 0–4, 1–7 or 0–8). A composite measure of CGS (COMCGS) scale was created for pre- and post-treatment data, where all scores were recoded on to a 0–8 scale according to the anchor points used in the original studies (e.g. 1 = 'normal' to 0 = 'absent'; '5 = moderate/definite impairment' to '4 = moderate/definitely disturbing/disabling'; '6 = severe' to '6 = severe/markedly disabling/disturbing'; '7 = extreme' to '8 = very severe/very severely disturbing/disabling').

Social adjustment

All but one of the original studies (Trial 1) had used either the SAS or the Sheehan Disability Scale to assess social functioning in a number of domains, of which functioning in the social environment and functioning within the family were common. The *z*-scores were calculated using all available data on these two measures to produce combined scores for a composite measure of adjustment in the family domain (CFAM) and a composite measure of adjustment in the social domain (CSOC).

Long-term outcome factor (LTOF)

A principle component analysis was performed on the seven scaled long-term outcome variables (i.e. CGS, CGI, BSI-GSI, PA, NA, SF-36 PC and SF-36 MC) which had been used across all eight anxiety disorder studies. A one-component solution was obtained which explained 62% of the variance. All tests loaded highly on this component (see *Table 13*). The negative loadings of PA and the SF-36 occur because high scores on these scales are favourable whereas the opposite is the case for the other scales. The lowest loading was the physical component of the SF-36 but even that was almost 0.5. The extracted component was used as the combined quantitative long-term outcome variable (termed long-term outcome factor).

TABLE 13 *Factor loadings of outcome measures on the extracted LTOF for Trials 1–8*

Long-term follow-up measure	Loading
CGS	0.925
CGI	0.706
BSI-GSI	0.890
PA	–0.684
NA	–0.864
SF-36 PC	–0.490
SF-36 MC	–0.844

Psychosis studies (Trials 9 and 10)

As the two psychosis studies had used essentially the same measures both at pretreatment and at long-term follow-up, there was no need to generate composite measures for these two studies.

Additional data collected

In addition to basic demographic data on age, gender and marital status, a range of additional data regarding demographic information, patient views of original treatment and how the patient had fared in the interim period, including amount of interim treatment, were also collected during the study. The additional variables which were used in the current analysis are described in more detail below.

Deprivation category[123]

Deprivation category is a measure of social deprivation in Scotland based on postcode sector. The calculation of deprivation scores is based on male unemployment, low social class, car ownership and overcrowding as calculated for a number of properties in a postcode sector represented by the first part of the postcode (e.g. PH4, FK15), plus the first digit after the space (0–9). These scores have then been divided into categories. The range of categories for Scotland is from 1 to 7 where 1 is the most affluent category and 7 the most deprived. Both the raw scores (deprivation score) and the categories have been used in this report.

Employment status

Employment status at long-term follow-up was categorised as follows: working (including full- and part-time, seasonal, self-employed but not voluntary work); retired due to age; retired or not working due to ill-health (includes those on incapacity benefit and those who took early retirement on health grounds – some of the latter may have since reached normal retirement age); housewife (including never having worked, or having given up work to look after children); full-time student and unemployed (i.e. not working for any other reason).

Views of treatment

Attitudes to participation in the original treatment trials were assessed during the interview, using a semi-structured approach. As part of this assessment participants were asked to rate the helpfulness of the original treatment on a 0–8 scale (0 = 'not at all helpful', 8 = 'extremely helpful') and their expectation of coping in the future on a 0–8 scale (0 = 'not at all hopeful', 8 = 'extremely hopeful'). They were also asked to judge whether they were now 'much better', 'much better but with periods of anxiety', 'slightly better', 'worse' or 'no different' compared with their state at the time of the trial, and how this compared with their state immediately after treatment. If some improvement had been effected over this time, they were asked what they thought were the reasons for this. They were also asked how much they remembered of the treatment and to state how often they practised any of the techniques they had learnt during treatment.

Interim treatment

A four-point scale was used to categorise the amount of patient-reported interim treatment with either medication or psychological therapy: (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 responses on amount of interim treatment were cross-checked with medical case notes for 36% (n = 12) of the sample in Trial 1. In eight cases (67%) the results tallied exactly, and in the remaining four cases three had underestimated the amount of treatment received by one category and one had underestimated it by two categories, primarily owing to the patient excluding medication which they may have believed had been prescribed for problems other than anxiety.

Statistical analysis

This section summarises the statistical analysis undertaken. The remainder of this chapter contains comparisons of aggregated CBT and non-CBT on pretreatment characteristics using χ^2 and t-tests, comparisons of long-term follow-up participants with non-participants on pre- and post-treatment variables and tables showing the

patterns of missing data. Chapter 6 is primarily a description of the status of the participants at long-term follow-up. The outcomes for different disorders are compared.

The aim of the analyses in Chapter 7 was to assess whether the outcomes were better for the patients receiving CBT as opposed to other treatments. One problem with this was that the original trials had used different outcome variables. It was therefore not possible to assess how the patients had changed on the same variable for every trial. As a result, the basic analyses consisted of comparing the scores of the CBT and non-CBT groups on common long-term follow-up outcome measures using χ^2 and *t*-tests (see *Tables 24* and *26*) and combining the *t*-tests using a MANOVA. The power of this test was maximised by combining the data from all the available trials including those that had only used different CBT conditions. Having found statistically significant differences on at least some of the quantitative measures, attempts were made to see if these differences could be explained away because of difficulties associated with missing data and pooling data from different trials.

The pooled comparisons treat all the patients assigned to the CBT and non-CBT treatments as if they were a random sample of patients assigned to these conditions. This will not be the case if the missing data occur non-randomly or if there are differences between the trials which confound differences between the CBT and non-CBT groups, e.g. trials which only compared different CBT treatments might differ from those which also included non-CBT treatments. The missing data problem was addressed by estimating the means of the outcome variables based on the predictor variables specified in *Table 40* using the direct maximum likelihood (DML) rather than the Expectation Maximisation (EM) method (used in Chapter 9) because it provided a test of the differences between the CBT and non-CBT groups. The problem of pooling was addressed by breaking the data down by trial and producing forest plots as used in meta-analyses. Meta-analyses also made it possible to perform analyses which combined changes on different measures between pre- and long-term follow-up variables and so made it possible to test differences between CBT and non-CBT groups on change scores.

In Chapter 8 the aim was to compare different intensities of CBT. Two analyses were performed: the first was restricted to data from Trials 2 and 7, which had been designed to investigate differences between standard and high intensity CBT, and the second involved Trials 4 and 5, which used standard and low intensity CBT groups. Again the analyses consisted of comparing the scores of the different intensity CBT groups on common long-term follow-up outcome measures using χ^2 and *t*-tests to test for differences and combining the *t*-tests using MANOVAs. As there were no clear indications of intensity effects present and the numbers were relatively small, it was not considered worthwhile to perform missing data analyses or Forest plots.

In Chapter 9 the aim was to predict long-term outcome from measures available both before and immediately after treatment. Here the basic analysis involved correlating the latter variables with the composite outcome measure and binary measures of clinical status at long-term follow-up. The effects of missing data on these correlations were assessed by recalculating them once the missing data had been imputed using the EM method. The DML method produced virtually the same results. In addition regression analyses were used to predict the outcome variables from all the predictor variables and interesting subsets of them.

Representativeness of participants in relation to original cohort

In view of the large number of patients who participated in the original clinical trials but who were unavailable to follow-up, or who failed to respond to an invitation to take part, an exhaustive analysis was carried out of the characteristics of participants and non-participants in the long-term follow-up study in respect of baseline demographics, allocation to treatment conditions, baseline symptomatic state and response to treatment.

The GAD studies (Trials 1, 2 and 7)

There were no significant differences between long-term follow-up participants and non-participants with regard to age, gender or marital status recorded at the time of the trial. There were also no significant differences between the proportions of participants and non-participants at long-term follow-up with respect to original treatment conditions. However, considerably higher proportions of long-term follow-up participants had completed initial treatment, as compared with either failed to start or dropped out of treatment, and this was significant for Trial 1, with 88% of participants having completed treatment versus only 69% of non-participants

[$\chi^2(1) = 4.3$, $p = 0.039$], and Trial 2, with 84% of participants having completed treatment versus only 59% of non-participants [$\chi^2(1) = 8.2$, $p = 0.004$]. In addition, long-term follow-up participants from Trial 1 were more likely to have come from a lower (i.e. less deprived) deprivation category than non-participants [$\chi^2(5) = 11.2$, $p = 0.048$].

There were no significant differences between participants and non-participants with regard to frequency of visits to GPs (primary care physicians) before and after the original trials, in the proportions on medication for anxiety at the time of the original trials, or with regard to previous psychological or psychiatric treatment.

In addition, there were no significant differences between participants and non-participants at long-term follow-up with regard to the presence of psychiatric diagnoses other than GAD at the time of the original trial, family history of mental health problems, whether the patient was a primary or secondary referral (Trial 2), the type of anxiety episode (e.g. recurrence or first episode – Trial 1) and duration of illness (Trial 2). However, in Trial 1, long-term follow-up participants were significantly more likely to have had a shorter duration of current episode of anxiety at the time of the trial (either of less than 1 month or of between 1 and 3 months versus 3–6 months or 6 months or more) than non-participants [71.4 versus 41.7%; $\chi^2(3) = 7.9$, $p = 0.048$]. There were no significant differences between participants and non-participants at long-term follow-up on any of the measures of psychological distress used in the original trials at pretreatment, post-treatment or short-term follow-up (up to 12 months) for Trials 1 and 2. However, for Trial 7, long-term follow-up participants had significantly lower pretreatment scores than non-participants on the CGS [participants mean = 5.2 versus non-participants mean = 5.7; 95% confidence interval (CI) for difference 0.01 to 0.8, $p = 0.044$]; BSI (participants mean = 1.45 versus non-participants mean = 1.81; 95% CI for difference 0.1 to 0.6, $p = 0.016$); HAM-A (participants mean = 16.3 versus non-participants mean = 18.5; 95% CI for difference 0.1 to 4.2, $p = 0.041$) and the marital subscale of the SAS (participants mean = 2.11 versus non-participants mean = 2.55; 95% CI for difference 0.1 to 0.7, $p = 0.004$). There were no significant differences between participants and non-participants at long-term follow-up on any of the main psychological distress measures used in the original trial either at post-treatment or short-term follow-up for Trial 7.

The PD trials (Trials 3, 4, 5 and 8)

There were no significant differences between long-term follow-up participants and non-participants with regard to age, gender, marital status or deprivation category for Trials 3, 4 and 8. However, patients from Trial 5 who participated at long-term follow-up had been significantly older at the time of the trial than non-participants [participants mean age = 41.9 years (SD = 12.7) versus non-participants mean age = 35.4 years (SD = 10.0); 95% CI for difference 0.6 to 12.4, $p = 0.032$]. In addition, significantly more men than women from Trial 5 took part at long-term follow-up [45 versus 31% in the original trial, $\chi^2(1) = 8.4$, $p = 0.004$).

There were no significant differences between the proportions of participants and non-participants at long-term follow-up with respect to original treatment conditions in Trials 4, 5 and 8. However, there was a significant difference with regard to treatment type in Trial 3 [$\chi^2(4) = 11.6$, $p = 0.021$], with long-term follow-up participants being less likely to have been in one of the two placebo treatment groups. This is a common finding in follow-up studies.

In all four trials, higher proportions of long-term follow-up participants had completed initial treatment, compared with either failed to start or dropped out of treatment, and this was significant for Trial 3, with 87% having completed the original treatment as compared to 71% of non-participants [$\chi^2(1) = 6.9$, $p = 0.009$].

There were no significant differences between participants and non-participants at long-term follow-up for any of the panic disorder trials with regard to previous panic episodes, avoidance, previous diagnoses of anxiety, depression, psychoses or other disorders, drug wash-out, the taking of benzodiazepine, antidepressant, beta-blocker, antipsychotic or other medication, previous psychotherapy, current medical condition, concurrent medication, type of panic disorder (with/without remission/agoraphobia), severity of panic disorder (mild, moderate, severe), or family psychiatric history. For Trials 3, 4 and 5 there were no differences with regard to the status or type of current episode at the time of the trial (e.g. continuation, recurrence, first episode), age of onset of panic attacks, number of years since first panic attack, ago of onset of avoidance or number of years of avoidance (data not recorded for Trial 8). For Trials 3, 5 and 8 there were no differences with respect to duration of current episode. However, in Trial 4, participants at long-

term follow-up had a significantly longer duration of current episode of panic disorder at the time of the trial than non-participants (65.1 versus 34.5 months; 95% CI for difference 3.1 to 58.0, $p = 0.029$).

When the measures of psychological well-being for the four panic disorder trials were examined, there were some differences between participants and non-participants in the long-term follow-up study with respect to scores at the time of the trial. Specifically, non-participants in Trial 3 had higher (i.e. 'worse') pretreatment scores on a measure of depression (the MADRS scale) and higher scores on a measure of general well-being (the GHQ[145]) both pre- and post-treatment (all $p < 0.05$). At 6-month follow-up there were no significant differences on the outcome measures between participants and non-participants in the long-term follow-up study for Trial 3.

There were also some differences between participants and non-participants in the follow-up study with respect to scores at the time of the trial for those in Trial 4. Non-participants at long-term follow-up had higher (i.e. 'worse') pretreatment scores on the HAM-A anxiety scale and there was also a significant difference on the pretreatment clinician measure of CGS, with non-participants scoring worse (both $p < 0.05$). Non-participants also scored worse in the STAI-T, for both state and trait measures taken at pretreatment ($p < 0.05$). However, there were no differences in any of these measures either at post-treatment or at 3- or 6-month follow-up.

There were no significant differences between long-term follow-up participants and non-participants with respect to pretreatment, post-treatment or short-term follow-up outcome measures used in Trial 5. There were also no significant differences between long-term follow-up participants and non-participants in Trials 3, 4 and 5 with regard to whether or not they still reported having panic attacks, either at day 84 or at short-term follow-up, or with respect to whether they had received intervening treatment between day 84 and short-term follow-up (data not recorded for Trial 8).

For Trial 8 there were no significant differences between long-term follow-up participants and non-participants on any of the outcome measures taken at pretreatment or at short-term follow-up. However, at post-treatment, participants had significantly better scores than non-participants on the Sheehan family subscale (participants mean = 2.3, non-participants mean = 3.8; 95% CI for difference 0.002 to 3.13, $p = 0.050$); the mental health summary component of the SF-36 (participants mean = 42.2, non-participants mean = 35.0; 95% CI for difference 0.7 to 13.8, $p = 0.031$) and the Bodily Sensations Questionnaire[146] (participants mean = 1.97, non-participants mean = 2.45; 95% CI for difference 0.02 to 0.9, $p = 0.040$), but not on any of the other outcome measures.

The PTSD study (Trial 6)

There were no differences between long-term follow-up participants and non-participants with regard to the majority of measures recorded at the time of the trial. Specifically, there were no differences with respect to gender, marital status, occupation, religion, previous psychiatric history of patient, family psychiatric history, co-morbid depression, medication at the time of the trial, previous history of trauma, type of therapy (i.e. EMDR or E+CR) or number of treatment sessions received during the trial. In addition there were no differences in the characteristics of the traumatic event which had led to the development of PTSD, including type of trauma (e.g. assault, vehicle accident), injury to self or others (including head injury and injuries which were disabling or disfiguring), warning of event, multiple traumatic events, whether or not others were involved, whether or not there were any fatalities, whether or not court proceedings were expected, the time from the trauma to the trial and the time from the trauma to the onset of PTSD.

There were a few differences between long-term follow-up participants and non-participants on measures recorded at the time of the study. Participants were older (mean age = 41.0 years) at the time of the original trial than non-participants (mean age = 33.8 years) (95% CI for difference 2.4 to 11.9, $p = 0.004$). In addition, participants had scored higher (i.e. worse) on the arousal subscale of the PSTD symptom checklist[147] (PTSD-SCL) than did non-participants, and this applied both pretreatment (participants 19.4 versus non-participants 17.7; 95% CI for difference 0.1 to 3.3), $p = 0.048$) and at 6-month follow-up (participants 12.7 versus non-participants 6.5; 95% CI 1.6 to 10.8, $p = 0.010$). Finally, participants had reported better social functioning at 6-month follow-up than non-participants on the Sheehan disability scale, in both their work (participants 4.8 versus 2.1; 95% CI for difference 0.1 to 5.3, $p = 0.042$) and family environments (participants 5.2 versus 2.2; 95% CI for difference 0.7 to 5.4, $p = 0.013$).

Psychosis studies (Trials 9 and 10)

There were no differences between long-term follow-up participants and non-participants with regard to the majority of measures recorded at the time of the trial in the two psychosis studies. Specifically, in Trial 9, there were no differences with respect to age at the time of the trial, age left school, care at referral (inpatient or outpatient), current residence, employment status, original diagnosis, co-morbidity, previous antipsychotic medication response, significant medical history, medication history (e.g. oral/depot), whether the family was available for interview at trial entry, duration of illness, number of treatment sessions, completion status and chlorpromazine equivalents throughout the trial. However, significantly more long-term follow-up participants from Trial 9 were female [$\chi^2(1) = 5.6$, $p = 0.018$] (with 67% of females taking part versus 36% of males), and long-term follow-up participants were more likely to be married/cohabiting or divorced/separated at the time of the trial as compared with being single [$\chi^2(2) = 7.1$, $p = 0.029$] (with 31% of participants married/cohabiting; 17% divorced/separated and 52% single versus non-participants: 14% married/cohabiting; 6% divorced/separated and 80% single). There appears to be an association between the findings with regard to gender and marital status, however, as significantly more males were single: 80 versus 38% of females [$\chi^2(1) = 16.0$, $p < 0.001$].

With regard to the outcome measures used in Trial 9, there were no differences between long-term follow-up participants and non-participants with regard to scores on the PANSS, CGS, CGI, SES and the PBIQ either at pre- or post-treatment or at short-term follow-up. There were no differences on the BSI-GSI at pretreatment or at short-term follow-up. However, at post-treatment, long-term follow-up participants [mean = 1.05 (SD = 0.59)] had significantly lower (i.e. more favourable) scores on the BSI-GSI than non-participants [mean = 1.76 (SD = 0.84); 95% CI for difference 0.2 to 1.3, $p = 0.011$]. With regard to global assessment of functioning, there were no differences between long-term follow-up participants and non-participants at pre- or post-treatment. However, long-term follow-up participants had had higher (i.e. more favourable) scores at short-term follow-up [mean = 38.7 (SD = 8.7) versus 33.3 (SD = 8.2); 95% CI for difference 0.4 to 10.4, $p = 0.035$]. Finally, for Trial 9, the only difference with regard to the seven subscales of the Social Functioning Scale (SFS) was that long-term follow-up participants had higher (i.e. more favourable) scores on the independence performance subscale at trial entry [mean 104.4 (SD = 13.8) versus 95.3 (SD = 14.6); 95% CI for difference 2.1 to 16.2, $p = 0.012$]. There were no other differences on the social functioning scales at pre- or post-treatment or at short-term follow-up.

In Trial 10, there were no differences between long-term follow-up participants and non-participants with respect to age at the time of the trial, gender, original diagnosis, presence of co-morbidity or type of co-morbid disorder, duration of illness, centre of referral or chlorpromazine equivalents throughout the trial. However, significantly more long-term follow-up participants had completed trial treatment [$\chi^2(1) = 5.8$, $p = 0.016$] (with 98% of long-term follow-up participants having completed treatment versus 88% of non-participants). In addition, long-term follow-up participants (mean = 4.9, SD = 0.4) had attended more treatment sessions during the original trial than non-participants [mean = 4.1 SD = 1.6); 95% CI for difference 0.05 to 1.17, $p = 0.034$], although data on this measure were available for only 71 patients (i.e. 49% of the sample).

With regard to the outcome measures used in Trial 10, there were no differences between long-term follow-up participants and non-participants with regard to scores on the global assessment of functioning scale, SES, the BSI-GSI and the PBIQ subscales either at pre- or post-treatment or at short-term follow-up. There were no differences on the PANSS at post-treatment or at short-term follow-up. However, at pretreatment long-term follow-up participants had significantly lower (i.e. more favourable) scores [mean = 10.1 (SD = 2.5)] on the PANSS positive symptom subscale than non-participants [mean = 11.2 (SD = 3.2); 95% CI for difference 0.1 to 2.0, $p = 0.029$]. In addition, on the withdrawal subscale of social functioning, long-term follow-up participants had had higher (i.e. more favourable) scores at short-term follow-up [mean = 104.7 (SD = 10.6) versus 100.6 (SD = 12.4); 95% CI for difference 0.1 to 8.1, $p = 0.045$]. There were no other differences on any of the main outcome measures at pre- or post-treatment or at short-term follow-up for long-term follow-up participants and non-participants from Trial 10.

Summary of differences between long-term follow-up participants and non-participants

Although some differences in patient characteristics, as recorded at the time of the original trial, were found between participants and

non-participants in the long-term follow-up study, for the most part these differences were inconsistent across studies, and in some cases contradicted each other (e.g. shorter duration of current episode of disorder in Trial 1 and longer duration of current episode in Trial 4). There was one consistent finding, however, namely that those who completed treatment were more likely to come back at follow-up. This is not surprising, as it is probable that these patients may be characterised by a willingness to comply with requests to participate in treatment or research programmes.

Despite these differences, however, the long-term follow-up sample does appear to be fairly representative of the original samples with regard to demographics, treatment conditions and treatment outcome. The evidence that non-participants often had worse pretreatment scores on the measures of outcome, that treatment completers were more likely to return at long-term follow-up and that, in Trial 3, those who responded positively to treatment were more likely to participate, all suggest that if there is any bias in the follow-up sample, it is likely to be towards a more, rather than less, favourable picture of long-term outcome.

Comparison of aggregate CBT on non-CBT groups on pretreatment characteristics for anxiety disorder studies (Trials 1–8)

As less than 50% of those entered into the original trials took part at long-term follow-up, it cannot be assumed that the randomisation to the CBT versus non-CBT conditions holds in the long-term follow-up participants. The aggregate CBT and non-CBT groups for Trials 1–8 were therefore compared on pretreatment characteristics, both for the whole sample and for long-term follow-up participants, using independent *t*-tests for quantitative variables and χ^2 tests for categorical data. Results for both the whole sample and the long-term follow-up participants were similar, so only those for the long-term follow-up participants are reported here.

With regard to the outcome measures, there were significant differences on pretreatment MADRS scores (95% CI for difference 1.4 to 7.0, $p = 0.004$), with the CBT group having lower depression scores (mean = 14.6, $n = 160$) than the non-CBT group (mean = 18.8, $n = 45$) and on the HAM-A (95% CI for difference 0.7 to 4.4, $p = 0.008$) with the CBT group having lower anxiety scores (mean = 22.3, $n = 296$) than the non-CBT group (mean = 24.9, $n = 91$). However, both of these results were confounded by study: the MADRS, as this was only used in half of the original studies, and the HAM-A, as the pretreatment scoring in Trial 2 was erroneously high. When the HAM-A scores for Trial 2 were adjusted for the suspected scoring bias, there were no differences on the HAM-A between the CBT and non-CBT groups. Comparisons on the amalgamated pretreatment scores of depression, anxiety and clinical severity (i.e. CDEP, CANX and COMCGS) were not significant, although CDEP approached significance ($p = 0.051$) with the CBT group having lower scores. There were no differences between the CBT and non-CBT groups on any of the other outcome measures used at pretreatment.

Duration of current episode was significantly longer in the CBT (mean = 52.7 months, $n = 292$) versus the non-CBT group (mean = 36.0 months, $n = 90$) (95% CI for difference 1.9 to 31.6, $p = 0.027$). The only other significant result with regard to the pretreatment demographic variables was deprivation category, where both deprivation category [$\chi^2(6) = 20.6$, $p = 0.002$] and deprivation score (95% CI for difference 0.5 to 1.7, $p < 0.001$) were significantly more favourable for the CBT group (i.e. the CBT group came from less deprived areas). This was true for both the whole sample and the long-term follow-up participants. There were no differences between the CBT and non-CBT groups with respect to the pretreatment variables of age, gender, marital or job status, completion status in original trial, co-morbidity, avoidance, number of treatment sessions or therapist contact hours.

Finally, time to long-term follow-up was significantly lower in the CBT group (mean = 6.5 years) than the non-CBT group (mean = 8.0 years) (95% CI for difference 0.9 to 2.2, $p < 0.001$) (however, extensive analysis revealed that there was no association between length of time to follow-up and long-term outcome).

Missing data

Anxiety disorder studies (Trials 1–8)

There was a noticeable amount of missing data across the variables used in the current analysis. *Table 14* shows the variables with more than 20%

TABLE 14 *Patterns of missing data for Trials 1–8*

		Data recorded at post-treatment				Data recorded at long-term follow-up							
		CDEP	**CANX**	**CGS**	**HAM-AD**	**Interim treatment**	**BSI-GSI**	**SF-36 PC**	**SF-36 MC**	**PA**	**NA**	**CGS**	**CGI**
Recorded at post-treatment	CDEP post	23.0											
	CANX post	1.6	24.1										
	CGS post	2.3	4.0	24.4									
	HAM-AD post	10.3	10.3	11.9	32.8								
Recorded at long-term follow-up	Interim treatment	45.4	44.5	44.5	45.2	55.0							
	BSI-GSI	45.9	45.0	45.0	45.9	0.7	55.3						
	SF-36 PC	47.1	46.1	46.1	46.6	2.3	2.1	56.9					
	SF-36 MC	47.1	46.1	46.1	46.6	2.3	2.1	0.0	56.9				
	PA	47.7	46.5	46.7	47.2	4.8	4.6	6.2	6.2	59.4			
	NA	47.7	46.5	46.7	47.2	4.8	4.6	6.2	6.2	0.0	59.4		
	CGS	48.4	47.4	47.4	47.4	4.8	5.3	6.4	6.4	4.2	4.2	59.8	
	CGI	48.5	47.5	47.5	47.8	4.9	5.4	6.8	6.8	3.9	3.9	0.4	60.0

The boxed diagonal elements are the total percentages of missing data on the column variable. The unboxed, shaded elements show the percentages of variables that have missing data between the respective post-treatment variables (pale grey) and between the respective long-term follow-up variables (dark grey). The unshaded elements are the percentages of missing data between the respective post-treatment and long-term follow-up variables. The table only includes variables with 20% or more missing values overall.

missing data and the percentage of each of the cases where they have a value where the other variables have missing values. The values on the diagonal are the overall percentages of missing data for each column variable (e.g. 23% of the sample had no data on the post-treatment CDEP and 55.3% had no data on the BSI-GSI at long-term follow-up). Values within the upper left quadrant reflect missing data within the post-treatment variables, values in the lower right quadrant reflect missing data within the long-term follow-up variables and those in the lower left quadrant reflect missing data between the post-treatment and long-term follow-up variables.

For example, with respect to patients attending post-treatment, 1.6% did not have scores on both the CANX and the CDEP (i.e. these patients had scores on one of these scales but not the other); and 45.9% of patients did not have both a CDEP score post-treatment **and** a BSI-GSI score at long-term follow-up (note: this figure is mainly due to attrition). As an example with regard to long-term follow-up participants, 2.1% failed to complete both the BSI-GSI and the SF-36.

Given the large amount of missing data, *t*-tests were performed between those patients who had missing values on the LTOF and those who had not. The results are shown in *Table 15*.

Those with missing values on the LTOF had at least one of the CGS score, the CGI score, the BSI-GSI score or one of the PANAS subscales or SF-36 summary components missing at long-term follow-up. Twenty-one *t*-tests were performed of which only two were significant and those only at the $p < 0.05$ level, which suggests that there was little difference between the patients with missing values at long-term follow-up and those without. If the significant tests are taken at face value (ignoring the problems of multiple testing), the missing patients are younger and have a longer duration of presenting episodes than the patients with data, but even then the effects are slight (see *Table 15*).

In addition, χ^2 tests were performed on whether having missing data on the long-term outcome factor was related to the pretreatment categorical variables – gender, taking concurrent psychotropic

TABLE 15 *Separate variance t-tests depending on whether all the variables contributing to the long-term outcome factor were present or one was missing for Trials 1–8*

	t	df	*p* (2-tailed)	*n* present	*n* missing	Mean (present)	Mean (missing)
Pretreatment							
Age	2.3	695.7	0.024	312	526	39.02	37.17
Deprivation score	–0.9	565.8	0.382	309	532	–0.09	0.06
Duration of current episode	2.0	685.0	0.043	308	518	4.29	4.14
CFAM	–1.0	571.1	0.298	281	442	–0.05	0.03
CSOC	0.6	609.4	0.543	273	430	0.03	–0.01
CDEP	–1.9	683.5	0.058	307	515	0.08	0.23
CGS	–0.1	640.5	0.958	310	531	4.87	4.87
HAM-AD	–0.5	623.6	0.634	311	533	21.00	21.20
CANX	–1.4	625.8	0.156	303	509	–0.03	0.07
Post-treatment							
CDEP	–1.3	624.6	0.205	268	390	0.00	0.10
CGS	0.5	569.1	0.604	266	380	2.73	2.65
HAM-AD	0.6	508.4	0.574	233	341	10.31	9.91
CANX	–0.1	591.0	0.913	268	380	–0.06	0.00
Long-term follow-up							
Interim treatment	0.2	99.6	0.814	312	72	2.26	2.22
CGS	–1.8	35.6	0.083	312	31	3.27	3.90
CGI	–1.5	33.6	0.141	312	30	2.24	2.77
BSI-GSI	0.1	97.3	0.931	312	70	1.21	1.20
PA	–0.5	39.7	0.644	312	35	29.27	30.09
NA	0.3	42.4	0.766	312	35	24.19	23.69
SF-36 PC	–1.4	74.0	0.171	312	56	40.43	43.24
SF-36 MC	–0.4	72.4	0.726	312	56	40.26	40.92

df, degrees of freedom.

medication, employment status, marital status and evidence of avoidance. Equivalent χ^2 tests were performed on whether CBT was given and whether the course of treatment was completed. Only two of these seven tests were significant ($p < 0.05$) and only the tendency for those patients who completed their treatment to return at long-term follow-up was appreciable. *Table 16* shows the significant results.

Summary of missing data in anxiety disorder studies (Trials 1–8)

The missing data analyses show that having missing data at long-term follow-up appears to be unrelated to most other measures. All of the effects of missing data are small. Only age, duration of current episode, taking concurrent medication and completion status showed significant differences between those with and without missing data. The only appreciable difference was completion status, which was not surprising given that we found in the section 'Statistical analysis' (p. 31) that treatment completers were more likely to return at long-term follow-up.

Psychosis studies (Trials 9 and 10)

There was a noticeable amount of missing data across the variables used in the current analysis. *Table 17* shows the variables used as outcome

TABLE 16 *Breakdown of presence or absence of long-term follow-up outcome measures with categorical variables for Trials 1–8*

Pretreatment					
Concurrent medication	No	Yes	χ^2	*p*	Cramer's *V*
Present	125	186	5.88	0.015	0.083
Absent	263	276			
Trial variables					
Non-completer	No	Yes	χ^2	*p*	Cramer's *V*
Present	54	258	18.4	<0.001	0.147
Absent	166	376			

TABLE 17 *Patterns of missing data for Trials 9 and 10*

		Data recorded at post-treatment			Data recorded at long-term follow-up					
		PANSS	**Social functioning**	**BSI-GSI**	**PANSS**	**CGS**	**CGI**	**BSI-GSI**	**SF-36 PC**	**SF-36 MC**
Post-treatment	PANSS	8.6								
	Social functioning	1.0	9.0							
	BSI-GSI	19.3	18.9	25.2						
Recorded at long-term follow-up	PANSS	52.6	52.9	54.2	55.7					
	CGS	52.6	52.9	54.2	0.0	55.7				
	CGI	53.6	53.9	54.8	2.2	2.2	56.7			
	BSI-GSI	54.7	55.0	56.1	4.3	4.3	3.3	57.6		
	SF-36 PC	55.7	56.0	57.4	6.5	6.5	5.5	6.7	58.6	
	SF-36 MC	55.7	56.0	57.4	6.5	6.5	5.5	6.7	0.0	58.6

The boxed diagonal elements are the total percentages of missing data on the column variable. The unboxed, shaded elements show the percentages of variables that have missing data between the respective post-treatment variables (pale grey) and between the respective long-term follow-up variables (dark grey). The unshaded elements are the percentages of missing data between the respective post-treatment and long-term follow-up variables. The table only includes outcome measures used in the analyses in Chapters 7 and 9. Where differences occurred with respect to the amount of missing data within the subscales of either the PANSS or social functioning measures, the highest levels of missing data are reported in this table.

measures in this report and the percentage of each of the cases where they have a value where the other variables have missing values.

The values on the diagonal are the overall percentages of missing data for each column variable (e.g. 8.6% of the sample had no data on the post-treatment PANSS scale and 55.7% had no data on the PANSS scale at long-term follow-up).

Values within the upper left quadrant reflect missing data within the post-treatment variables, values in the lower right quadrant reflect missing data within the long-term follow-up variables and those in the lower left quadrant reflect missing data between the post-treatment and long-term follow-up variables. For example, with respect to patients attending post-treatment, 1.0% did not have scores on both the PANSS and social

TABLE 18 *Separate variance t-tests depending on whether all of the long-term follow-up outcome measures were present or one or more was missing for Trials 9 and 10*

	t	df	*p* (2-tailed)	*n* present	*n* missing	Mean (present)	Mean (missing)
Pretreatment							
Age	0.09	208	0.931	80	130	36.2	36.3
Deprivation score	1.58	199	0.117	78	123	1.4	2.1
Duration of illness	0.05	197	0.964	79	120	128.2	128.8
Chlorpromazine equivalent	0.06	194	0.953	78	118	464.8	468.0
PANSS positive	0.20	201	0.844	79	124	15.0	15.2
PANSS negative	0.41	201	0.686	79	124	16.7	17.2
PANSS general	0.17	201	0.863	79	124	35.2	35.4
PANSS total	0.40	201	0.699	79	124	66.9	68.2
BSI-GSI	1.65	184	0.101	75	111	1.26	1.47
Social functioning							
Withdrawal	–0.14	200	0.893	79	123	94.5	94.2
Interpersonal communication	–1.10	200	0.272	79	123	115.4	112.7
Independence performance	–2.34	200	0.020	79	123	104.9	100.6
Independence competence	–0.79	199	0.431	79	122	108.6	107.1
Recreation	–1.48	200	0.140	79	123	97.2	94.1
Prosocial	–0.42	200	0.674	79	123	100.1	99.3
Employment	0.10	198	0.919	79	121	93.1	93.3
Post-treatment							
PANSS positive	0.14	190	0.888	78	114	13.7	13.9
PANSS negative	–0.16	190	0.872	78	114	15.7	15.5
PANSS general	0.36	190	0.719	78	114	32.1	32.7
PANSS total	0.20	190	0.839	78	114	61.5	62.3
BSI-GSI	2.73	155	0.007	60	97	0.89	1.25
Social functioning							
Withdrawal	0.07	189	0.942	78	113	95.6	95.8
Interpersonal communication	–1.28	189	0.203	78	113	118.6	115.0
Independence performance	–2.02	190	0.045	78	114	106.0	102.3
Independence competence	–1.07	190	0.285	78	114	110.6	108.4
Recreation	–0.56	190	0.578	78	114	97.0	95.8
Prosocial	–0.55	189	0.580	77	114	101.8	100.6
Employment	–0.38	188	0.702	78	112	95.6	94.9
Long-term follow-up							
CGS	1.42	91	0.158	80	13	3.96	4.69
CGI	2.38	89	0.037	80	11	2.29	3.82
BSI-GSI	0.78	87	0.435	80	9	1.10	1.33
PANSS positive	–1.51	91	0.134	80	13	17.0	14.2
PANSS negative	–1.15	91	0.252	80	13	16.9	14.8
PANSS general	–0.80	91	0.424	80	13	36.2	33.5
PANSS total	–1.29	91	0.199	80	13	70.0	62.4
SF-36 PC	–1.08	85	0.284	80	7	30.9	37.5
SF-36 MC	–0.96	85	0.342	80	7	43.0	37.2

functioning measures (i.e. these patients had scores on one of these scales but not the other) and 54.7% of patients with a post-treatment PANSS score did not have a BSI-GSI score at long-term follow-up (note: this figure is mainly due to attrition). As an example with regard to long-term follow-up participants, 6.7% failed to complete both the BSI-GSI and the SF-36.

Given the large amount of missing data, *t*-tests were performed between those patients ($n = 130$) who had missing values on one or more of the long-term outcome measures listed in *Table 17* and those who had not ($n = 80$). The results are shown in *Table 18*. Those with missing values had at least one of the CGS score, the CGI score, the BSI-GSI score or one of the PANSS subscales or SF-36 summary components missing at long-term follow-up. Thirty-seven *t*-tests were performed, of which only four were significant, and only one of those with $p < 0.01$, which suggests that there was little difference between the patients with missing values at long-term follow-up and those without. If the significant tests are taken at face value (ignoring the problems of multiple testing), the missing patients had worse scores both at pre- and post-treatment with regard to independence performance on the social functioning scale, worse BSI-GSI scores at post-treatment (but not at pre-treatment or long-term follow-up) and considered themselves less improved (on the CGI) at long-term follow-up (see *Table 18*).

In addition χ^2 tests were performed on whether having missing data on one or more of the long-term outcome measures was related to gender, CBT group or completion of original treatment. Only treatment completion was significant, with 96.3% of those with no missing data on the long-term outcome measures having completed treatment versus 87.7% of those with missing data [$\chi^2(1) = 4.4$, $p = 0.036$].

Summary of missing data in psychosis studies (Trials 9 and 10)

The missing data analyses show that having missing data at long-term follow-up appears to be unrelated to most other measures. All of the effects of missing data are small. Only social functioning with regard to independence performance, post-treatment BSI-GSI scores, long-term follow-up CGI and completion status showed significant differences between those with and without missing data. The only appreciable difference was on the BSI-GSI recorded at post-treatment; however, there were no significant differences on the pretreatment or long-term follow-up scores for this measure. The fact that all of the significant differences showed preferential results for those with no missing data on the long-term outcome variables suggests that, if anything, those with missing data at long-term follow-up would have been expected to have slightly poorer outcome.

Chapter 6

Summary of long-term outcome by disorders and for aggregate data

Observer-rated measures

Diagnostic status

Anxiety disorder studies (Trials 1–8)

Overall, just over half (52%) of patients still had at least one clinical diagnosis as assessed via the ADIS-IV at long-term follow-up. The mean number of diagnoses for all patients was 1.4 (SD = 1.7), and the mean number of diagnoses for patients with at least one clinical diagnosis was 2.7 (SD = 1.8), demonstrating a high level of co-morbidity at long-term follow-up. The percentage of patients by number of long-term follow-up diagnoses is shown in *Figure 1*.

The percentage of long-term follow-up patients with each type of diagnosis is shown in *Figure 2*. In addition (not shown in *Figure 2*), one patient from Trial 2 had a diagnosis of somatisation and one patient from Trial 3 had a diagnosis of manic depression.

Across the whole sample, the most frequently occurring diagnoses were GAD (27% of long-term follow-up participants), agoraphobia (26%), panic disorder (23%) and depression (22%).

The proportion of patients originally treated for GAD (Trials 1, 2 and 7) who had at least one clinical diagnosis at long-term follow-up was 52%. The proportion of this group who still had GAD at long-term follow-up was 34% and, in each of these trials, GAD was the most likely diagnosis at long-term follow-up. There were also moderately high levels of depression (25%), agoraphobia (19%), dysthymia (18%), panic disorder (17%) and social phobia (15%) in patients from the GAD trials.

The proportion of patients who were originally treated for panic disorder (Trials 3, 4, 5 and 8) who had any clinical diagnosis at long-term follow-up was 48%. The proportion in this group who still had panic disorder at long-term follow-up was 26%, and the proportion of these patients with an agoraphobia diagnosis at long-term follow-up was 29%. The higher percentage of agoraphobia over panic disorder diagnoses may in part be due to a number of patients who appeared to be using

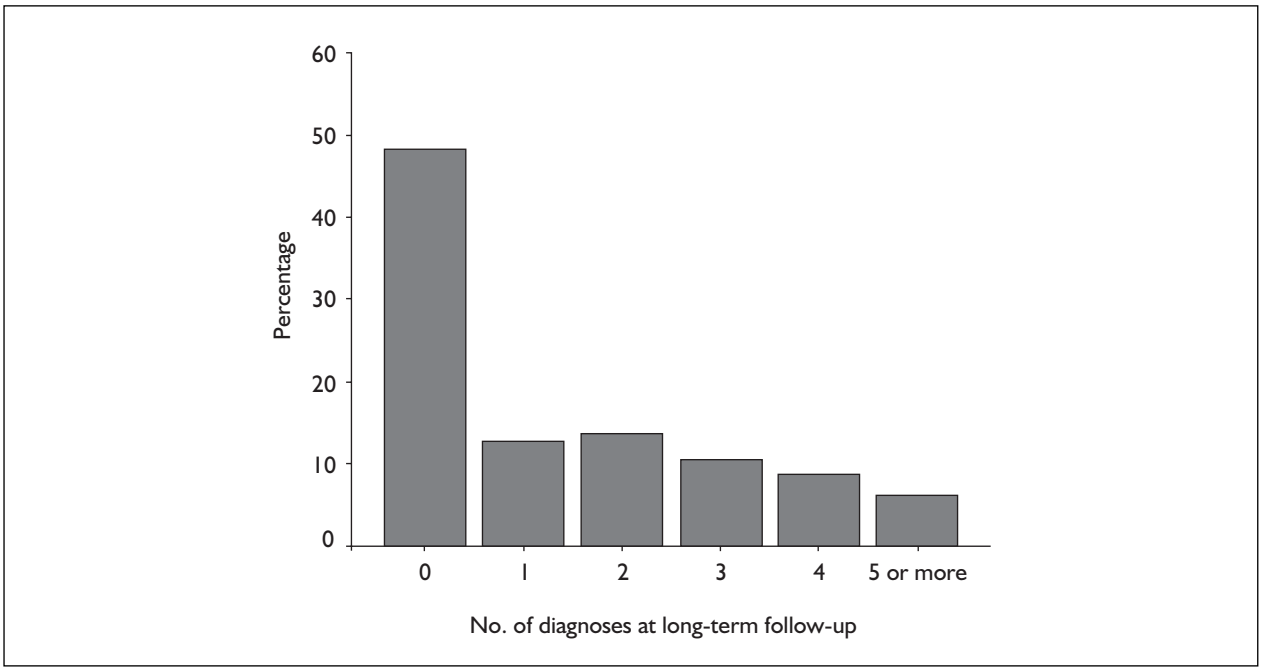


FIGURE 1 *Long-term follow-up patients (n = 344) by number of clinical diagnoses (assessed using ADIS-IV) for Trials 1–8*

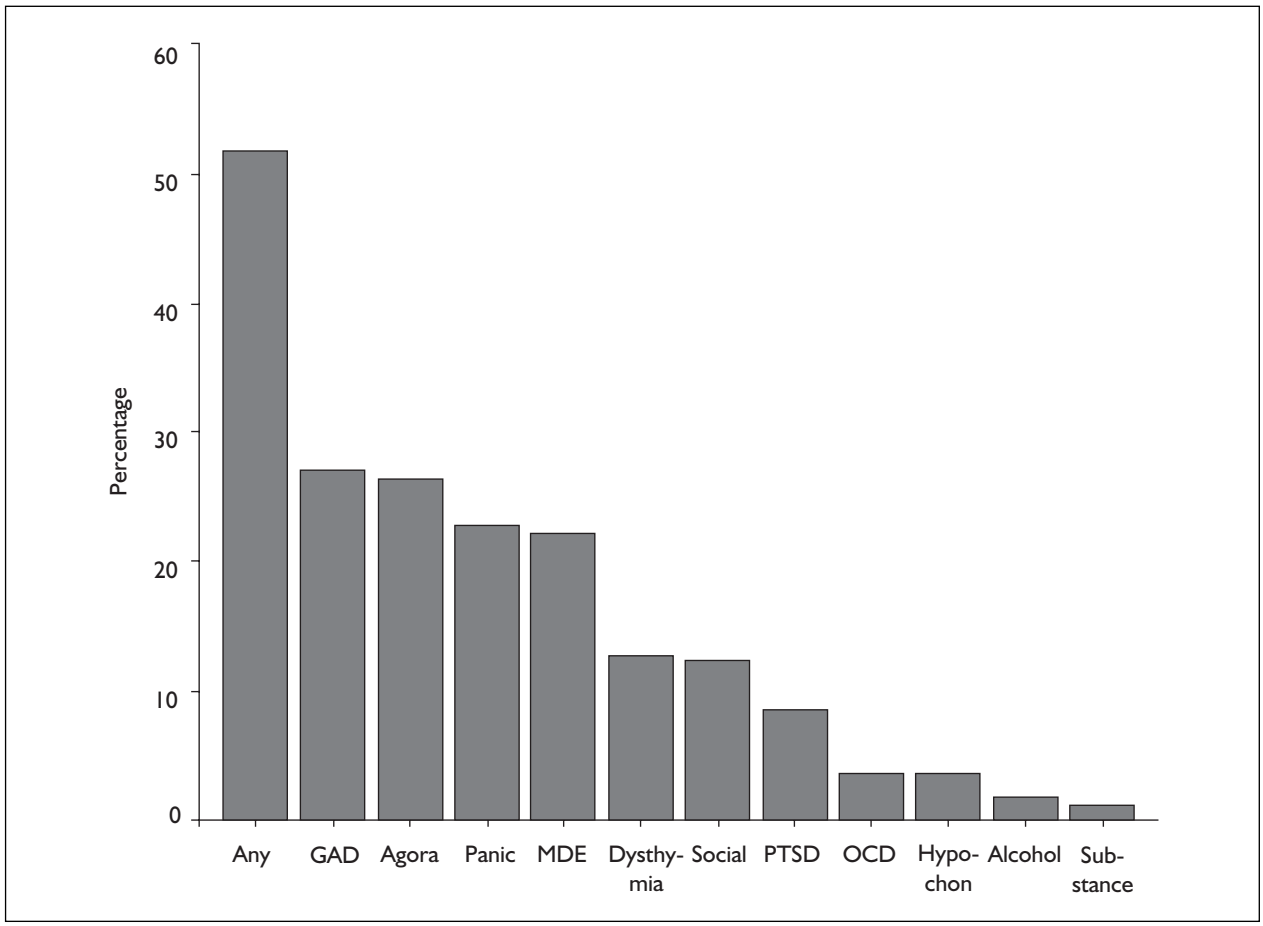


FIGURE 2 *Percentage of long-term follow-up patients (n = 344) with each diagnosis for Trials 1–8*

avoidance as a coping mechanism, and who had thus reduced their panic attacks to nil by totally avoiding all anxiety-provoking situations. In Trials 5 and 8, panic disorder and/or agoraphobia were the most likely diagnoses at long-term follow-up. In Trials 3 and 4, however, the proportion of those with a GAD diagnosis at long-term follow-up was 4–5% higher than those with panic disorder and 2–6% higher than agoraphobia. The overall occurrence of GAD at long-term follow-up in patients originally treated for panic disorder was 23%; there were also moderately high levels of depression (18%), social phobia (12%) and dysthymia (9%) in this group.

The proportion of patients who were originally treated for PTSD (Trial 6) who had any clinical diagnosis at long-term follow-up was 74%. Over half (55%) of the patients from Trial 6 still had a diagnosis of PTSD at long-term follow-up. In addition, fairly large percentages of PTSD patients had diagnoses of agoraphobia (39%), depression (36%), panic disorder (29%), GAD (23%), social phobia (23%) and dysthymia (13%) at long-term follow-up. Indeed, the patients originally treated for PTSD appeared to fare worst overall.

These results indicate that, irrespective of original diagnosis, patients originally diagnosed with anxiety disorders who retained a clinical diagnosis over periods of 2–14 years, were not only likely to have multiple diagnoses, but also may fluctuate between different disorders. Perhaps not surprisingly in a group of patients with such chronicity, depression was a common co-morbid disorder (43% of those with any clinical diagnosis at long-term follow-up, of which only 3% had no other diagnosis).

Psychosis studies (Trials 9 and 10)

Diagnostic status was not assessed for Trials 9 and 10 at long-term follow-up. Diagnostic status is not recognised as a key outcome for psychological treatments for people diagnosed with schizophrenia, rather the aims of these treatments is to reduce the severity of symptomatology.

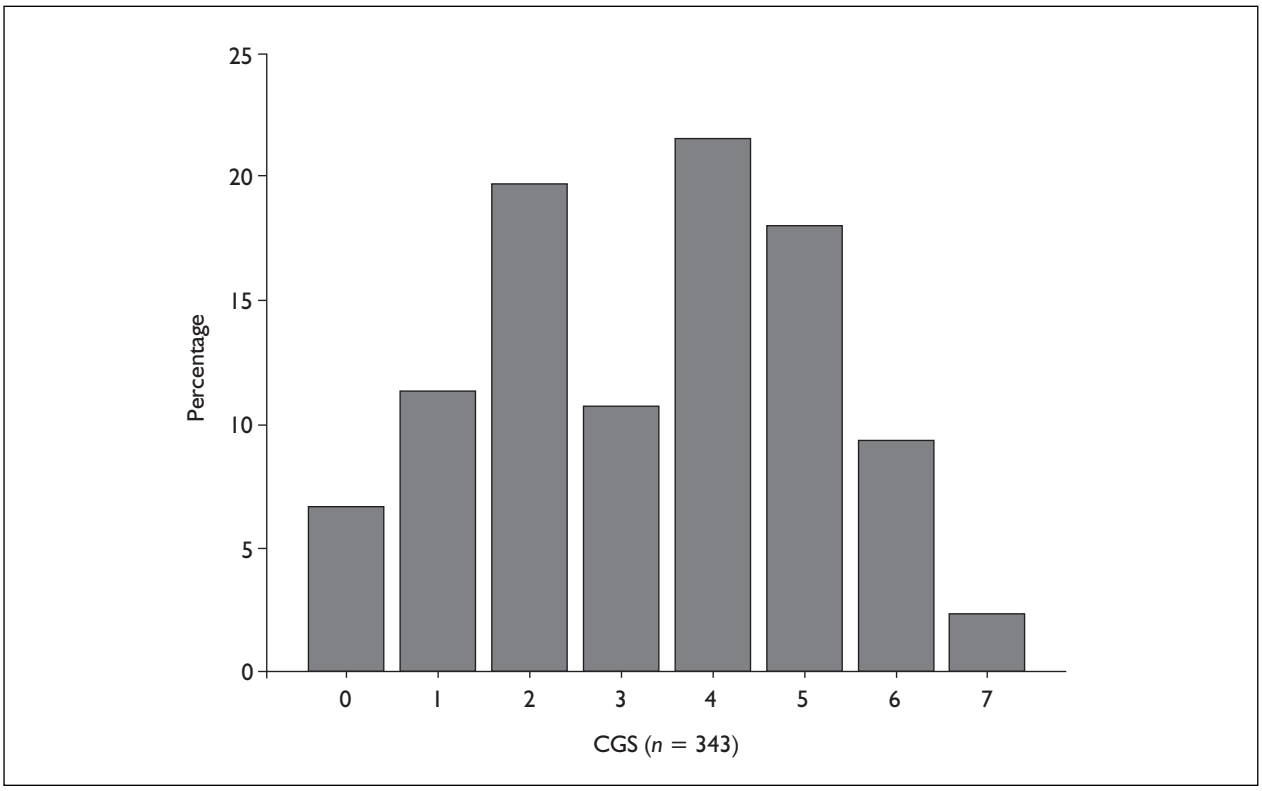


FIGURE 3 *Clinical global severity at long-term follow-up for Trials 1–8*

Clinical Global Severity scores

Anxiety disorder studies (Trials 1–8)

Figure 3 shows a bar chart of CGS scores at long-term follow-up for the eight anxiety disorder studies. The mean CGS score of 3.3 (SD 1.8) for the whole sample is below the cut-off point of 4, which equates to having a clinical diagnosis. Only 7% of long-term follow-up participants had a CGS score of zero, that is, had no symptoms of any anxiety or depressive disorder, a further 11% had very mild symptoms and, at the other end of the spectrum, 12% had symptoms considered severely distressing or disabling. In addition, nearly one-third of patients had subclinical levels of symptoms (20% with mild symptoms and 11% with a score just below the diagnostic threshold), and it was felt that many of these patients were likely to be prone to relapse in the light of difficult life circumstances (indeed, one patient who had taken part in multiple trials originally did switch diagnostic status between her two follow-up visits, which were held at an interval of 22 months).

Psychosis studies (Trials 9 and 10)

Figure 4 shows a bar chart of CGS scores at long-term follow-up for the two psychosis studies. The mean CGS score for the whole sample is 4.1 (SD 1.7). Only 1% of long-term follow-up participants had a CGS score of zero, that is, had no symptoms, and only 9% had very mild symptoms. In contrast, 23% had symptoms considered severely distressing or disabling (or worse) and another 40% had symptoms which were definitely disabling (i.e. CGS ≥4).

Comparison of anxiety disorder and psychosis studies

An independent *t*-test of the differences in scores of clinical global severity at long-term follow-up between the anxiety disorder and psychosis studies was significant ($t = 3.5$, 95% CI for difference 0.4 to 1.1, $p < 0.001$), with the psychosis patients having the worst scores. A χ^2 test by individual CGS category was also significant, with more of the psychosis patients having scores in the higher categories and fewer in the lower categories [$\chi^2(7) = 18.8$, $p = 0.009$].

Clinical status

Jacobson method applied to primary outcome measures

Anxiety disorder studies (Trials 1–8)

Jacobson criteria[148] were used to calculate clinically significant change at long-term follow-up on the main outcome measures used in each of the anxiety disorder studies. Two methods were used

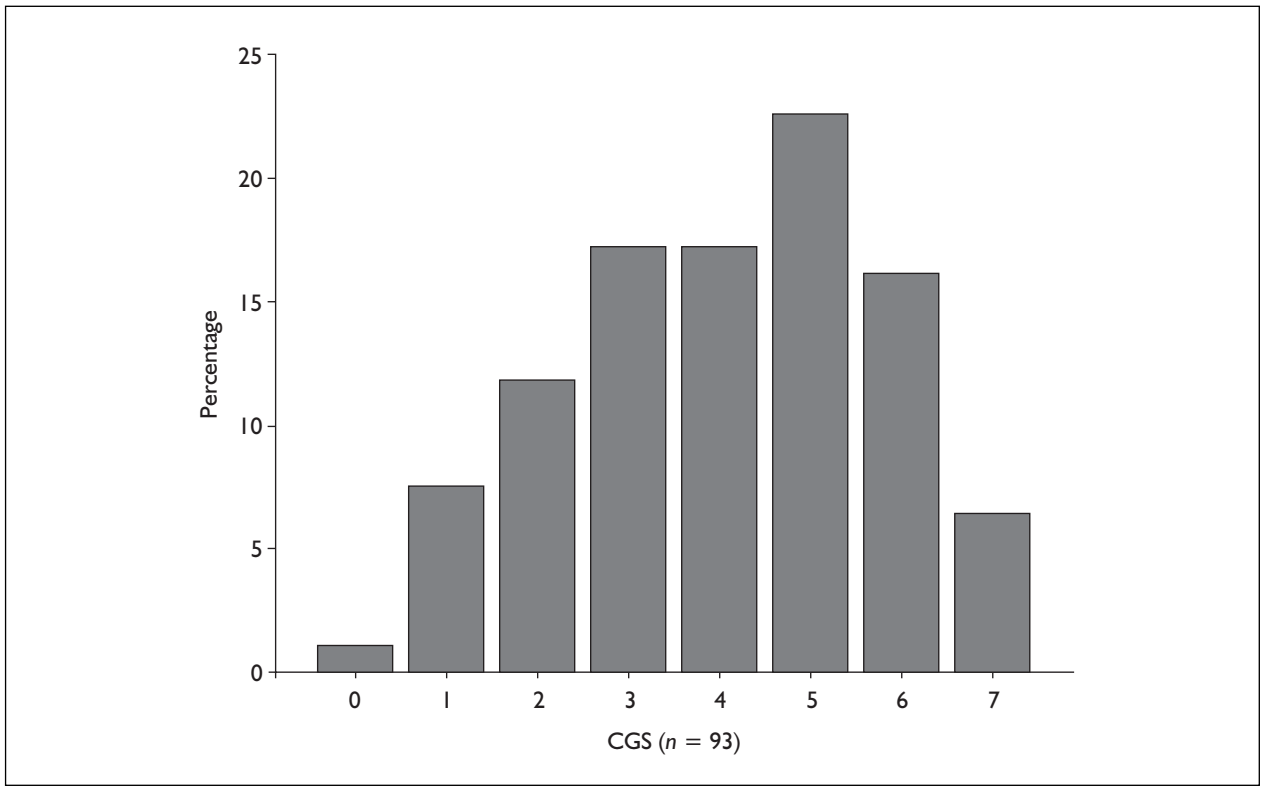


FIGURE 4 *Clinical global severity at long-term follow-up for Trials 9 and 10*

depending on whether or not normative data were available. For those measures with no normative data, Jacobson criterion (a) was used, that is, with a cut-off point for achieving clinically significant change equal to two SDs below the pretreatment mean. Criterion (c) was used for the measures where population norms were available. Normative functioning according to Jacobson criterion (c) is defined as demonstrating a reliable change between pretreatment and follow-up scores, in addition to having a score on the relative scale below a cut-off point. For criterion (c), the cut-off is calculated as the mid-point between two SDs below the pretreatment mean and two SDs above the mean for the normal population.

The Jacobson criteria applied in each of the original studies was used to determine clinically significant change at long-term follow-up on the main outcome measures. (The analysis was repeated using recalculated Jacobson criteria based on pretreatment data across all studies. The resulting percentages achieving clinically significant change were remarkably similar on all measures to those reported above, with one exception: on the HAM-A the percentages were reduced to 67% at short-term follow-up and 42% overall at long-term follow-up. This difference was due to the small pretreatment variances on the HAM-A in Trials 1 and 3–5 resulting in fairly high cut-offs in the original studies). Jacobson criteria were determined as follows: HAM-A – criteria (a) (Trial 1, <12; Trials 3–5, <13; Trial 6, <14; Trial 7, <8; Trial 8 [Jacobson criteria were not applied to the HAM-A in the original report for Trial 8, so these have been calculated using the same criteria (a) as the remaining studies at long-term follow-up], <7); SRT – criteria (a) (Trial 1, <16; Trial 3, <6; Trial 4, <13; Trial 5, <11); STAI-T – criteria (c) plus reliable change (rc) (Trials 2 and 7, <47; rc > 7; Trial 8 rc > 7 only); BSI-GSI – criteria (c) plus reliable change (Trial 2, ≤ 0.98, rc ≥ 0.51; Trial 7, ≤ 0.91, rc ≥ 0.64); FQ-Agora – criteria (c) (Trials 3–5 cut-off only, <9; Trial 8, rc > 9 only).

Table 19 shows the percentage of patients achieving clinically significant change on these Jacobson criteria at short-term and long-term follow-up.

Table 20 shows those patients achieving Jacobson criteria on the main outcome measures used at long-term follow-up, irrespective of whether they also attended short-term follow-up. The results show the percentage achieving clinically significant change according to the Jacobson criteria used within each of the original trials for the HAM-A, SRT and FQ-Agora.

TABLE 19 *Recovery rates at short-term (3–12 months) and long-term (2–14 years) follow-up using Jacobson criteria applied to the primary self-report outcome measures for Trials 1–8 (based on the calculations used in the original studies)*

	Short-term follow-up		Long-term follow-up					
	Recovered		Recovery maintained		Recovery achieved		Recovery rate overall	
	n	%[a]	*n*	%[a]	*n*	%[a]	*n*	%[a]
HAM-A	143	80	87	49	6	3	93	52
SRT	68	48	23	16	7	5	30	21
STAI-T	30	29	20	19	24	23	44	43
BSI-GSI	24	30	13	17	8	10	21	27
FQ-Agora	99	72	73	52	13	9	86	62

[a] % represents the number achieving or maintaining recovery as a percentage of long-term follow-up participants also attending short-term follow-up for each outcome measure as follows: HAM-A, $n = 178$, Trials 1, 3–8; SRT, $n = 142$, Trials 1, 3–5; STAI-T, $n = 103$, Trials 2, 7, 8; BSI-GSI, $n = 79$, Trials 2, 7; FQ-Agora, $n = 138$, Trials 3–5, 8.

TABLE 20 *Patients achieving clinically significant change on main outcome measures at long-term follow-up for Trials 1–8*

	Patients achieving clinically significant change at long-term follow-up	
	n	%
HAM-A	290	45
SRT	209	18
STAI-T[a]	336	44
BSI-GSI[a]	389	57
FQ-Agora	214	56

[a] Cut-offs [Jacobson criterion (c) without reliable change] calculated using all available pretreatment data, applied universally across all trials at long-term follow-up.

On the BSI-GSI there were no pretreatment values available for Trials 1–6 and 8, and the STAI-T was not used pretreatment in Trials 1, 3, 5 and 6, so cut-off points were calculated based on available data and then applied to long-term follow-up scores for all studies. On the BSI-GSI, the cut-off point as calculated according to Jacobson criterion (c) for patients in Trials 2 and 7 (i.e. 0.94) was used across all studies, as an indicator of 'good' and 'poor' scores. Similarly, for the STAI-T, the cut-off score of 46 (without reliable change) was applied across all studies. The results are shown in *Table 20*. The percentages achieving clinically significant change are similar to those shown in *Table 19*, with the exception of the BSI, where the recalculated values are much higher. This is mainly due to only the cut-off point and not the reliable change being applied in the calculation for all studies.

Psychosis studies (Trials 9 and 10)

The BSI-GSI was used at both pretreatment and long-term follow-up in Trials 9 and 10. We were therefore able to calculate Jacobson criterion (c) on the BSI-GSI for the psychosis patients. This resulted in a cut-off point of 0.77 and a reliable change criterion of 0.76. Using this criterion, 38% of patients (of $n = 89$) from the psychosis studies had scores less than the specified cut-off point on the BSI-GSI at long-term follow-up, but only 12.4% had achieved Jacobson criterion (c) with reliable change. The latter was due in part to the large variation in pretreatment scores in this group of patients, resulting in a relatively large reliable change value.

Comparison of anxiety disorder and psychosis studies

In order to compare anxiety and psychosis patients with regard to Jacobson criteria on the BSI-GSI, we applied (1) the cut-off point of 0.94, calculated from all available data from the anxiety disorder studies and (2) the cut-off of 0.77 calculated from the psychosis patients, to both groups of patients. This resulted in (1) 47% of the psychosis patients achieving the cut-off versus 57% of the anxiety disorder patients [$\chi^2(1) = 2.9$, $p = 0.091$] and (2) 36% of the anxiety patients versus 38% of the psychosis patients [$\chi^2(1) = 0.2$,

TABLE 21 *Mean PANSS scores at long-term follow-up and numbers of patients achieving 25 and 50% improvement from pre-treatment (n = 93) for Trials 9 and 10*

PANSS	Long-term follow-up: mean (SD)	Reduction in scores from pretreatment to long-term follow-up			
		Patients achieving 25% reduction		Patient achieving 50% reduction	
		n	%	*n*	%
Positive symptoms	16.6 (6.3)	8	8.7	1	1.1
Negative symptoms	16.6 (6.1)	22	23.9	8	8.7
General symptoms	35.8 (10.9)	13	14.1	2	2.2
Total score	68.9 (19.8)	9	9.8	1	1.1

$p = 0.646$]. Hence we can see that applying consistent Jacobson criteria across all studies revealed no significant differences between the two groups of patients with regard to clinical status on the BSI-GSI.

Percentage change in PANSS scores for psychosis studies (Trials 9 and 10)

As a measure of clinical change in the two psychosis studies, we examined the proportion of patients who showed at least a 25 and 50% decrease in symptom severity on the PANSS at long-term follow-up. These figures were chosen in order to make meaningful comparisons with the three main clinical trials published in this area.[62–64] Although both figures are somewhat arbitrary, most clinicians with experience of chronic schizophrenia are likely to regard a 25% improvement as being worthwhile and a 50% improvement as representing an important clinical change.

The mean PANSS scores at long-term follow-up for the two psychosis studies are shown in *Table 21*, along with the number achieving 25 and 50% changes in scores from pretreatment values.

It can be seen from *Table 21* that the biggest reductions were seen in the negative symptom subscale with nearly one-quarter (24%) of patients achieving a ≥25% reduction in scores from pretreatment values, and 9% achieving a 50% reduction. Overall, however, the percentages of patients achieving a notable (i.e. ≥25%) reduction in total PANSS symptoms from pretreatment to long-term follow-up was low (<10%).

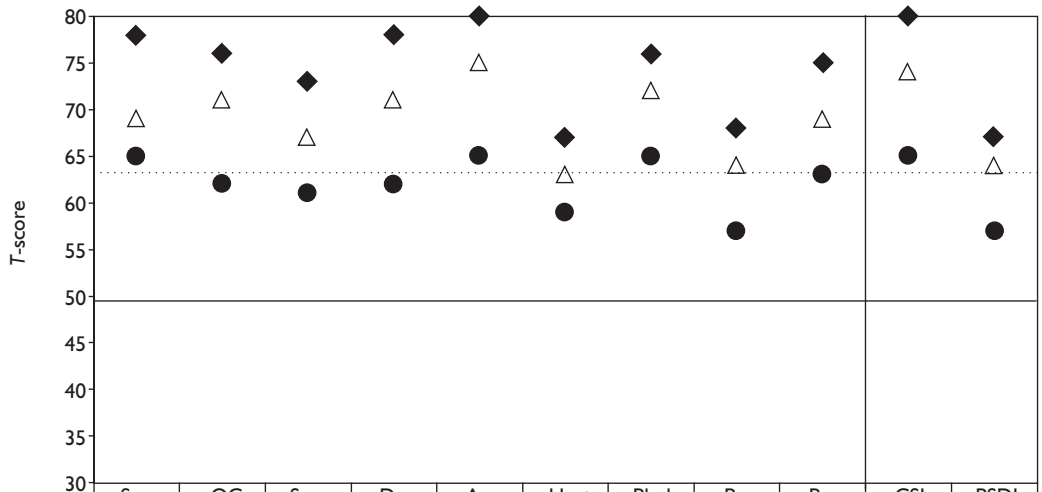
Clinical profiles on Brief Symptom Inventory

Anxiety disorder studies (Trials 1–8)

The profile of long-term follow-up participants on the BSI subscales, in comparison with the population norms for adult non-patient data[135] for the eight anxiety disorder studies are shown in *Figures 5* and *6*. *Figure 5* shows *T*-scores for male patients for the whole sample (all data), for those with no clinical diagnosis ('no diag'), and for those with at least one diagnosis ('any diag') at long-term follow-up. *Figure 6* shows the same data for female patients. In each figure actual *T*-scores and raw score values are shown in a table at the foot of the diagram. The dotted line at *T*-score = 63 represents caseness for each subscale, as defined by Derogatis,[135] that is, a score on or above the line would be equivalent to a positive diagnosis or case. The solid line at *T*-score = 50 represents the mean of the normative data (SD = 10).

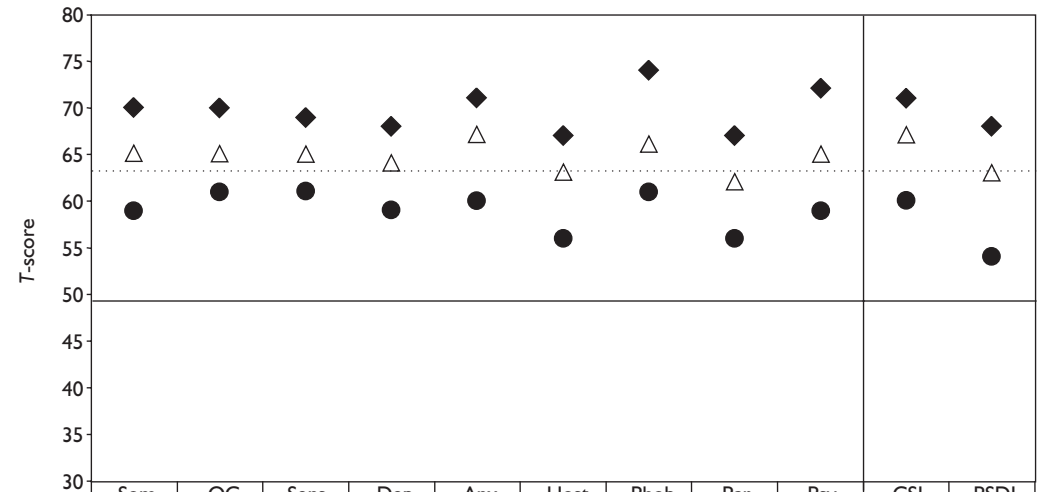
It can be seen from *Figure 5* that the *T*-scores of the mean raw scores for the whole sample are higher than the caseness cut-off points for each of the BSI subscales, and also for the GSI and Positive Symptom Distress Index (PSDI). Male participants with no diagnosis at long-term follow-up also had *T*-scores on or slightly above the caseness cut-off point on the somatisation, anxiety, phobic anxiety and psychoticism subscales, in addition to the GSI. Male patients who had at least one diagnosis at long-term follow-up have *T*-scores which are noticeably above the caseness cut-off points for each of the BSI subscales.

It can be seen from *Figure 6* that female patients with no diagnosis have *T*-scores which are all below the caseness cut-off points, and female patients with any clinical diagnosis have T-scores above the caseness cut-off points, for all BSI subscales. *T*-scores for the whole sample of female participants at long-term follow-up are higher than the caseness cut-off points for all of the BSI subscales, with the one exception of paranoid ideation (mean = 62), which is just below the cut-off. Although raw long-term follow-up scores on the BSI subscales are



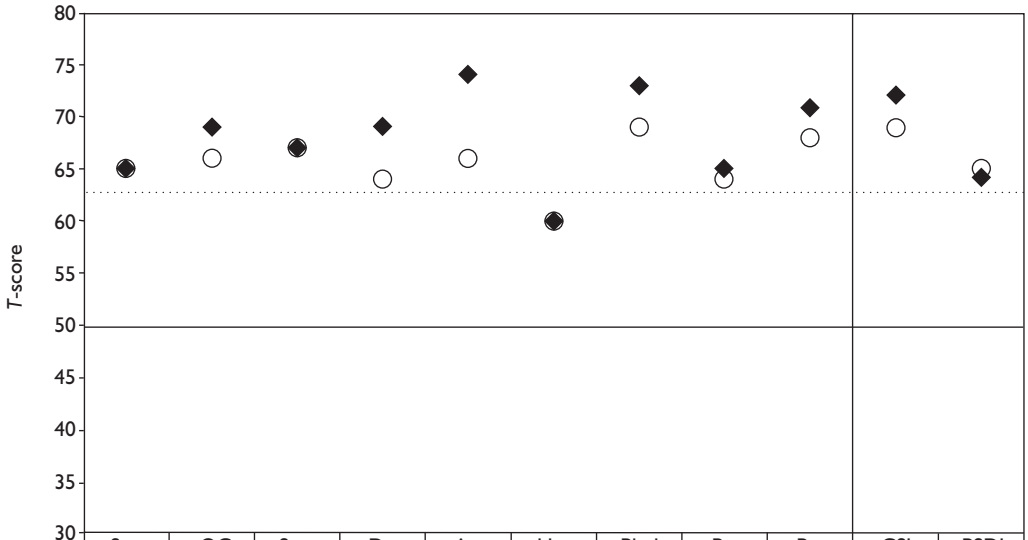
	Som	OC	Sens	Dep	Anx	Host	Phob	Par	Psy	GSI	PSDI
◆ diag *T*-score	78	76	73	78	80	67	76	68	75	80	67
raw score	1.54	2.17	1.60	1.84	2.10	1.26	1.77	1.34	1.28	1.68	2.22
● no diag *T*-score	65	62	61	62	65	59	65	57	63	65	57
raw score	0.75	0.89	0.63	0.56	0.85	0.59	0.52	0.55	0.43	0.66	1.44
△ all *T*-score	69	71	67	71	75	63	72	64	69	74	64
raw score	1.17	1.55	1.16	1.24	1.51	0.95	1.19	0.95	0.89	1.20	1.85

FIGURE 5 *Long-term follow-up scores on the BSI subscales as compared with adult non-patient norms for male participants (*n *= 148), and also by diagnostic status (any diagnosis* n *= 76; no diagnosis* n *= 59) for Trials 1–8.* Anx, Anxiety; Dep, Depression; GSI, General Severity Index; Host, Hostility; OC, Obsessive-compulsive; Par, Paranoid ideation; Phob, Phobic anxiety; PSDI, Positive Symptom Distress Index; Psy, Psychoticism; Sens, Interpersonal sensitivity; Som, Somatisation.



	Som	OC	Sens	Dep	Anx	Host	Phob	Par	Psy	GSI	PSDI
◆ diag *T*-score	70	70	69	68	71	67	74	67	72	71	68
raw score	1.91	2.10	1.92	1.74	2.43	1.31	2.29	1.40	1.26	1.84	2.39
● no diag *T*-score	59	61	61	59	60	56	61	56	59	60	54
raw score	0.66	0.95	0.80	0.57	0.87	0.48	0.57	0.48	0.36	0.64	1.41
△ all *T*-score	65	65	65	64	67	63	66	62	65	67	63
raw score	1.24	1.50	1.35	1.16	1.61	0.87	1.40	0.93	0.82	1.22	1.87

FIGURE 6 *Long-term follow-up scores on the BSI subscales as compared with adult non-patient norms for female participants (*n *= 235), and also by diagnostic status (any diagnosis* n *= 106; no diagnosis* n *= 100) for Trials 1–8.* Abbreviations as in *Figure 5*.



	Som	OC	Sens	Dep	Anx	Host	Phob	Par	Psy	GSI	PSDI
◆ males *T*-score	65	69	67	69	74	60	73	65	71	72	64
raw score	0.83	1.31	1.19	1.07	1.21	0.61	1.26	1.12	1.01	1.06	1.89
○ females *T*-score	65	66	67	64	66	60	69	64	68	69	65
raw score	1.17	1.49	1.61	1.22	1.39	0.68	1.72	1.03	1.06	1.26	2.03

FIGURE 7 *Long-term follow-up scores on the BSI subscales as compared with adult non-patient norms for male (*n *= 61) and female (*n *= 28) participants for Trials 9 and 10.* Abbreviations as in *Figure 5*.

higher for female than male patients in the current study, when compared with the normative data, male patients appear to have fared worse at long-term follow-up than female patients.

Psychosis studies (Trials 9 and 10)

The profile of long-term follow-up participants on the BSI subscales, in comparison with the population norms for adult non-patient data,[135] for the two psychosis studies are shown in *Figure 7*. With the exception of the hostility subscale, all scores are above the caseness cut-off points for both male and female patients. It can be also seen that, although female patients generally had higher (i.e. worse) raw scores on the BSI subscales than male patients, when compared with population norms, the male patients had generally fared worse. The exceptions to this were the somatisation, sensitisation and hostility subscales, where male and female patients had identical *T*-scores, and the PSDI, where females had scores which were slightly worse.

Comparison of anxiety disorder and psychosis studies

There were some differences in raw scores on the BSI-GSI between the anxiety and psychosis patients at long-term follow-up. For the whole sample, the psychosis patients had lower (i.e. better) scores on the subscales representing somatisation (95% CI for difference 0.06 to 0.49, $p = 0.015$), anxiety (95% CI for difference 0.05 to 0.56, $p = 0.021$) and hostility (95% CI for difference 0.09 to 0.45, $p = 0.004$). When examined separately for male and female patients, only the somatisation (95% CI for difference 0.09 to 0.61, $p = 0.015$) and hostility scales (95% CI for difference 0.11 to 0.59, $p = 0.007$) for male patients reached significance, with the psychosis patients having lower scores in both instances.

User perspective

Patient perspectives of overall improvement

Patients were asked to rate how much they had improved (or deteriorated) on a CGI scale (1–7), comparing how they felt at long-term follow-up with respect to how they felt at the start of the original trial.

Anxiety disorder studies (Trials 1–8)

Figure 8 shows a bar chart of patients' responses for the eight anxiety disorder studies.

It can be seen from *Figure 8* that most of the anxiety patients (80%) believed they had improved

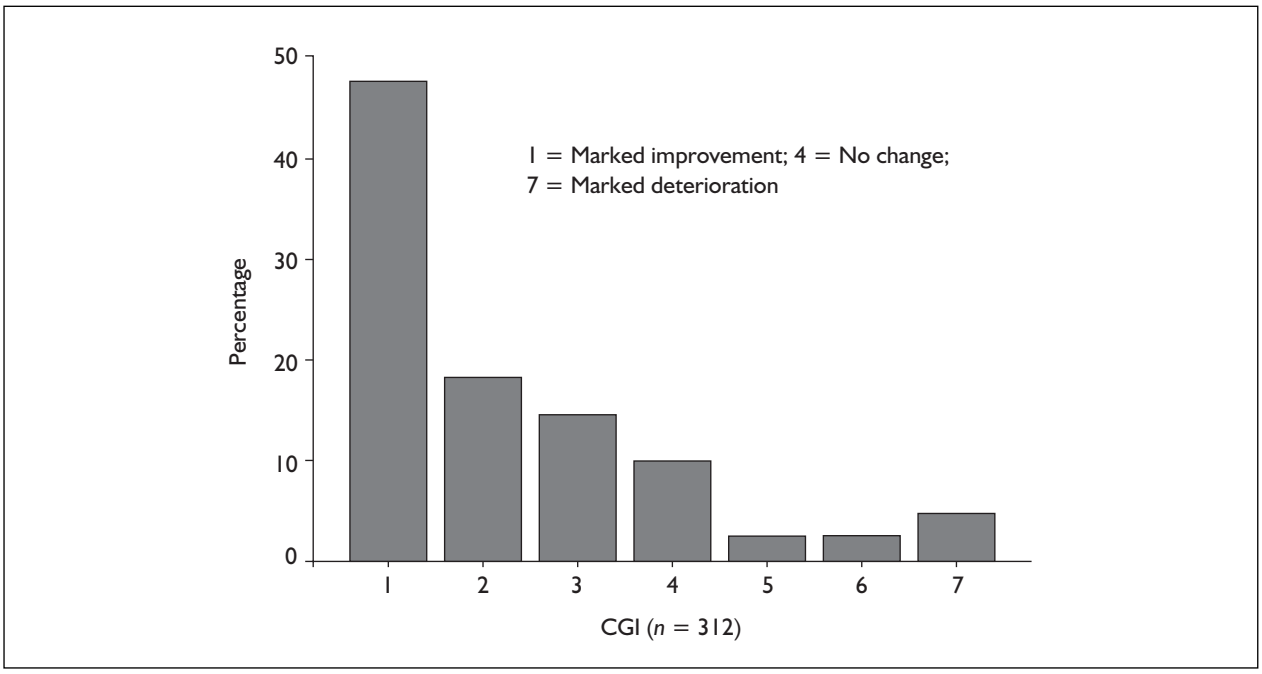


FIGURE 8 *Long-term follow-up scores of CGI (patients' view) for Trials 1–8*

since the time of the trial; 10% reported no change and a further 10% reported that they felt worse. It would seem that, even though 52% of the sample still had at least one clinical diagnosis at long-term follow-up, the majority of the anxiety patients did believe they had shown some improvement since the original trial. All but two of the patients (99%) who no longer had a clinical diagnosis reported being improved (and 85% markedly so); whereas 63% of those who still had at least one clinical diagnosis reported some improvement, 18% reported being worse and 19% reported no difference. Only 13% of those who retained a clinical diagnosis at long-term follow-up reported a marked improvement.

Psychosis studies (Trials 9 and 10)

Figure 9 shows the psychosis patients' view of how they had changed with regard to their illness since the original trial. As with the anxiety disorder patients, the majority of the psychosis patients (81%) did believe they had shown some improvement since the original trial. However, only 27% reported this as a marked improvement.

Comparison of anxiety disorder and psychosis studies

A χ^2 test by CGI category between the anxiety and psychosis patients was significant, with the most noticeable difference between the two groups being that more of the anxiety patients said they were markedly better (47 versus 27%) [$\chi^2(7) = 18.6$, $p = 0.002$].

Quality of life using SF-36

Long-term follow-up participants who either attended interview or returned questionnaire data completed the SF-36 (version II) health status measure. Scores on the eight subscales of the SF-36 were then converted to *z*-scores (mean = 0, SD = 1) in comparison with the UK normative data published by Jenkinson and colleagues.[99] This takes into account both gender and age, according to the age bands 18–24, 25–34, 35–44, 45–54 and 55–64 years. There were a few patients in the current study who were aged 65 years or older at the time of the long-term follow-up; these have been converted to *z*-scores based on the 55–64 years category. Although the SF-36 physical and mental component scores are already standardised [compared with a mean of 50 (SD 10)], long-term follow-up scores were also converted to *z*-scores based on the normative data accounting for age and gender, as this provides a more accurate picture, particularly as SF-36 scores vary with age.

Anxiety disorder studies (Trials 1–8)

The *z*-scores for the whole sample of anxiety disorder patients are represented graphically in *Figure 10*, and are also shown broken down by presence or absence of a clinical diagnosis at long-term follow-up.

It can be seen from *Figure 10* that quality of life scores on all SF-36 subscales were lower (i.e. worse) than the mean of the normative sample, including

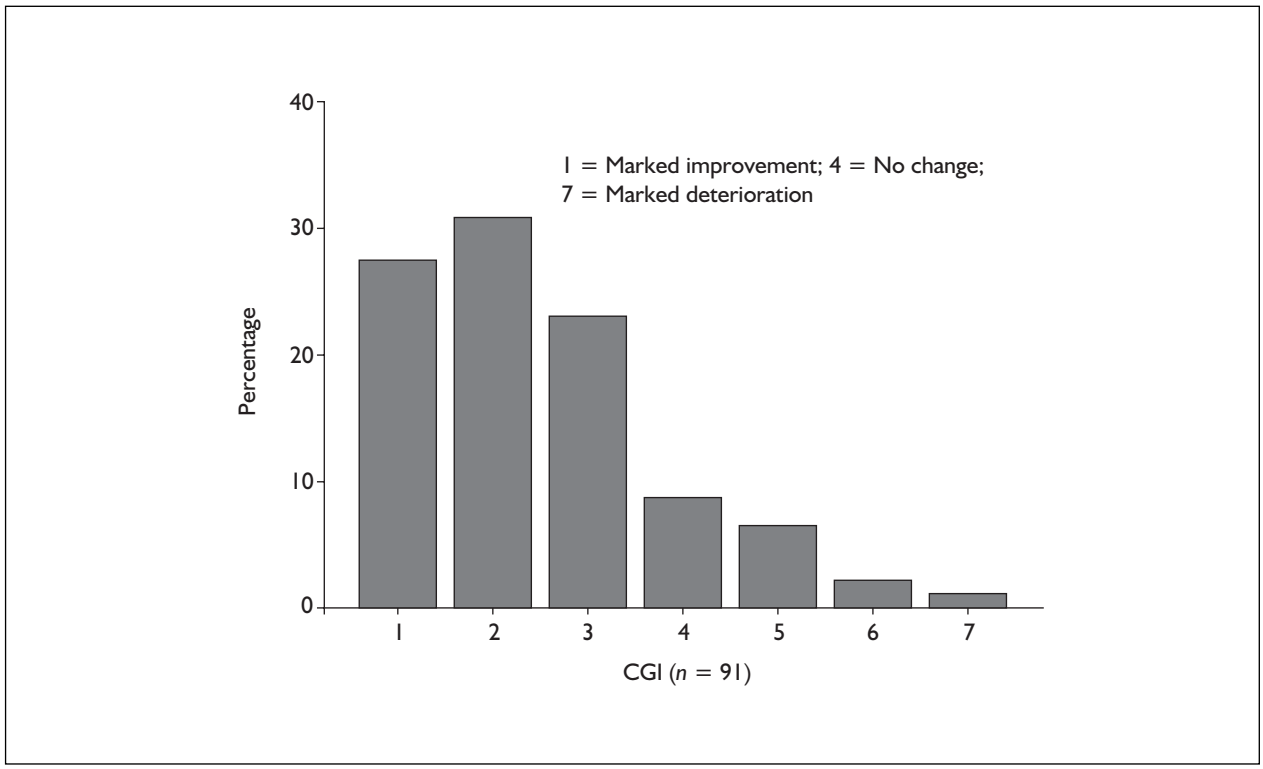


FIGURE 9 *Long-term follow-up scores of clinical global improvement (patients' view) for Trials 9 and 10*

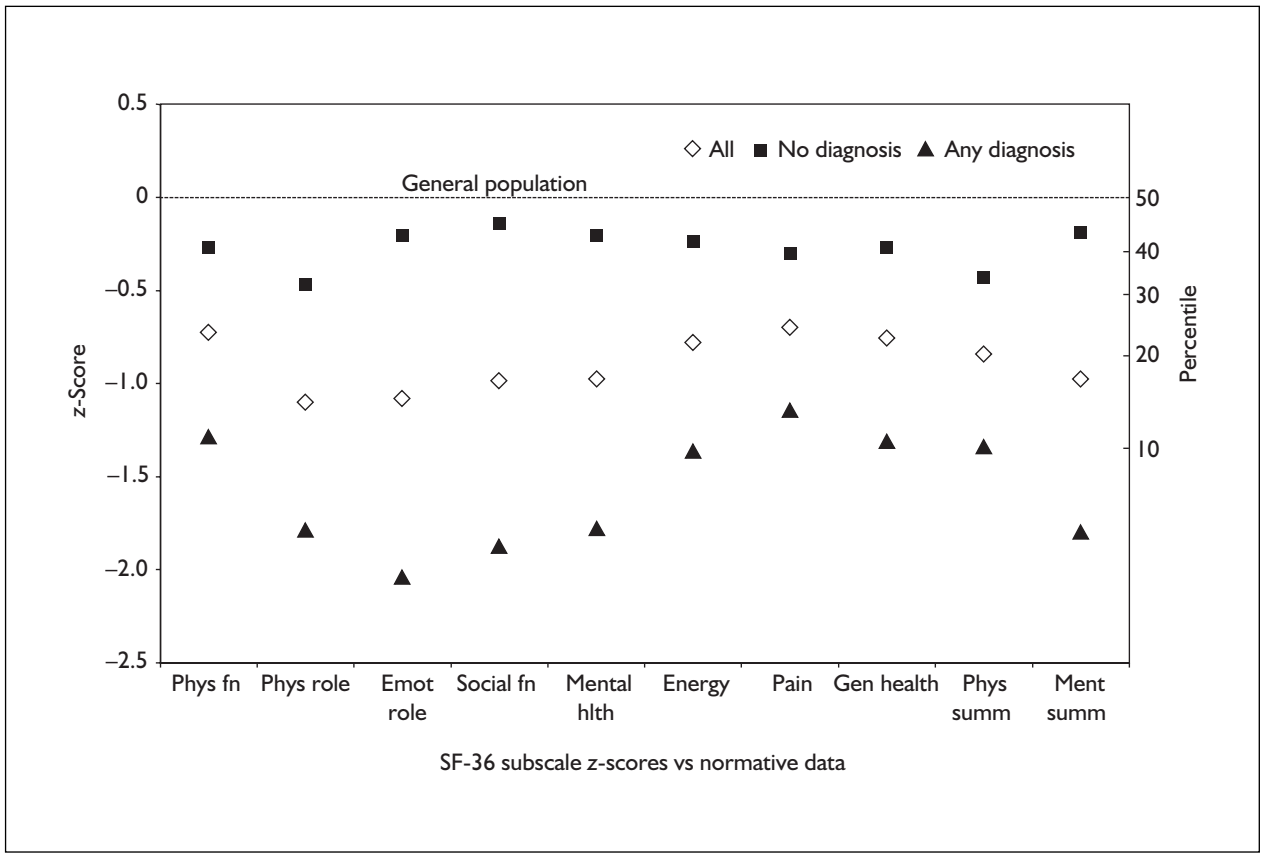


FIGURE 10 *SF-36 subscale scores for long-term follow-up participants, broken down by diagnostic status for Trials 1–8*

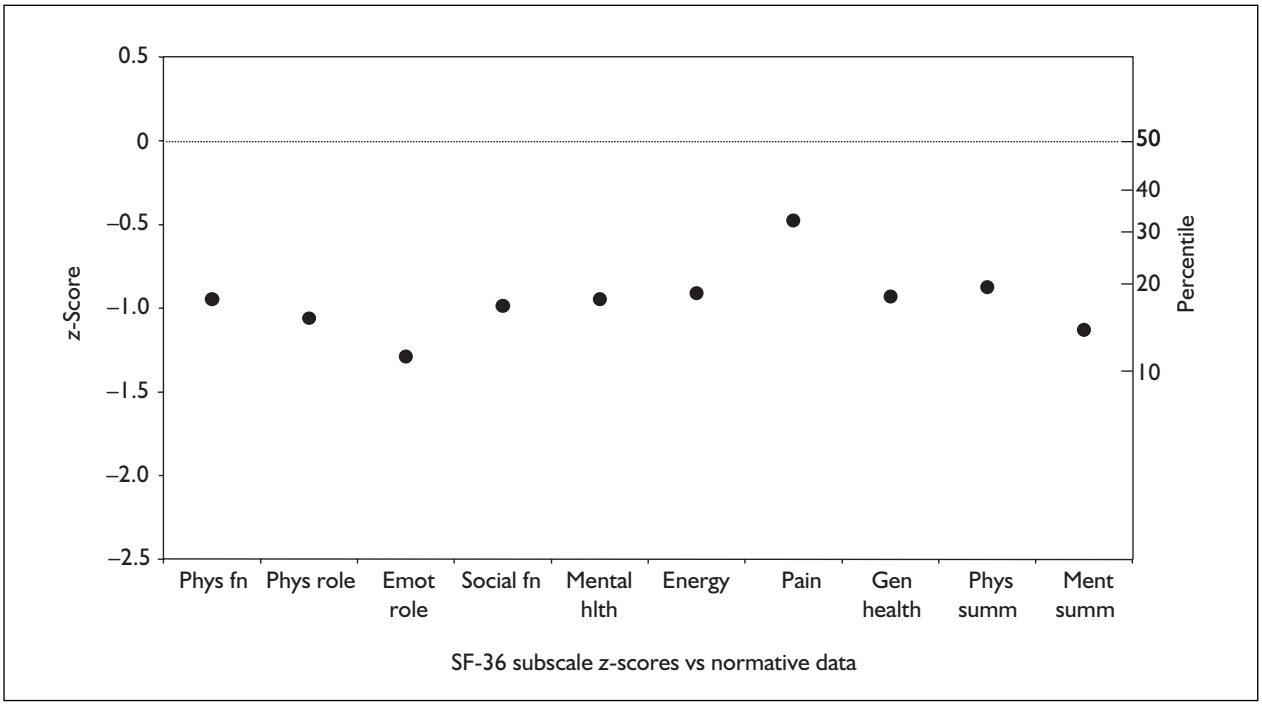


FIGURE 11 *SF-36 subscale scores for long-term follow-up for Trials 9 and 10*

for those patients who no longer had a clinical diagnosis. This applied to measures of both physical and mental well-being. Overall, mean scores for long-term follow-up participants were equivalent to those scored by the lowest 25% of the general population. For patients who still had a clinical diagnosis at long-term follow-up, all scores except pain fell in the range of the lowest (i.e. worst) scoring 10% of the general population, and the emotional role scale corresponded to the lowest 2.1%. In contrast, patients who no longer had a clinical diagnosis had mean scores which fell into the range scored by the highest 68% of the general population on all subscales.

Psychosis studies (Trials 9 and 10)

Long-term follow-up *z*-scores on the SF-36 for the two psychosis studies are shown in *Figure 11*. It can be seen that scores on all SF-36 subscales were lower (i.e. worse) for the psychosis patients than the mean of the normative sample. This applied to measures of both physical and mental well-being. With the exception of the pain subscale, mean scores for long-term follow-up participants were equivalent to those scored by the lowest 19% of the general population.

SF-36 scores by original diagnosis

Mean *z*-scores on the SF-36 are also shown by type of original diagnosis in *Figure 12*. There were no significant differences on any of the scales between the patients originally treated for GAD, panic disorder and psychosis, with mean scores for all groups in the worst 27% of the population on all scales, with one exception of the pain subscale for the psychosis patients (which fell into the worst 32%).

The PTSD patients had the worst scores overall, and these were significantly worse than the GAD and panic disorder patients on the physical role, emotional role, social functioning, pain and general health subscales, and also on both the physical and mental summary components ($0.002 < p < 0.024$). The PTSD group also scored significantly worse than the psychosis patients on the physical role, social functioning and pain subscales ($0.003 < p < 0.043$), and approached significance on the physical summary component ($p = 0.058$). The PTSD patients had long-term follow-up SF-36 scores within the worst 14% of the population on all subscales, and the worst 4% for physical and emotional role.

Patient attitudes to treatment received

Towards the end of the semi-structured interview, data were collected on responses to a number of open-ended questions regarding the patients' view of the treatment received and their progress since the original trial. These responses were categorised according to the criteria described in the section 'Composite measures' (p. 29). Patients

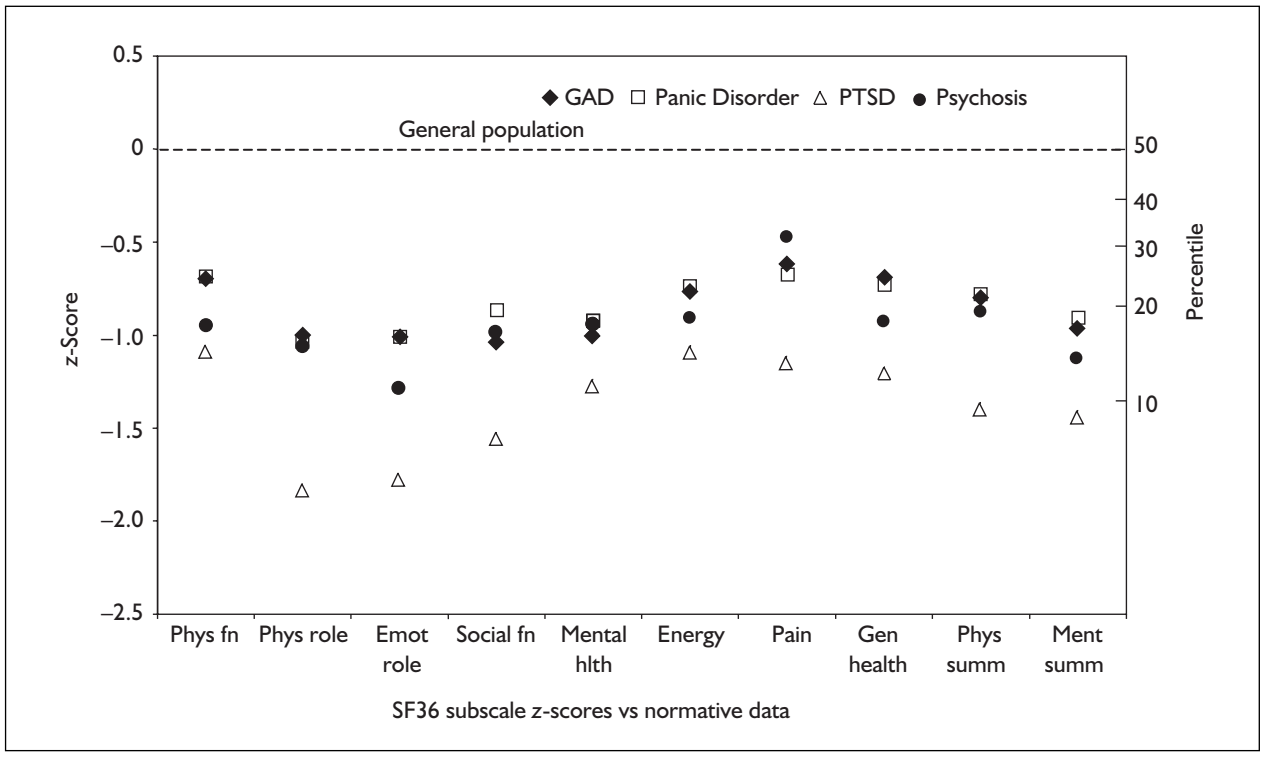


FIGURE 12 *SF-36 subscale scores for long-term follow-up participants, broken down by diagnostic category for which patients were originally treated*

completing questionnaire data only also completed a questionnaire version of these items.

Anxiety disorder studies (Trials 1–8)

Data in this section are based on $n = 384$. The majority of patients (68%) reported a fairly good memory of the original trial treatment, 24% were able to remember at least some details of the treatment (e.g. the therapist) and only 9% were unable to recall any aspects of the treatment or appeared to misattribute it to another source of treatment. Patients with a clinical diagnosis at long-term follow-up had a poorer memory of treatment than those with no diagnosis [$\chi^2(2) = 21.4$, $p < 0.001$; example: 59% versus 81% reporting a good memory of treatment].

Patients were asked how they were with regard to anxiety compared with before the original trial treatment. Overall, 81% of patients reported being better (32% much better, 35% much better although they still had periods of marked anxiety, 15% slightly better), 10% reported no change and 9% reported being worse. These figures (with a larger response rate of +42) are very similar to those reported in the section 'Jacobson method applied to primary outcome measures' (p. 45) on the CGI. Not surprisingly, patients with a clinical diagnosis were significantly less likely to report improvement than those with no diagnosis [$\chi^2(4) = 178.9$, $p < 0.001$; example: 65% versus 2% reporting being much better (with no marked anxiety)].

Overall, nearly one-fifth (19%) of patients reported having almost constant treatment for anxiety (either medication or psychological) in the interim period since the original trial, 24% reported having a moderate amount (over one or more extended periods totalling more than 1 year, but not continuous), 21% reported having some treatment in the interim period (not exceeding 1 year in total) and 36% reported having no interim treatment since the original trial. Patients who retained a clinical diagnosis at long-term follow-up were significantly more likely to have had more interim treatment than patients who no longer had a diagnosis [$\chi^2(3) = 69.9$, $p < 0.001$; example: 35% versus 4% reporting almost constant interim treatment].

Nearly one-quarter of patients (24%) reported that they continued regularly to use techniques which they had learnt during the original trial for managing their anxiety, 27% reported sometimes using techniques they had learnt, 10% reported occasionally using some technique learnt from the original trial and 39% reported never using anything they had learnt. The last category included some individuals who felt they had learnt

nothing and some who considered themselves so much better that they had no need to practise any anxiety management techniques. Once again, those who still had a diagnosis at long-term follow-up were less likely to employ any techniques learnt [($\chi^2(3) = 20.1$, $p < 0.001$; example: 51% versus 31% reporting never using any techniques].

A subset of patients attending interview ($n = 256$) were asked their reasons for agreeing to take part at long-term follow-up. Nearly three-quarters (71%) of these participants said they had agreed to take part at long-term follow-up because they wanted to help, 5% felt they ought to take part, 12% said they attended because they wanted further help for themselves and 13% said they came because they wanted someone to talk to about their problems. Not surprisingly, patients who still had a clinical diagnosis were more likely to state that they came because they wanted help or to talk to someone, and less likely because they wanted to help, than patients who no longer had a clinical diagnosis [$\chi^2(3) = 32.1$, $p < 0.001$; example: 21% versus 4% reporting they came because they wanted help].

Patients who reported being better with regard to anxiety were also asked why they thought they had improved. A number of patients gave more than one reason, in which case only their first answer is reported here. Data are available on 273 patients. Well over half (61%) of patients asked gave the trial treatment as their main reason for their improvement, 16% put it down to either time or a change in circumstances (such as a change in employment), 10% attributed their improvement to either their self-determination or the help of others – usually close family (this excludes psychological treatment), 8% said medication was responsible and only 5% quoted later psychological treatment as a main reason for any improvement in anxiety. Although slightly fewer patients with a clinical diagnosis at long-term follow-up reported the trial treatment as the main reason for any improvement (55 versus 66% of those with no diagnosis), the differences in reasons for improvement given by the two groups were not significant (ns) [$\chi^2(4) = 3.6$, ns].

Overall, patients viewed the trial treatment as being helpful to them over the long term [mean = 4.6 (SD 3.1) on a scale of 0–8 where 4 = moderately helpful] and were fairly hopeful of coping with any future problems [mean = 5.2 (SD 2.3) on a scale of 0–8 where 4 = moderately hopeful]. Patients with a clinical diagnosis at long-term follow-up viewed the trial treatment as having been less helpful than patients with no diagnosis (3.3 versus 6.0; 95% CI for difference 2.0 to 3.2, $p < 0.001$) and were also less hopeful of coping with future problems (3.9 versus 6.6; 95% CI for difference 2.3 to 3.1, $p < 0.001$).

Psychosis studies (Trials 9 and 10)

Data in this section are based on a maximum $n = 92$. Only 44% of the psychosis patients reported having a good memory of the trial treatment, 37% were able to recall some element of the treatment and 20% had little or no memory of treatment. Almost all (93%) of the psychosis patients reported receiving virtually continuous treatment (either medication or psychological) since the original trial, and only 1% reported having had no interim treatment. Over two-thirds of patients reported being much better now than at the time of the trial (much better 27%, much better but with periods of being markedly unwell 45%) and a further 16% reported being slightly better. Only 9% reported being unchanged with regard to their illness since the original trial, and 3% reported being worse.

Although the psychosis patients were asked whether they continued to use anything learnt during the trial, the fact that one of the treatment conditions was TAU appears to have led to some confusion in patient responses (e.g. with patients stating that contact with a community psychiatric nurse or psychiatrist continues to be helpful), so results are not reported here.

Patients in the psychosis studies generally reported the trial treatment as being helpful over the long-term [mean = 5.4 (SD 2.5), where 4 = moderately helpful and 8 = very helpful] and were also fairly hopeful of coping with any future problems [mean = 5.5 (SD 2.1), where 4 = moderately hopeful and 8 = very hopeful].

Comparison of anxiety disorder and psychosis studies

Patients in the anxiety disorder studies were likely to have better recall of the trial treatment than the psychosis patients [$\chi^2(2) = 19.6$, $p < 0.001$] and also reported significantly less interim treatment [$\chi^2(3) = 178.8$, $p < 0.001$]. There were no differences between the two groups of patients with regard to how they reported being at long-term follow-up compared with before the trial treatment, or with regard to how hopeful they were of coping in the future. The psychosis patients did report that they had found the trial treatment more helpful than did the anxiety patients (95% CI for difference 0.2 to 1.5,

$p = 0.020$). However, it is possible that this was attributable to the medication received by the psychosis patients and not to any active treatment.

Summary of long-term outcome

Anxiety disorder studies (Trials 1–8)

Overall, just over half (52%) of patients returning at long-term follow-up (2–14 years after the original trial) still had at least one clinical diagnosis of an anxiety or depressive disorder, as assessed via a structured clinical diagnostic tool. Although the majority of these still had their original diagnosis, around one-third of those with at least one diagnosis no longer met criteria for the disorder for which they were originally treated, but instead had one or more other diagnoses. Levels of co-morbidity were high, with patients with a diagnosis having a mean of 1.7 additional disorders. In addition, over two-thirds of patients who did not meet diagnostic criteria still had some symptoms (43% with mild symptoms and 28% just failing to meet diagnostic levels). Only 7% of all patients had no symptoms of anxiety or depression at long-term follow-up. The proportion of patients achieving clinically significant change on the main outcome measures varied from around 20% to just over 60%, depending on the measure used. Despite these relatively low levels of recovery, 80% of all patients did believe they had improved at long-term follow-up, and only 10% felt they were worse than before the original trial. Comparison with normative data on measures of overall symptomatology and health status showed that even those with no clinical diagnosis at long-term follow-up had scores which were poorer than population means, and the whole sample had means which fell into the worst 12% of the population for symptomatology and the worst 24% for measures of physical health status. The results indicate that, although many patients are reported as faring well immediately after treatment for anxiety disorders, a relatively small number seem to maintain good levels of recovery over the longer term, and a proportion may experience a chronic, and often fluctuating, course of illness throughout their lifetime.

Psychosis studies (Trials 9 and 10)

The mean CGS of the psychosis patients at long-term follow-up was exactly 4.0, which is equivalent to having symptoms which are 'definitely disabling'. Only 10% of the sample had very mild or no symptoms. The proportion of patients achieving clinically significant improvement (i.e. 25% or more) on the PANSS total score was also only 10%. Further, only 12% had achieved Jacobson criteria (c) with reliable change on the BSI-GSI at long-term follow-up. Nonetheless, the majority of psychosis patients (81%) did believe they had made some improvement since the original trial, and they also were generally hopeful of coping with future problems. Their scores on the CGS and CGI were, however, significantly worse than the anxiety disorder patients.

Comparison with normative data on measures of overall symptomatology and health status showed that the psychosis patients had scores which were consistently poorer than normal population means. However, scores on the measures of symptomatology for the psychosis patients were in most instances significantly better than the PTSD patients, and on some subscales of the BSI they were also significantly better than the anxiety disorder patients as a whole. The results indicate that, although this group of psychosis patients had not done particularly well over the long term and were worse than the anxiety disorder patients with regard to clinical status, their levels of overall symptomatology were not dissimilar from those of patients with common anxiety disorders.

Use of healthcare resources

Anxiety disorder studies (Trials 1–8)

Of the 396 individuals who participated in the long-term follow-up study, 366 consented to access to their case notes (representing 92.4% of the follow-up population and 42.5% of the original study populations of 861). Of these, five individuals appeared in more than one trial and so analysis is presented for 361 cases. Cost data were therefore unavailable for 30 individuals who were part of the follow-up study. The balance of these 30 individuals across CBT and non-CBT based interventions reflected the balance of CBT across the whole population of 861 and was non-significant.

Figure 13 displays a histogram of total cost per patient over the 4-year period for which resource use data were collected. These costs ranged from £39.62 to £43,908.92 on a per patient basis, with a mean of £2369.13.

Clearly these resource use data are highly skewed. These total costs are obviously composite variables with a number of component parts, as outlined in the section 'Economic evaluation methods' in Chapter 2.

Table 22 shows the contribution of each type of resource use to cost over the two periods. This

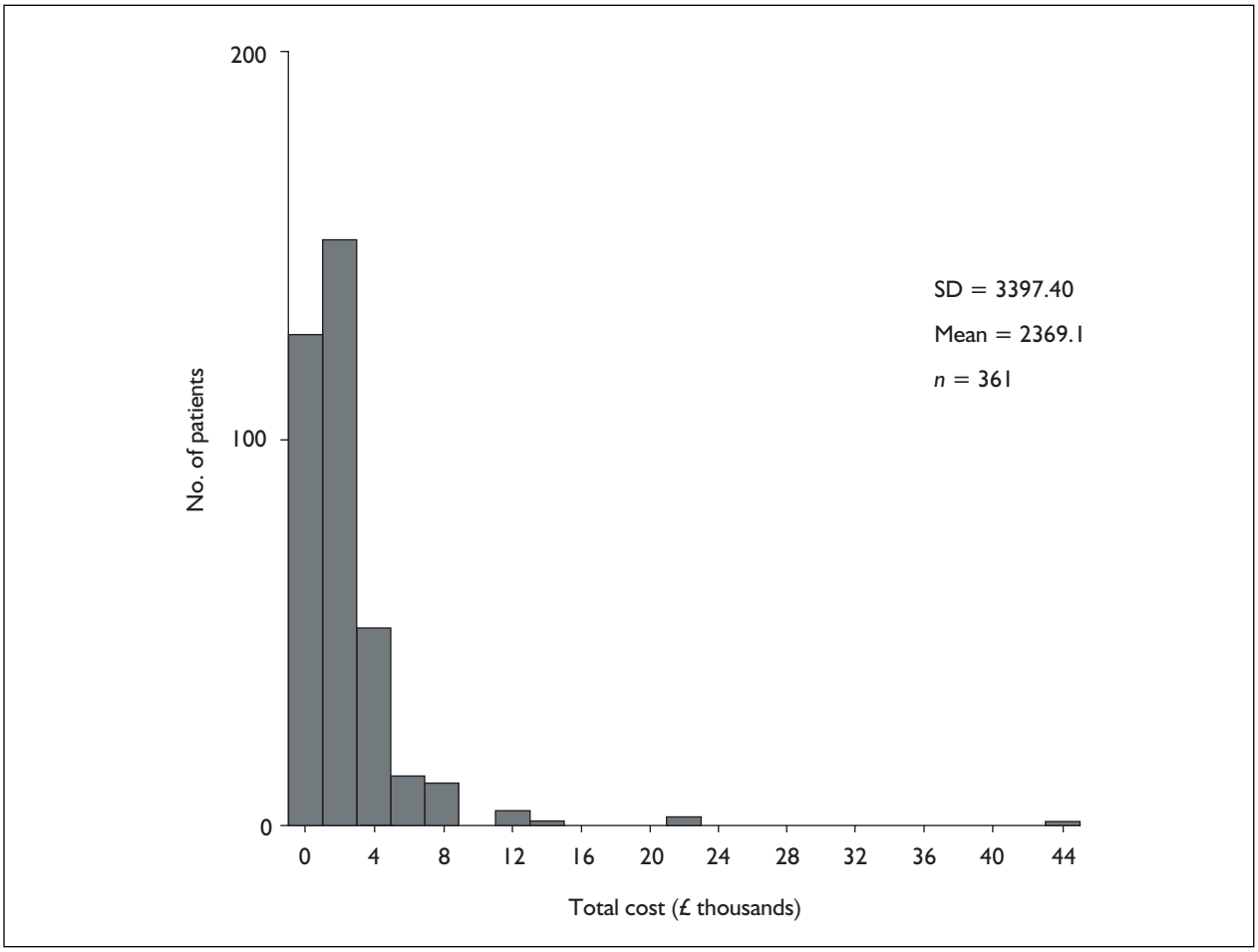


FIGURE 13 *Total cost per patient over four years (histogram) for Trials 1–8*

table also demonstrates a 40% real increase in the value of resources utilised by these patient groups over the two periods. Clearly, this cannot be attributed solely to an increase in consumption by this patient group, and is likely to be due to a more general increase in health interventions (either in a greater number of interventions or the use of more costly interventions or both). The majority of this increase (over 52%) can be attributed to a single category of resource use – prescribing (note that all costs were calculated using the same base year and any changes cannot therefore be attributed to inflation). This table also demonstrates some key differences by diagnostic group, with particularly high levels of resource use by the PTSD group.

Figure 14 is a graphical representation of the aggregate data in *Table 22* and shows the important contribution of prescribing, referrals and inpatient stays to total costs in both periods.

Although there has clearly been a significant increase in mean costs over the two periods, this does mask significant variation at an individual level. *Figure 15* shows the difference in cost for each individual, where a positive level indicates an increase in cost over the two periods. The mean change in cost is also shown as a fixed line on the chart. A total of 14 individuals experienced a **reduction** in resource use, equivalent to >£2000. In comparison, a total of 39 individuals experienced a **rise** in resource use of a similar level.

Psychosis studies (Trials 9 and 10)

Resource use data from case note review were available for 94 patients (35 in Trial 9 and 59 in Trial 10).

Figure 16 displays a histogram of total cost per patient over the 4-year period for which resource data were collected. These costs ranged from £664.07 to £251,046.47. Clearly these resource use

TABLE 22 *Mean cost per patient (£) by category, time period and diagnostic group for Trials 1–8 (bias corrected percentile based bootstrapped 95% CIs based on 1000 replications)*[a]

	Period	
Category	**Pre-trial**	**Pre-follow-up**
GP telephone	**3 (2 to 4)**	**5 (3 to 8)**
GAD	1 (0 to 2)	5 (2 to 14)
Panic	3 (2 to 5)	5 (3 to 7)
PTSD	8 (4 to 15)	7 (3 to 13)
GP out of hours	**7 (5 to 9)**	**20 (12 to 33)**
GAD	7 (3 to 12)	22 (7 to 54)
Panic	6 (4 to 9)	20 (11 to 37)
PTSD	10 (3 to 19)	7 (2 to 17)
Practice nurse	**8 (6 to 11)**	**14 (11 to 17)**
GAD	5 (4 to 8)	12 (9 to 17)
Panic	8 (6 to 11)	14 (11 to 19)
PTSD	18 (6 to 42)	19 (11 to 30)
GP home	**15 (6 to 34)**	**16 (7 to 33)**
GAD	7 (3 to 13)	12 (5 to 29)
Panic	10 (5 to 22)	8 (5 to 13)
PTSD	77 (4 to 287)	71 (2 to 273)
Tests	**32 (27 to 37)**	**61 (44 to 95)**
GAD	26 (19 to 35)	80 (39 to 179)
Panic	36 (29 to 44)	50 (40 to 63)
PTSD	25 (14 to 39)	56 (32 to 91)
GP attendance	**170 (160 to 181)**	**143 (132 to 154)**
GAD	180 (160 to 201)	155 (136 to 177)
Panic	165 (153 to 179)	135 (120 to 150)
PTSD	161 (132 to 192)	145 (111 to 189)
Prescribing	**171 (140 to 204)**	**375 (323 to 434)**
GAD	142 (111 to 182)	387 (310 to 480)
Panic	165 (131 to 216)	344 (284 to 427)
PTSD	316 (193 to 509)	520 (315 to 851)
Referrals	**265 (220 to 339)**	**368 (295 to 481)**
GAD	277 (221 to 358)	386 (307 to 466)
Panic	181 (148 to 221)	271 (216 to 373)
PTSD	729 (396 to 1384)	887 (361 to 1979)
Inpatient and daycase	**319 (227 to 447)**	**378 (235 to 575)**
GAD	400 (197 to 757)	627 (304 to 1068)
Panic	227 (136 to 344)	230 (125 to 421)
PTSD	571 (297 to 942)	323 (56 to 764)
Total cost	**990 (853 to 1195)**	**1379 (1156 to 1672)**
GAD	1046 (795 to 1424)	1686 (1287 to 2232)
Panic	802 (681 to 965)	1078 (913 to 1330)
PTSD	1916 (1169 to 3285)	2034 (1046 to 4043)

[a]For a fuller description of the use of bootstrapping as a method to capture levels of uncertainty in point estimates, see the section 'Cost-effectiveness analysis' (p. 95).

data are highly skewed. These total costs are obviously composite variables with a number of component parts, as outlined in the section 'Economic evaluation methods' in Chapter 2.

Table 23 shows the contribution of each type of resource use to cost over the two periods. This table suggests that costs have fallen over the two periods at an aggregate level, although these differences are non-significant. Prescribing has increased by 102% and social work input by 319%, whereas in-patient resource use has fallen by 36%.

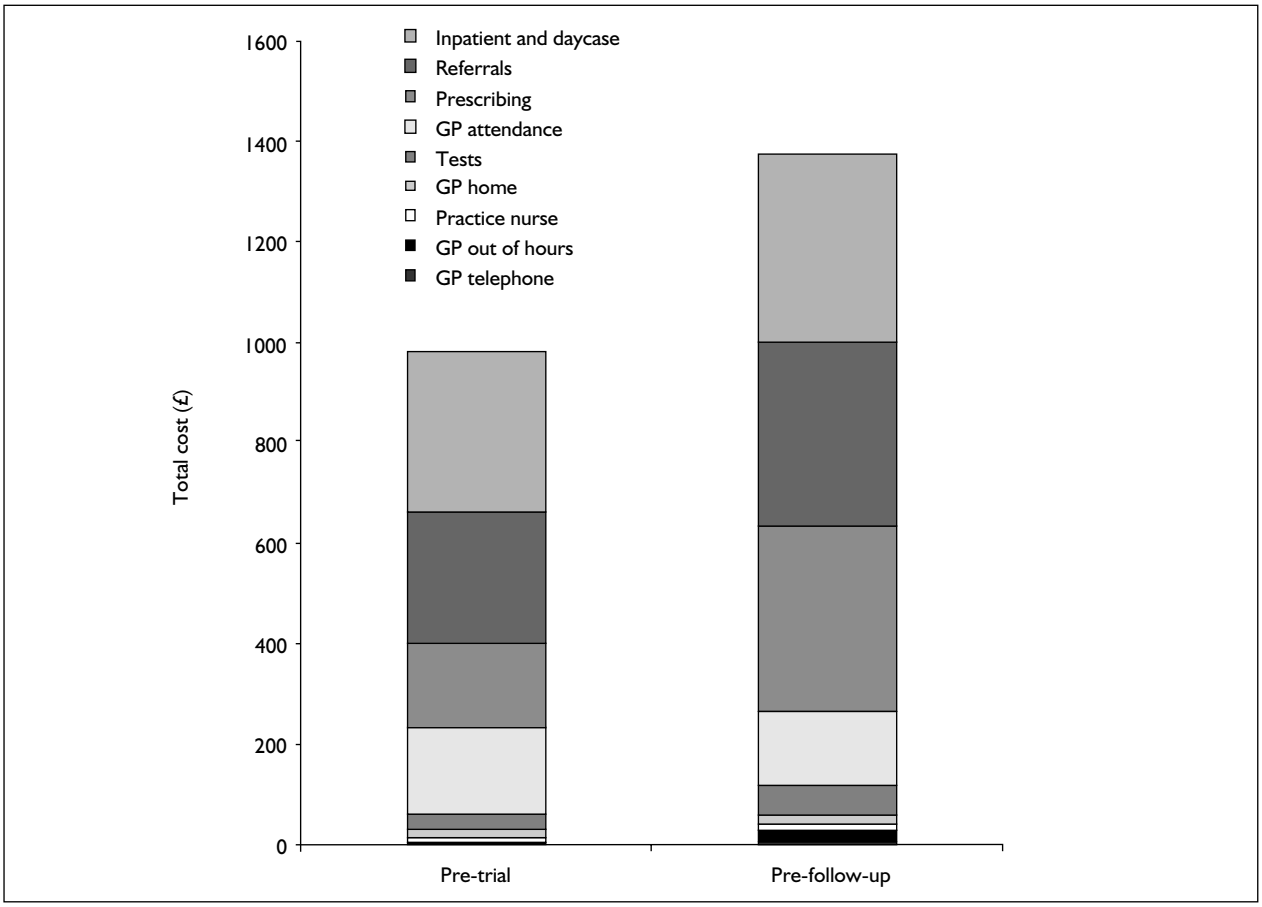


FIGURE 14 *Total cost by category and period for Trials 1–8*

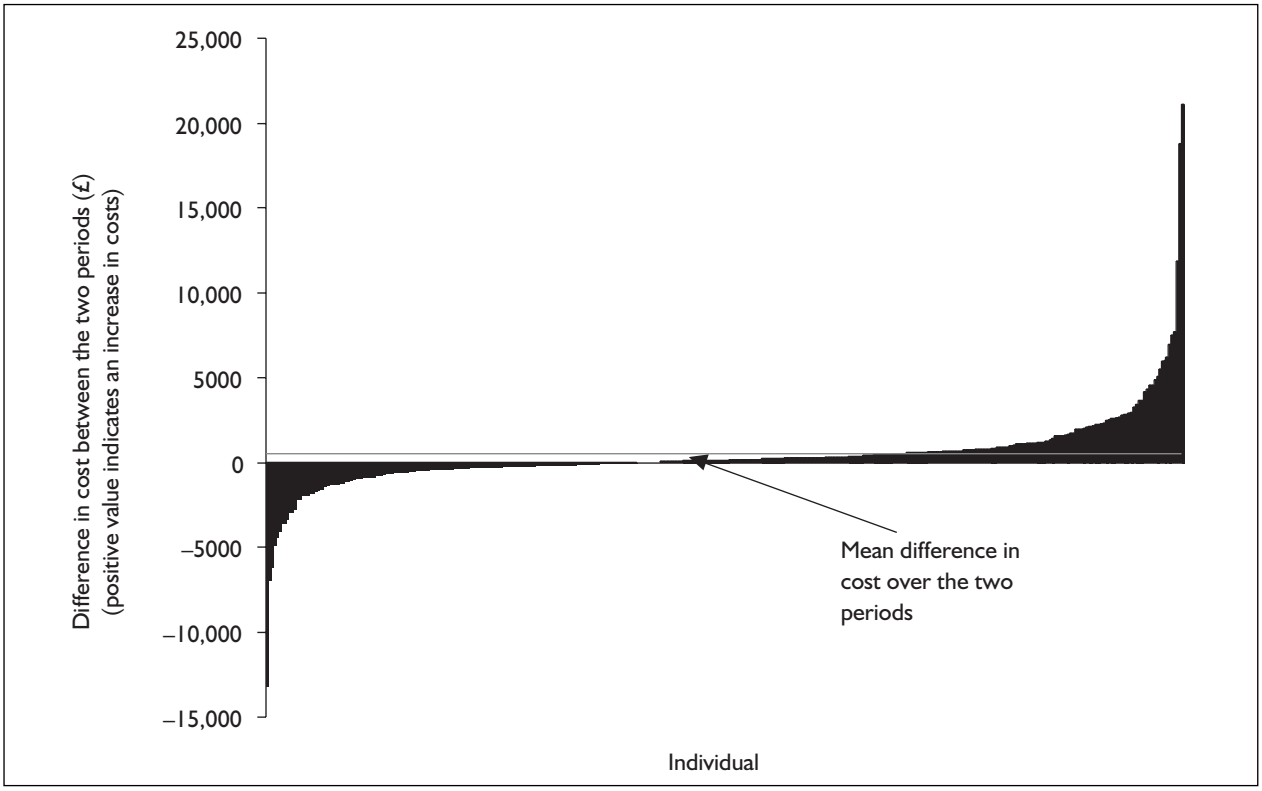


FIGURE 15 *Frequency distribution of change in resource use between the two periods for Trials 1–8*

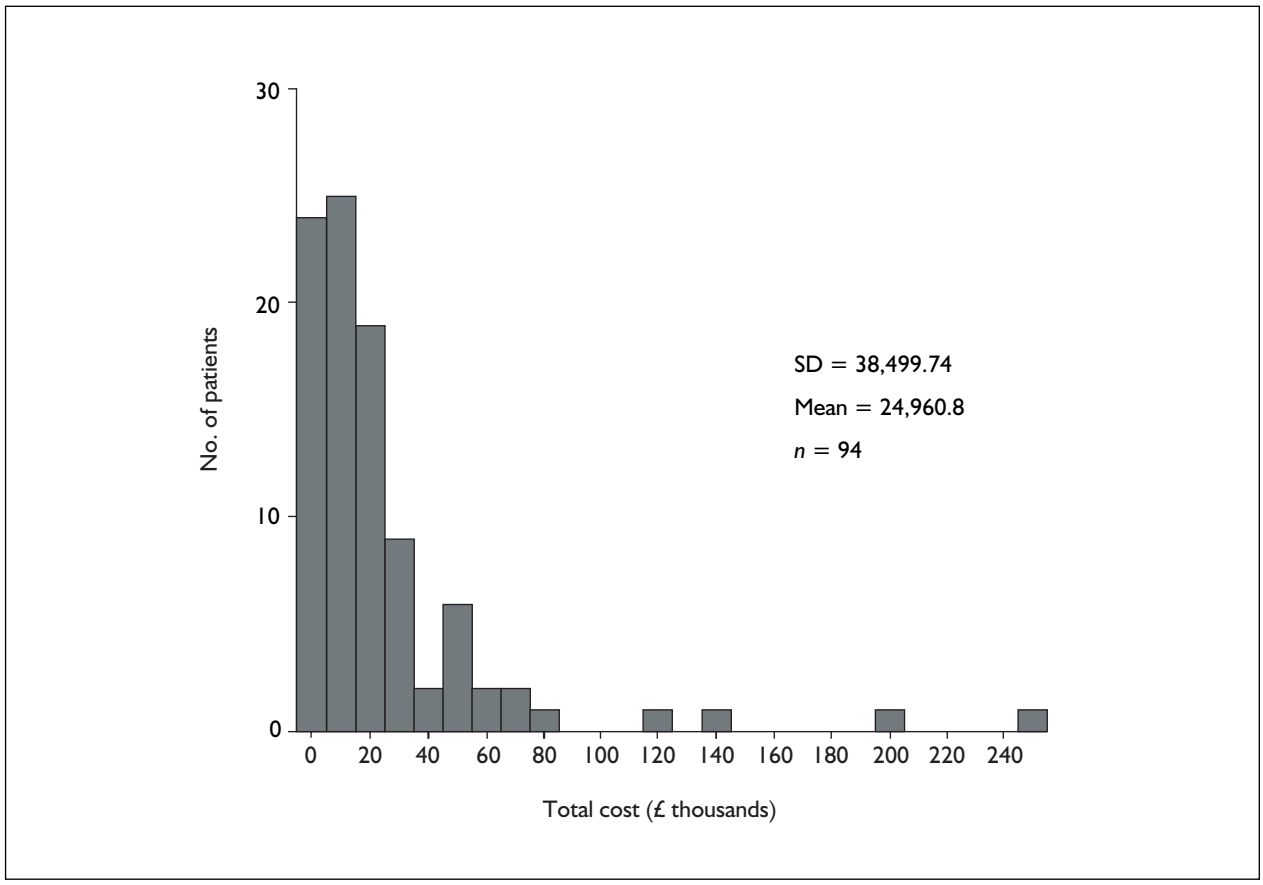


FIGURE 16 *Total cost per patient over 4 years (histogram) for Trials 9 and 10*

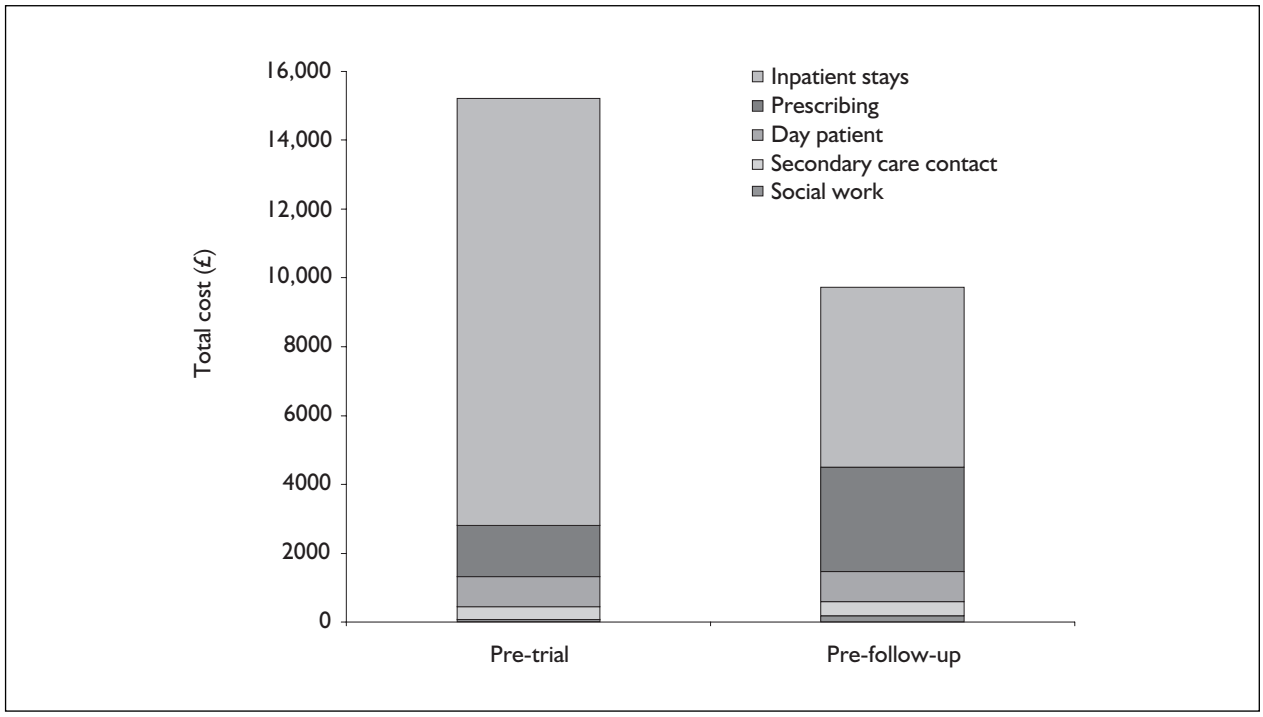


FIGURE 17 *Total cost by category and period for Trials 9 and 10*

TABLE 23 *Mean cost per patient (£) by category and period for Trials 9 and 10 (bias-corrected percentile-based bootstrapped 95% CIs based on 1000 replications)*[a]

Category	Period	
	Pre-trial	**Pre-follow-up**
Social work	37 (15 to 66)	155 (32 to 478)
Secondary care contacts	350 (296 to 412)	377 (316 to 453)
Day patient	859 (311 to 1,941)	901 (454 to 1,564)
Prescribing	1,493 (1,194 to 1,814)	3,016 (2,567 to 3,487)
Inpatient stays	12,497 (7,849 to 19,775)	5,277 (2,806 to 8,394)
Total cost	15,235 (10,417 to 22,364)	9,725 (7,129 to 13,409)

[a]For a fuller description of the use of bootstrapping as a method to capture levels of uncertainty in point estimates, see the section 'Cost-effectiveness analysis' (p. 95).

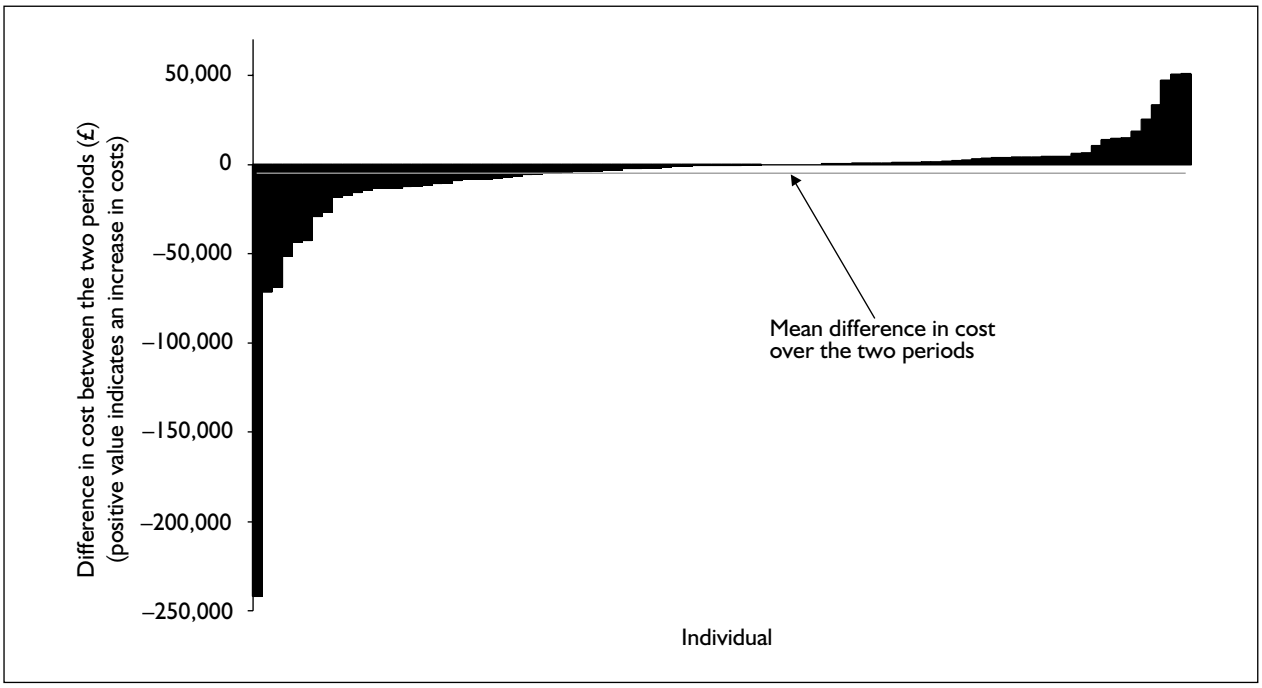


FIGURE 18 *Distribution of change in resource use between the two periods of study for Trials 9 and 10*

Figure 17 is a graphical representation of the aggregate data in *Table 23* and shows the important contribution of inpatient stays to total costs in both time periods.

Figure 18 shows the difference in cost for each individual, where a positive level indicates an increase in cost over the two periods. The mean change in cost is a reduction of £5510. A total of 56 individuals experienced a **reduction** in resource use. In comparison, a total of 38 individuals experienced a **rise** in resource use.

Chapter 7

Efficacy of CBT versus non-CBT

Introduction

Three studies are reported in this chapter. In the first section we report an analysis of CBT versus non-CBT conditions for the data aggregated across all the anxiety disorder studies (Trials 1–8). The next section gives a brief summary of a published paper[149] examining this question for Trials 1 and 2 only. The first two trials on CBT for GAD were both direct comparisons of CBT and non-CBT treatment conditions and they provided the clearest opportunity to test the hypothesis that CBT for GAD confers an advantage over the longer term in comparison with alternatives. Some evidence for the long-term influence of CBT was found and also evidence of a marked difference in outcome between the two trials. In the final section we report an analysis of the CBT versus non-CBT conditions for data aggregated across the two trials of CBT for psychosis (Trials 9 and 10).

CBT versus non-CBT for the eight anxiety disorder studies (Trials 1–8)

Introduction

All eight anxiety disorder studies reported in this section had shown favourable results for CBT treatment both at post-treatment and short-term follow-up (3–12 months) of the original trials. This section reports on the comparison between CBT and other treatments at long-term follow-up of between 2 and 14 years.

Method

The sample was assigned to CBT versus non-CBT treatment groups according to the criteria described in the section 'Treatment groups' (p. 17). The resulting groups used in the current analysis are as follows.

Non-CBT

This includes the DZ and PL groups from Trial 1; the AP treatment group from Trial 2; the FL and PL groups from Trial 3; EMDR and WL groups from Trial 6; and the delayed group from Trial 8. Most of these groups had some therapist contact over and above the assessments. The mean number of hours of therapist contact for patients in the non-CBT group was 4.79 (SD = 4.02).

CBT

This includes the CBT, CBT+DZ and CBT+PL groups from Trial 1; the CT and AMT groups from Trial 2; the CBT, CBT+FL and CBT+PL groups from Trial 3; the maximum, minimum and bibliotherapy CBT groups from Trial 4; the group CBT and individual CBT groups from Trial 5; the E+CR group from Trial 6, the brief, standard and intensive CBT groups from Trial 7; and the CBT6, CBT6-CA and CBT12 groups from Trial 8. The mean number of hours of therapist contact for patients in the CBT group was 5.64 (SD = 3.91).

Results

First, the outcome measures at long-term follow-up which were common across all studies were compared using independent *t*-tests by CBT group. The results are shown in *Table 24*.

The CBT group had more favourable scores on all outcome measures at long-term follow-up, but only the BSI-GSI, PA and NA reached significance. A MANOVA by CBT versus non-CBT treatment on these seven common outcome measures also showed a significant effect of CBT group [$F(7,303) = 2.2$, $p = 0.033$] [a similar result was obtained on an independent *t*-test on the long-term outcome factor (non-CBT mean = 0.20, CBT mean = –0.07; 95% CI for difference 0.01 to 0.52, $p = 0.038$], with those receiving CBT treatment ($n = 230$) having more favourable long-term follow-up scores on all measures than those in the non-CBT group ($n = 81$). As comparisons of pre-treatment patient characteristics (see the section 'Representativeness of participants in relation to original cohort', p. 32) had revealed some significant differences between the aggregated CBT and non-CBT groups, the MANOVA was repeated controlling for the measures where differences were found, namely pretreatment deprivation score and duration of current episode. The effect in favour of the CBT group remained significant [$F(7,295) = 2.1$, $p = 0.045$] and there were also significant effects for both deprivation score [$F(7,295) = 3.1$, $p = 0.003$] and duration of current episode [$F(7,295) = 2.1$, $p = 0.048$) {the

TABLE 24 *Means (SD) of main outcome measures at long-term follow-up by CBT group and results of independent t-tests for Trials 1–8*

	Treatment group				Independent t-test	
	Non-CBT		CBT			
Long-term follow-up	***n***	**Mean (SD)**	***n***	**Mean (SD)**	**95% CI for difference**	***p***
CGS	83	3.46 (1.84)	258	3.27 (1.82)	–0.26 to 0.64	0.408
CGI	82	2.27 (1.69)	258	2.26 (1.61)	–0.40 to 0.42	0.967
BSI-GSI	91	1.40 (0.87)	289	1.15 (0.82)	0.05 to 0.45	0.013
PA	87	27.77 (8.85)	259	29.89 (8.45)	–4.20 to –0.03	0.046
NA	87	26.66 (10.53)	259	23.29 (9.25)	1.03 to 5.71	0.009
SF-36 PC	90	39.30 (14.14)	276	41.38 (13.37)	–5.32 to 1.16	0.207
SF-36 MC	90	38.82 (11.96)	276	40.96 (12.07)	–5.01 to 0.74	0.145

TABLE 25 *Means (SD) of main outcome measures at post-treatment by CBT group and results of independent t-test, for long-term follow-up participants for Trials 1–8*

	Treatment group				Independent t-test	
	Non-CBT		CBT			
Post-treatment	***n***	**Mean (SD)**	***n***	**Mean (SD)**	**95% CI for difference**	***p***
CGS (1–7)	37	3.22 (1.7)	144	2.48 (1.3)	0.13 to 1.34	0.018
CGS (0–8)	94	2.67 (2.2)	26	4.27 (2.3)	0.64 to 2.56	0.001
CGI (1–7)	25	2.20 (1.5)	152	2.12 (1.3)	–0.48 to 0.64	0.775
CGI (0–5)[a]	18	2.72 (0.8)	32	3.25 (1.0)	–1.09 to 0.08	0.066
BSI-GSI	19	1.64 (1.14)	62	1.12 (0.9)	0.18 to 1.02	0.006
SF-36 PC[a]	6	47.53 (8.3)	28	46.76 (11.2)	–9.10 to 10.65	0.874
SF-36 MC[a]	6	31.88 (11.1)	28	44.46 (11.5)	–23.10 to 2.06	0.021

[a] High scores on this measure are favourable.

pretreatment deprivation score was also consistently related to long-term outcome in the prediction analysis [see the section 'Predictors for the eight anxiety disorder studies (Trials 1–8)', p. 83]}.

As reported in the section 'Representativeness of participants in relation to original cohort' (p. 32), there were no significant differences between long-term follow-up participants in the CBT and non-CBT groups on any of the composite pretreatment measures. However, as may be expected from the favourable results for CBT treatments in each of the original trials, long-term follow-up participants who had been in the CBT group had significantly better outcome in comparison with the non-CBT group on all of the common outcome measures which were used post-treatment (i.e. CGS, CGI, BSI-GSI, STAI-T and SF-36), with the exception of the CGI and the SF-36 PC. The results are shown in *Table 25*. Similar patterns were found on outcome measures which were not common across studies. The results indicate that the differences between the CBT and non-CBT groups were much smaller at long-term follow-up than at post-treatment.

CBT group by clinical status is shown in *Table 26*.

It can be seen from *Table 26* that there are few differences between the CBT and non-CBT groups with regard to clinical status at long-term follow-up. The only significant effect is that more patients in the CBT group have scored below the cut-off score of 0.94 (calculated according to Jacobson criteria on all data from Trials 2 and 7) on the BSI-GSI. However, as we were unable to apply Jacobson reliable change criteria on this measure (owing to the BSI not being used pretreatment in six of the eight studies), this result would appear to reflect the differences between the CBT and non-CBT groups in long-term follow-up symptom scores on the BSI-GSI, rather

TABLE 26 *Clinical status by CBT group at long-term follow-up for Trials 1–8*

	CBT		Non-CBT			
	n	%	*n*	%	$\chi^2(1)$	*p*
No clinical diagnosis	259	49.0	83	47.0	0.11	0.745
Achieved Jacobson[a]	256	44.1	83	37.3	1.18	0.277
Below BSI-GSI cut-off[b]	289	46.0	91	34.1	4.03	0.045

[a] Jacobson criterion (a) applied to HAM-A for all studies except Trial 2, which used Jacobson (c) on BSI-GSI (with reliable change) – all based on criteria used in the original trials.
[b] Cut-off [Jacobson criterion (c) without reliable change] calculated using all available pretreatment data, applied universally across all trials at long-term follow-up.

TABLE 27 *Percentage within CBT group by clinically significant change at post-treatment and long-term follow-up for Trials 1–8*

	Achieved clinically significant change post-treatment on HAM-A (%)[a]			
	CBT group (*n* = 191)		Non-CBT group (*n* = 67)	
	Yes	No	Yes	No
Achieved Jacobson at long-term follow-up[b]	38.7	9.4	29.9	10.4
Did not achieve Jacobson at long-term follow-up	26.7	25.1	37.3	22.4
	$\chi^2(1) = 17.6$, $p < 0.001$		$\chi^2(1) = 1.0$, $p = 0.322$	

[a] Percentages are within CBT group.
[b] Jacobson criterion (a) applied to HAM-A for all studies except Trial 2, which used Jacobson (c) on BSI-GSI (with reliable change) – all based on criteria used in the original trials.

than a true assessment of achieving clinically significant change from their pretreatment state.

CBT group by response to trial treatment

We have seen that overall, there was a slight effect of CBT group with regard to self-report symptoms at long-term follow-up, but there were no differences with respect to clinical status for the whole sample. Next, we examined whether there were any differences with regard to Jacobson clinically significant change at long-term follow-up for treatment completers who had achieved clinically significant change at post-treatment on the HAM-A (this measure was chosen as it was used in seven of the eight trials at post-treatment, excluding Trial 8). The results are shown in *Table 27*.

It can be seen from *Table 27* that, in the CBT group, a greater number of patients who had achieved clinically significant change (i.e. recovered) on the HAM-A post-treatment also achieved clinically significant change at long-term follow-up, and the χ^2 test within the CBT group was highly significant. A moderately large percentage of patients in both treatment groups (27% of the CBT and 37% of the non-CBT) had relapsed with regard to clinically significant change from post-treatment to long-term follow-up. In addition, very few of those who had not achieved clinically significant change post-treatment went on to achieve clinically significant change at long-term follow-up, in both treatment conditions.

In the CBT group, the odds ratio of maintaining clinical status from post-treatment to long-term follow-up was significant. Within the CBT group, patients who had achieved clinically significant change at post-treatment were 1.56 times (95% CI 1.3 to 1.9) more likely to maintain clinically significant change at long-term follow-up than those who had not recovered post-treatment. Similarly, those in the CBT group who did not achieve clinically significant change post-treatment were 2.48 times (95% CI 1.6 to 3.9) more likely to remain unrecovered at long-term follow-up than those in this group who had achieved clinically significant change post-treatment. The odds ratio for the non-CBT group was not significant.

The results suggest that patients who had responded to CBT treatment (i.e. achieved

TABLE 28 *Direct maximum likelihood estimates of the means of the outcome variables and their standard errors (SE) using all the predictor variables in the analyses for Trials 1–8*

	Non-CBT		CBT			
Long-term follow-up	**Mean**	**SE**	**Mean**	**SE**	$\chi^2(1)$	***p***
CGS	3.59	0.17	3.29	0.10	1.94	0.16
CGI	2.42	0.16	2.34	0.09	0.19	0.66
BSI-GSI	1.41	0.08	1.21	0.04	4.02	0.05
PA	27.77	0.82	29.33	0.49	2.22	0.13
NA	27.29	0.97	23.97	0.51	7.51	0.01
SF-36 PC	39.41	1.20	41.34	0.75	1.59	0.21
SF-36 MC	38.12	1.08	40.25	0.67	2.39	0.12

clinically significant change) may have been more likely to maintain clinically significant change from post-treatment to long-term follow-up than those who completed non-CBT treatment, with 59% of CBT responders and 44% of treatment responders in the non-CBT group maintaining clinically significant change over this period. However, using a log-linear model, a three-way test of association between CBT group, responder post-treatment and responder at long-term follow-up was non-significant [likelihood $\chi^2(1) = 1.6$, $p = 0.208$]. Nonetheless, it should be noted that the odds of remaining unrecovered were higher than those of maintaining recovery from post-treatment to long-term follow-up, and this also applied across the whole sample. In other words, although those who had responded to CBT had a higher chance of maintaining improvement to long-term follow-up than those who had not responded to CBT, overall those who had failed to respond to any initial treatment were very unlikely to have achieved clinically significant change at long-term follow-up.

CBT group by views of treatment variables

There was a highly significant result with regard to how much patients used anything they had learnt during treatment (e.g. relaxation techniques, breathing control, distraction techniques), with those in the CBT group being much more likely to use something than those in the non-CBT groups [$\chi^2(3) = 20.5$, $p < 0.001$; with 33% of the CBT group versus 58% of the non-CBT group reporting never using anything they learnt]. Also, for those patients who reported being better, significantly more in the CBT group gave the trial treatment as their main reason for being better (as opposed to any other reason) than did the non-CBT group [$\chi^2(1) = 7.2$, $p = 0.007$; with 66% of the CBT group versus 47% of the non-CBT group stating trial treatment]. Patients in the CBT group reported the trial treatment as being more helpful to them over the long term than did the non-CBT group [CBT group mean = 4.8 (SD 3.0) versus non-CBT mean = 4.0 (SD 3.0); 95% CI for difference 0.02 to 1.48, $p = 0.044$]. Although those in the CBT group generally had more favourable scores on all of the views of treatment variables, there were no significant differences between the CBT and non-CBT groups with regard to the amount of interim treatment received since the original trial, how much they remembered of the trial treatment, how hopeful they were of coping in the future or how they viewed themselves with regard to anxiety at long-term follow-up as compared with how they were before the trial treatment.

Missing data analysis

As there were missing data on most of the long-term outcome variables, analyses were undertaken using AMOS,[150] which provides estimates and significance tests of differences between the means of two groups on one or more variables using an estimation procedure which takes account of any information that is available on the cases whose data are missing. AMOS was used to compute the DML estimates (DML analysis was reported rather than the EM method because it permits the use of significance tests between the CBT and non-CBT groups whereas the EM analysis does not) of the means of the long-term outcome variables for both the CBT and non-CBT groups (see *Table 28*) using the information in all 21 predictor variables specified in the section 'Predictors for the eight anxiety disorder studies (Trials 1–8)' (p. 83) in the estimation process.

The effect of the missing data appears to be small as the computed means did not differ greatly from the unadjusted means already shown in *Table 24*. In both cases the CBT means indicate a better outcome than the non-CBT means on all the variables. The equality of the means was then

TABLE 29 *Results of the meta-analyses on the mean differences between the CBT and non-CBT groups on the long-term follow-up measures*

Measure	Trials 1, 2, 3, 6, 8		Trials 1, 2, 3, 8	
	Pooled standardised estimate of mean CBT – mean non-CBT	95% CI	Pooled standardised estimate of mean CBT – mean non-CBT	95% CI
CGS	–0.09	–0.57 to 0.39	–0.31	–0.85 to 0.23
CGI	0.011	–0.32 to 0.55	0.12	–0.35 to 0.58
BSI-GSI	–0.18	–0.38 to 0.03	–0.26*	–0.47 to –0.04
PA	1.72	–0.59 to 4.03	2.26	–0.21 to 4.74
NA	–3.19*	–5.89 to –0.48	–3.77*	–6.66 to –0.89
SF-36 PC	1.14	–2.38 to 4.67	–0.11	–4.15 to 3.93
SF-36 MC	3.30*	0.09 to 6.50	3.91*	0.47 to 7.35

*$p < 0.05$.

tested by constraining the CBT and non-CBT groups to have the same means on all the long-term outcome variables and seeing if the presence of the constraint resulted in a significantly poorer fit. This yielded an insignificant ($\alpha = 0.05$) combined test of all the differences [$\chi^2(7) = 10.76$, $p = 0.194$]. Nevertheless, using univariate tests, significant ($\alpha = 0.05$) CBT effects were found on the BSI-GSI and on the negative affect PANAS scale (see *Table 28*). Therefore, although the composite test was insignificant, the finding that all the differences showed that the CBT group had a better outcome than the non-CBT group and that univariate significance was achieved on two of the variables suggests that CBT did continue to have an effect, even after the effect of the missing data had been taken into account.

Meta-analyses

As the current study involved following up eight different anxiety disorder studies with eight different randomisations, we also performed a meta-analysis comparing the CBT and non-CBT groups for those studies with patients in both groups. The meta-analyses reported below were performed using RevMan 4.2.[151] A fixed rather than a random effects model was used since the CBT studies analysed were all the trials conducted by the authors over the period under investigation rather than being a random sample from some bigger population. The meta-analyses involved combining the differences between the mean CBT and non-CBT scores from the different trials. Hence data from Trials 4, 5 and 7, which involved only CBT conditions, were not used. For each trial an exact CI for the differences between the mean CBT and non-CBT groups was computed and the pooled mean effect size estimate was calculated.

Meta-analyses were performed on the long-term outcome variables used in the pooled analyses and reported in *Table 24*. The combined effects and associated CIs appear in *Table 29*. With the exception of the CGI, the combined effect showed that the CBT group performed better than the non-CBT group although significant ($p < 0.05$) overall differences were only found for the PANAS NA and the SF-36 MC scales. The Forest plots for these effects are shown in *Figure 19*. In Forest plots, horizontal lines represent the 95% CIs and CIs which do not include zero indicate that the effect in question is significantly different from zero at the $p = 0.05$ level. The estimates of the effects appear as squares in the CIs. The pooled estimate is marked as a diamond and its width is the pooled 95% CI. The Forest plots showed small beneficial effects of CBT at long-term follow-up on the PANAS NA and the SF-36 MC, but that was not enough to yield statistical significance when a Forest plot was produced which combined the change scores on the different outcome variables used by the different studies.

Trial 6 was unique in that it used patients that were reacting badly to a traumatic event and the control group in this trial involved EMDR which included some elements of CBT and it was found that in Trial 6 the non-CBT group was better than the CBT group on all the measures analysed except the SF-36 physical component. Accordingly, the meta-analyses were re-run excluding this trial. Although the results were broadly in line with the previous analyses (see *Table 29*), the pooled BSI-GSI effect was now significant at the 5% level. This suggests that there is a small positive long-term effect of CBT in GAD and panic disorder patients on some self-report measures but they provide no

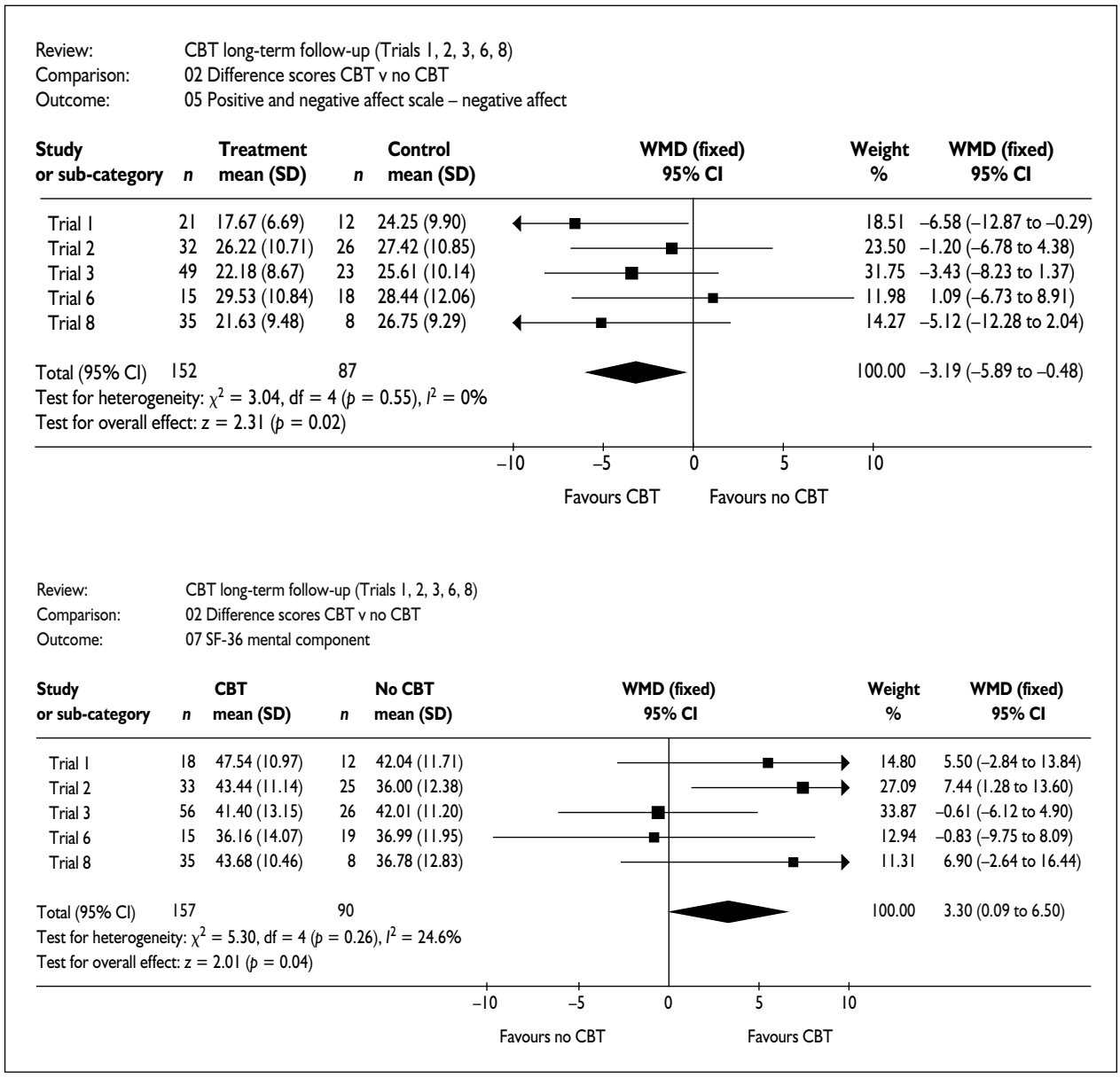


FIGURE 19 *Forest plots of the mean differences between the CBT and non-CBT groups on the long-term follow-up outcome measures yielding significant ($p < 0.05$) overall effects*

evidence for the beneficial effects of CBT on the PTSD patients. Tests of the heterogeneity of the differences were performed on all the meta-analyses and none of them was close to being significant. Although such tests are not very powerful (the tests did not detect differences between the PTSD group and the rest), they provide some justification for pooling the differences.

A meta-analysis was also run on difference between the pretreatment and long-term follow-up scores for the main self-report outcome measure for each trial. These were the SRT for Trials 1 and 3, the BSI-GSI for Trial 2, the PTSD-SCL for Trial 6 and the STAI-T for Trial 8. The pooled effect favoured the CBT group but it was not significant ($p < 0.05$) and this remained so when Trial 6, involving the PTSD patients, was dropped (see *Figure 20*).

In summary, therefore, it can be concluded that the results of the meta-analyses and the pooled data analyses are broadly similar. The statistical significance of the differences between the CBT and non-CBT groups at long-term follow-up in the meta-analyses is slightly less than in the pooled analyses, as might be expected since fewer trials could be included in the meta-analyses. The meta-analyses also allowed the change scores to be

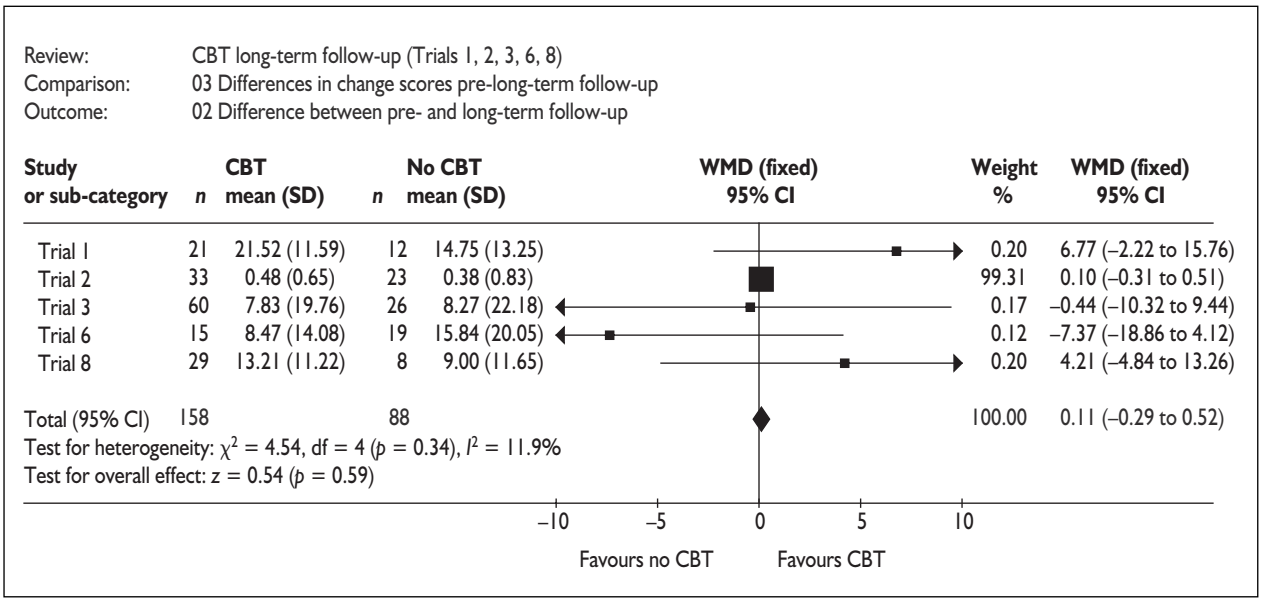


FIGURE 20 *Forest plots of the mean differences between the CBT and non-CBT groups on pretreatment to long-term follow-up differences on the main self-report outcome measure*

analysed by combining differences on different measures across trials. However, this analysis yielded no significant pooled difference between the CBT and non-CBT groups. Although the Forest plots provided some support for the beneficial effects of CBT at long-term follow-up in patients treated for GAD or panic disorder, they provided no support for favourable effects of CBT in PTSD patients.

Discussion

There was an effect of CBT group on symptomatology, with a MANOVA by CBT versus non-CBT treatment on the seven long-term outcome measures common to all eight anxiety disorder studies showing a significant effect, with those receiving CBT having more favourable scores on all measures. However, the differences between the CBT and non-CBT groups were much smaller at long-term follow-up than were generally evident at post-treatment. In addition, there were no significant differences between the CBT and non-CBT groups with regard to diagnostic status or clinically significant change at long-term follow-up.

Patients who had achieved clinically significant change (i.e. 'recovered') following participation in the original trials, regardless of treatment modality, were more likely to remain 'recovered' from post-treatment to long-term follow-up than those who had not achieved clinically significant change at post-treatment. Patients who had failed to respond to initial treatment, whether CBT or non-CBT, were very unlikely to have achieved clinically significant change at long-term follow-up. Patients in the CBT group viewed the trial treatment as being more helpful to them over the long term, were more likely to continue to use what they had learnt and to attribute any improvement to the treatment received than those not receiving CBT.

The findings indicate that, although CBT does continue to result in some improvements in symptomatology over other forms of treatment for anxiety disorders over the long term (2–14 years), it does not appear to lead to greater levels of recovery. It would seem that, although CBT undoubtedly confers distinct advantages over other treatments for anxiety disorders over the short term, such improvements are likely to be eroded over longer periods in many patients.

Summary of CBT versus non-CBT in GAD (Trials 1 and 2)

This section summarises a published paper[149] on the first two GAD studies (Trials 1 and 2). Marked differences in long-term outcome were found between the two studies. In broad terms, about two-thirds of Trial 1 patients could be considered as substantially improved whereas only about one-third of Trial 2 patients could be categorised in this way. In addition, a significantly higher proportion of Trial 2 patients reported receiving a lot of additional treatment over the follow-up

period (3% in Trial 1 versus 23% in Trial 2). Interpretation of this contrasting outcome was problematic since there were also some significant differences in methods between the studies that might influence outcome (e.g. selection criteria, clinical setting, severity of baseline measures and demographic characteristics). It was suggested that the two cohorts of patients were probably best viewed as representing increasing degrees of complexity and severity in the broad spectrum of neurotic disorder seen in various clinical settings within the mental health services.

Notwithstanding the differences between the two studies, analyses of CBT versus non-CBT groups for the combined data set suggested that the CBT groups were generally better at long-term follow-up than the non-CBT groups. A MANOVA on the differences between the CBT and non-CBT groups for the two studies combined on the CGS, BSI-GSI and SF-36 MC measures of long-term outcome yielded a significant overall effect [$F(3,73) = 4.4$, $p = 0.007$], and also significant univariate effects on the SF-36 MC [$F(1,75) = 8.1$, $p = 0.006$] and the BSI-GSI [$F(1,75) = 5.8$, $p = 0.019$]. In all cases the CBT group had better long-term outcome scores than the non-CBT group. In addition, significantly fewer patients from the CBT as opposed to the non-CBT group had received a lot of treatment during the follow-up period [9% versus 26%; $\chi^2(1) = 5.1$, $p = 0.024$]. However, no differences were found between the CBT and non-CBT groups in terms of diagnostic status. It was concluded that CBT was associated with an attenuation of symptoms over the longer term but not with an increased probability of remission.

CBT versus non-CBT for the two psychosis studies (Trials 9 and 10)

Introduction

In the original trials, outcomes for the two psychosis studies were reported at 12 months post-randomisation. The results showed that CBT conducted by clinical nurse specialists in CBT is a helpful adjunct to routine care[66] and that CBT delivered during the early signs of relapse achieved important clinical benefits in terms of reduced relapse and improved symptomatic outcome at 12 months.[67] However, it is not known whether the effects of CBT in these trials endure over time. This section seeks to answer that question by examining the outcome of patients in Trials 9 and 10 at long-term follow-up of between 3 and 6 years post-randomisation.

Method

The sample was assigned to CBT versus non-CBT treatment groups according to the criteria described in the section 'Treatment groups' (p. 17). The resulting groups used in the current analysis are as follows.

Non-CBT

This includes the SP and TAU groups from Trial 9 and the TAU group from Trial 10 ($n = 49$).

CBT

This includes the CBT groups from Trials 9 and 10. The mean number of sessions of therapist contact for patients in the CBT group was 6.5 (SD = 4.9); $n = 44$.

Results

First, the outcome measures at long-term follow-up which were common across all studies were compared using independent *t*-tests by CBT group. The results are shown in *Table 30*.

The CBT group had more favourable scores on the CGS and all subscales of the PANSS at long-term follow-up, and the PANSS negative symptom subscale reached significance. However, the CBT group had less favourable scores on the CGI, BSI-GSI and SF-36, and the SF-36 mental summary component reached significance. [It was noted that the correlations of the self-report SF-36 MC summary component with the clinician-rated PANSS subscales were rather low ($0.01 < r < 0.35$) for scales purportedly measuring similar constructs. Data on SF-36 MC summary scores were missing for a number ($n = 6$) of patients owing to missing responses on one or more items, suggesting that patients had some difficulty completing this measure. It is possible that self-report measures, such as the SF-36, are less valid in psychosis patients, which may account for the differing findings on the SF-36 and PANSS measures.] A MANOVA by CBT versus non-CBT treatment on the outcome measures listed in *Table 30* (excluding PANSS total) just failed to reach overall significance for CBT group [$F(8,74) = 2.0$, $p = 0.060$], although there was an individual effect on the SF-36 MC ($F(1,81) = 7.2$, $p = 0.009$]. A MANOVA by CBT versus non-CBT treatment on the three PANSS subscales did reach overall significance for CBT group [$F(3,89) = 3.5$, $p = 0.019$] and there was an individual effect for the PANSS negative subscale at long-term follow-up [$F(1,91) = 6.0$, $p = 0.016$].

However, it was noted that there were some significant differences between long-term follow-

TABLE 30 *Means (SD) of main outcome measures at long-term follow-up by CBT group and results of independent t-tests for Trials 9 and 10*[a]

	Treatment group				Independent t-test	
	Non-CBT		CBT			
Long-term follow-up	***n***	**Mean (SD)**	***n***	**Mean (SD)**	**95% CI for difference**	***p***
CGS	49	4.31 (1.8)	44	3.80 (1.6)	(–0.2 to 1.2)	0.155
CGI	48	2.31 (1.3)	43	2.65 (1.5)	(–0.9 to 0.2)	0.241
BSI-GSI	45	1.06 (0.9)	44	1.19 (0.8)	(–0.5 to 0.2)	0.484
SF-36 PC	48	43.4 (12.8)	39	41.5 (13.5)	(–3.7 to 7.5)	0.496
SF-36 MC	48	43.0 (10.6)	39	37.7 (8.5)	(1.2 to 9.5)	0.012
PANSS:						
Positive	49	17.7 (6.3)	44	15.3 (6.1)	(–0.1 to 5.0)	0.061
Negative	49	18.0 (6.8)	44	15.0 (4.7)	(0.6 to 5.4)	0.014
General	49	37.0 (10.7)	44	34.5 (11.1)	(–2.0 to 7.0)	0.280
Total score	49	72.7 (19.2)	44	64.8 (19.8)	(–0.1 to 15.9)	0.054

[a] High scores are less favourable on all measures except the SF-36, where high scores are more favourable.

up participants in the CBT and non-CBT groups with regard to scores on the PANSS subscales taken at pretreatment. Specifically, significant differences occurred on the PANSS positive symptom subscale (95% CI for difference = 0.04 to 6.3, $p = 0.047$) and the PANSS negative symptom subscale (95% CI for difference = 0.7 to 7.3, $p = 0.018$). Further checks revealed that there were no significant differences between long-term follow-up participants in the CBT and non-CBT groups on other common pretreatment measures such as the BSI-GSI or the global assessment of functioning (the SF-36 and CGS were not used as common outcome measures at pretreatment). The MANOVA on all outcome measures listed in *Table 30* was therefore repeated controlling for the PANSS positive and negative symptom subscales recorded at pretreatment. Again there was no overall effect of the CBT group, although the individual effect of the SF-36 MC remained significant [$F(1,78) = 6.3$, $p = 0.014$], and the BSI-GSI also reached significance ($F(1,78) = 4.2$, $p = 0.045$], with the CBT group having worse scores in both cases (see *Table 30* for scores).

However, when the MANOVA on the three PANSS subscales was repeated controlling for pretreatment positive and negative subscale scores, the overall effect of the CBT group was no longer significant [$F(3,86) = 2.1$, $p = 0.106$], nor was the individual effect on the negative symptom subscale [$F(1,88) = 2.6$, $p = 0.112$]. On all three long-term follow-up subscales the estimated marginal means were higher for the CBT group and lower for the non-CBT group than the actual means reported in *Table 30*. This suggests, therefore, that the differences on the PANSS between the CBT and non-CBT groups found at long-term follow-up may be due, at least in part, to pretreatment differences on this measure.

Next we examined the CBT group by clinical status (see *Table 31*).

Higher percentages of those in the CBT group had achieved Jacobson criterion (c) with reliable change on the BSI-GSI and also a 25% reduction in scores on the PANSS negative and general symptom subscales and the total PANSS score. However, none of the differences between the CBT and non-CBT groups with regard to clinical status at long-term follow-up were significant.

CBT group by response to trial treatment

Next, we examined whether there were any differences with regard to achieving a 25% reduction (as compared with baseline scores) in PANSS total scores at long-term follow-up for treatment completers who had achieved a 25% reduction in PANSS total scores at post-treatment. The results are shown in *Table 32*.

It can be seen from *Table 32* that very few patients in either treatment group maintained a 25% reduction in scores from post-treatment to long-term follow-up on the PANSS total score. Similar patterns were found on the subscales of the PANSS. Only one significant result was found, namely that a greater percentage of those who had achieved a 25% reduction in general symptom

TABLE 31 *Clinical status by CBT group at long-term follow-up for Trials 9 and 10*

	CBT		Non-CBT			
	n	%	*n*	%	$\chi^2(1)$	*p*
Below BSI-GSI cut-off[a]	44	36.4	45	40.0	0.1	0.724
Jacobson criterion (c)[b]	44	15.9	45	8.9	1.0	0.314
25% improvement in PANSS						
Positive symptoms	44	6.8	48	10.4	0.4	0.541
Negative symptoms	44	25.0	48	22.9	0.1	0.815
General symptoms	44	18.2	48	10.4	1.1	0.285
Total score	44	13.6	48	6.3	1.4	0.234

[a] Cut-off [Jacobson criterion (c) without reliable change] calculated using pretreatment data from the two psychosis studies at long-term follow-up.
[b] Jacobson criterion (c) with reliable change for the two psychosis studies at long-term follow-up.

TABLE 32 *Percentage within CBT group by 25% reduction (from baseline) in PANSS subscale scores at post-treatment and long-term follow-up for Trials 9 and 10*

		Achieved 25% reduction at post-treatment (%)[a]			
		CBT group (n = 44)		Non-CBT group (n = 47)	
PANSS	25% reduction at long-term follow-up	Yes	No	Yes	No
Positive symptoms	Yes	2.3	4.5	2.1	8.5
	No	13.6	79.5	12.8	76.6
		$\chi^2(1) = 0.7$, $p = 0.393$		$\chi^2(1) = 0.1$, $p = 0.734$	
Negative symptoms	Yes	9.1	15.9	4.3	19.1
	No	20.5	54.5	14.9	61.7
		$\chi^2(1) = 0.3$, $p = 0.567$		$\chi^2(1) = 0.0$, $p = 0.926$	
General symptoms	Yes	11.4	6.8	2.1	8.5
	No	13.6	68.2	12.8	76.6
		$\chi^2(1) = 7.3$, $p = 0.007$		$\chi^2(1) = 0.1$, $p = 0.734$	
Total score	Yes	6.8	6.8	2.1	4.3
	No	18.2	68.2	10.6	83.0
		$\chi^2(1) = 2.3$, $p = 0.128$		$\chi^2(1) = 1.2$, $p = 0.270$	

[a] Percentages are within CBT group.

scores at post-treatment in the CBT group also had a 25% reduction in general symptom scores at long-term follow-up in comparison with those who had not achieved the reduction post-treatment (46 versus 9%). Using a log-linear model to test three-way effects revealed that there were no significant differences between the CBT and non-CBT groups for any of the results in *Table 32*.

CBT group by views of treatment variables

The only differences between long-term follow-up participants in the two psychosis studies with respect to the views of treatment variables were that those in the CBT group had a better memory of treatment than those in the non-CBT group [$\chi^2(2) = 8.8$, $p = 0.013$), with 58% of the CBT group (versus 31% of non-CBT) having a good memory of treatment, and 9% of the CBT group (versus 29% of non-CBT) having a poor memory of treatment.

There were no significant differences between the CBT and non-CBT groups with regard to the amount of interim treatment received since the original trial, how hopeful they were of coping in the future, how they viewed themselves with regard to anxiety at long-term follow-up as compared with before the trial treatment or how helpful they had found the original treatment.

TABLE 33 *Direct maximum likelihood estimates of the means of the outcome variables and their standard errors using the predictor variables in the analyses for Trials 9 and 10*

	Non-CBT		CBT			
Long-term follow-up	**Mean**	**SE**	**Mean**	**SE**	$\chi^2(1)$	p
CGS	4.28	0.22	4.05	0.24	0.43	0.511
CGI	2.31	0.16	2.69	0.22	1.64	0.200
BSI-GSI	1.11	0.11	1.28	0.11	1.04	0.308
SF-36 PC	43.5	1.61	40.3	1.87	1.44	0.230
SF-36 MC	42.0	1.43	37.5	1.32	4.42	0.036
PANSS						
Positive	17.9	0.73	15.8	0.89	2.84	0.092
Negative	17.8	0.86	15.0	0.67	5.66	0.017
General	37.0	1.33	35.1	1.45	0.86	0.354

Missing data analysis

As there were missing data on most of the long-term outcome variables, analyses were undertaken using AMOS, which provides estimates and significance tests of differences between the means of two groups on one or more variables using an estimation procedure which takes account of any information that is available on the cases whose data are missing. AMOS was used to compute the DML estimates (DML analysis was reported rather than the EM method because it permits the use of significance tests between the CBT and non-CBT groups whereas the EM analysis does not) of the means of the long-term outcome variables for both the CBT and non-CBT groups (see *Table 33*) using the information in a subset of the predictor variables specified in the section 'Predictors for the two psychosis studies (Trials 9 and 10)' (p. 89) in the estimation process. These were level of medication (chlorpromazine equivalent), age, duration of illness, pretreatment BSI-GSI and PANSS positive, negative and general symptom subscales, and post-treatment BSI-GSI and PANSS positive, negative and general symptom subscales. It was not possible to use all of the predictor variables in the estimation on account of a lack of data and over-fitting problems.

The effect of the missing data appears to be small as the computed means did not differ greatly from the unadjusted means already shown in *Table 30*. In addition, the patterns of differences between the CBT and non-CBT groups were the same in both the computed and unadjusted means.

The equality of the means was then tested by constraining the CBT and non-CBT groups to have the same means on all the long-term outcome variables and seeing if the presence of the constraint resulted in a significantly poorer fit. This yielded a significant ($\alpha = 0.05$) combined test of all the differences [$\chi^2(8) = 16.49$, $p = 0.036$]. In addition, using univariate tests, significant ($\alpha = 0.05$) CBT effects were found on the PANSS negative subscale and the SF-36 MC (see *Table 33*), which matched those found in the actual data in *Table 30*. Hence, the findings suggest that CBT did continue to have an effect, even after the effect of the missing data had been taken into account.

Discussion

At long-term follow-up, participants who had received CBT had more favourable scores on the CGS and all the PANSS subscales in comparison with the non-CBT group. The PANSS negative scale reached statistical significance. This statistically significant effect on negative symptoms for those who received CBT was lost after controlling for pretreatment scores on the PANSS positive and negative subscales. However, the CBT group had less favourable scores on the CGI, BSI-GSI and the SF-36. The SF-36 mental summary component reached statistical significance.

As less than half of the original sample took part at long-term follow-up, missing data analysis to compute DML estimates of the means of the main outcome measures was carried out. The results suggested that any effects of missing data were small.

Our study did not find that the shorter-term effects of CBT were maintained at longer-term follow-up. Previously in a 5-year follow-up of 40 participants randomised to either CBT or ATY (recreational activities with informal support), Drury and colleagues[91] found no differences between CBT and ATY with regard to number of

relapses or admissions, positive or negative symptoms or time spent in acute inpatient facilities. However, for those individuals who received CBT and did not have more than one relapse, the effects of CBT endured. These participants continued to show reduced self-rated delusional conviction, observer-rated delusional ideation, thought disorder, hallucinations and increased perception of control over psychosis. Therefore, one might hypothesise that psychological treatments which emphasise staying well after psychosis or additional interventions aimed at early detection of relapse and intensive preventative interventions may help sustain the effects of CBT and other treatments. For example, Trial 10 had the prevention of relapse and readmission as its central aim. Of those randomised to CBT (n = 72), 28 met criteria for early relapse and received the additional targeted CBT intervention. There is evidence from trials of assertive outreach[152] that, during active phase intervention, hospital admissions are reduced; however, at follow-up there is no difference between routine care and assertive outreach. Therefore, interventions which maintain vigilance for increased risk of relapse, such as early signs monitoring supplemented by booster sessions, may be of benefit to supporting the maintenance of treatment gains. Such an approach has been advocated previously by Kuipers and colleagues.[84] Furthermore, all CBT interventions with psychosis have relapse prevention interventions embedded within the protocol. However, the number of sessions focused on detection and prevention of relapse is usually limited to the final 2–3 sessions. CBT interventions in the future might benefit from a more elaborated intervention protocol focused on staying well after psychosis and maintenance of treatment gains.

Chapter 8

Relative efficacy of different intensities of CBT for anxiety disorders

Introduction

The case for developing brief therapies that emphasise patient self-management has been argued persuasively by a number of authors in recent years[153,154] and this approach is likely to be more efficient for common and less severe conditions. Conversely, clinical experience suggests that intensive therapy should be reserved for those patients with more severe problems and there is some evidence that supports this. Shapiro and colleagues[155] demonstrated a dose–effect relationship in the psychological treatment of depression. Patients with mild to moderately severe depression did as well with eight as with 16 sessions whereas clients with severe depression improved more over 16 than over eight sessions. In the NIMH Collaborative Research Program on the treatment of depression[156] it was found that psychological treatments were only clearly more effective than less intensive placebo therapy with severely depressed patients. In another study,[157] bibliotherapy was found to be as effective as eight sessions of group or individual CBT in panic disorder.

A number of the original trials (i.e. Trials 2, 4, 7 and 8) had examined whether different intensities of CBT were related to different outcomes, including whether brief CBT (generally up to six sessions) could be as effective as standard length CBT and also whether intensive CBT could confer any advantages over shorter lengths of treatment in patients presenting with more chronic or severe symptoms. In Trial 2, a controlled clinical trial of AP and CT with GAD at two different levels of intensity, Durham and colleagues[26] found evidence at 1 year follow-up that the best outcomes were achieved in the more intensive CT condition. In Trial 4, for panic disorder, Power and colleagues[111] found a totally self-administered bibliotherapy condition to be significantly less effective on a range of outcome measures compared with 'standard' therapist-delivered CBT. This may be related to inaccessibility and inconvenience of bibliotherapy during daily activities. In Trial 7, those patients (identified as having good prognosis) who were allocated to brief CBT had the most favourable outcome, whereas there was no significant difference between the standard CBT and intense CBT groups.[116] It was concluded from this study that patient characteristics are a more powerful influence on outcome than the length of therapy. Finally, in Trial 8, 12 sessions of therapist-delivered CBT were found to be more effective than six sessions of CBT, but not more effective than computer-augmented CBT over six sessions at post-treatment.[14] However, there were no differences between treatment conditions at short-term follow-up.

The aim of the current analysis was to examine the relative outcome of patients allocated to different intensities of CBT across the whole sample in order to see whether the differences in outcome observed at post-treatment and short-term follow-up were maintained at long-term follow-up. Not all of the original trials examined this issue and in the case of Trial 7 GAD patients in the low contact CBT group were not randomised to treatment and it was therefore thought prudent to exclude these patients from the analyses. There were no differences with regard to intensity of treatment in the PTSD study (Trial 6) and so these patients were also excluded from this chapter. The results are presented first for patients originally treated for GAD and second for panic disorder patients. The analysis for the two GAD studies which involved different intensities of CBT (i.e. Trials 2 and 7) examines the effects of standard contact CBT versus high contact CBT. The results presented for the panic disorder studies examine low contact CBT versus standard CBT. Only the two Stirling panic disorder studies with different intensities of CBT (i.e. Trials 4 and 5) have been included in the latter analysis as they both used the same outcome measures at pretreatment and at long-term follow-up, allowing more robust comparisons to be made.

Intensity of CBT in GAD (Trials 2 and 7)

Introduction

This section examines the long-term outcome of patients originally treated for GAD at two

TABLE 34 *Means (SD) of main outcome measures at long-term follow-up by intensity of CBT and results of independent t-tests for Trials 2 and 7*

	Treatment group				
	Standard CBT		High contact CBT		Independent t-test
Long-term follow-up	*n*	Mean (SD)	*n*	Mean (SD)	95% CI for difference
CGS	34	3.68 (1.79)	22	3.92 (1.73)	–1.1 to 0.8, $p = 0.771$
CGI	34	2.68 (1.68)	22	2.41 (1.50)	–0.6 to 1.2, $p = 0.548$
BSI-GSI	36	1.17 (0.72)	23	1.44 (0.87)	–0.7 to 0.2, $p = 0.214$
PA	32	28.06 (9.24)	24	29.88 (6.91)	–6.3 to 2.7, $p = 0.424$
NA	32	24.38 (9.32)	24	28.29 (10.09)	–9.1 to 1.3, $p = 0.139$
SF-36 PC	36	39.23 (13.63)	23	40.47 (15.01)	–8.8 to 6.3, $p = 0.745$
SF-36 MC	36	39.63 (13.05)	23	38.12 (11.67)	–5.2 to 8.2, $p = 0.654$

differing intensities of CBT (i.e. standard contact CBT versus high contact CBT). These patients had originally taken part in either Trial 2 or Trial 7.

Method

Treatment groups

The original treatment types were allocated to groups of different intensities of CBT as detailed in the section 'Treatment groups' (p. 17). The breakdown of groups in the current analysis is as follows.

Low contact CBT

There was only one GAD study which had patients allocated to low contact CBT, namely the brief CBT group from Trial 7. However, patients in this treatment condition were not randomised but were allocated to treatment on the basis of good prognosis, and so they have been excluded from the analysis in this section.

Standard CBT

All treatments in this condition followed a standard protocol for CBT for GAD. Treatment was delivered over an average of 8.2 (SD = 3.4) sessions, with an average total therapist contact time of 7.7 (SD = 3.2) hours including assessments. This standard CBT group includes the low intensity CT and AMT groups from Trial 2 and the standard CBT group from Trial 7.

High contact CBT

All treatments in this condition followed a standard protocol for CBT for GAD, which was conducted over an extended period. Treatment was delivered over an average of 12.9 (SD = 6.2) sessions, with an average total therapist contact time of 12.8 (SD = 6.2) hours including assessments. This high contact CBT group includes the high frequency CT group from Trial 2 and the intensive CBT group from Trial 7.

Outcome measures

Outcome measures used in the analysis for the current section are the HAM-A, CGS, CGI, BSI-GSI, PANAS and SF-36. These are described fully in the section 'Main outcome measures' (p. 27).

Results

Examination of the pretreatment composite scores revealed no significant differences between long-term follow-up participants by intensity of CBT on any of the composite pretreatment measures, or any measures common between the two studies (i.e. BSI-GSI, CGS and STAI-T).

The outcome measures at long-term follow-up, which were common across both studies, were compared for the standard and high contact CBT groups using independent *t*-tests. The results are shown in *Table 34*. There were no significant differences between the standard CBT and high contact CBT group on any of the outcome measures at long-term follow-up. Further, a MANOVA by intensity of CBT on these seven common outcome measures was not significant [$F(7,41) = 1.3$, $p = 0.161$]. An independent *t*-test between the standard and high contact CBT groups on the long-term outcome factor was also non-significant (means: standard CBT = 0.08, high contact CBT = 0.29; 95% CI for difference –0.8 to 0.3, $p = 0.456$).

Intensity of CBT group by clinical status at long-term follow-up in the GAD studies is shown in *Table 35*. It can be seen that the standard CBT group had slightly more favourable results than the high contact CBT group, but none of the results were significant.

TABLE 35 *Clinical status by intensity of CBT at long-term follow-up for Trials 2 and 7*

	Standard CBT		High contact CBT			
	n	%	*n*	%	$\chi^2(1)$	*p*
No clinical diagnosis	34	32.4	22	27.3	0.2	0.686
Achieved Jacobson[a]	36	30.6	23	26.1	0.1	0.712
Below BSI-GSI cut-off[b]	36	41.7	23	26.1	1.5	0.223

[a] Jacobson criterion (c) on BSI-GSI (with reliable change) – recalculated across Trials 2 and 7.
[a] Cut-off [Jacobson criterion (c) without reliable change] calculated using all available pretreatment data, applied universally across all trials at long-term follow-up.

TABLE 36 *Percentages[a] within CBT group by clinically significant change at post-treatment and at long-term follow-up for Trials 2 and 7*

	Achieved Jacobson criterion (c) with reliable change on BSI-GSI post-treatment[b]			
	Standard CBT group (*n* = 31)		High contact CBT group (*n* = 18)	
	Yes	No	Yes	No
Achieved Jacobson at long-term follow-up[b]	16.1	12.9	16.7	5.6
Did not achieve Jacobson at long-term follow-up	22.6	48.4	22.2	55.6
	$\chi^2(1) = 1.5$ $p = 0.218$		$\chi^2(1) = 2.8$ $p = 0.093$	

[a] Percentages are within CBT group.
[b] Jacobson criterion (c) on BSI-GSI (with reliable change) – recalculated across Trials 2 and 7.

CBT group by response to trial treatment

Next, we examined whether there were any differences with regard to Jacobson clinically significant change at long-term follow-up for those with post-treatment data, between those who had and had not achieved clinically significant change at post-treatment on the BSI-GSI (this measure was chosen as it was common across the two GAD trials). Jacobson criterion (c) with reliable change was used, based on the pretreatment means calculated across the two studies. The results are shown in *Table 36*.

It can be seen from *Table 36* that, in both the standard and high contact CBT groups, a greater number of patients who did not achieve clinically significant change on the HAM-A post-treatment also failed to achieve clinically significant change at long-term follow-up and very few of those who had not achieved clinically significant change post-treatment went on to achieve clinically significant change at long-term follow-up, in both treatment groups. In the standard CBT group, for patients achieving clinically significant change post-treatment the relative risk of maintaining that status to long-term follow-up (compared with those unrecovered post-treatment) was 1.75 (95% CI 0.8 to 4.1, and for those not achieving clinically significant change post-treatment the relative risk of not achieving clinically significant change at long-term follow-up (compared with those recovered post-treatment) was 1.5 (95% CI 0.8 to 3.4). For those patients in the high contact CBT group achieving clinically significant change post-treatment the relative risk of maintaining that status to long-term follow-up (compared with those unrecovered post-treatment) was 2.6 (95% CI 1.0 to 7.2), and for those not achieving clinically significant change post-treatment the relative risk of not achieving clinically significant change at long-term follow-up (compared with those recovered post-treatment) was 2.9 (95% CI 0.5 to 16.1). None of these results were significant. Using a log-linear model, a three-way test of association between CBT group, responder post-treatment and responder at long-term follow-up was also non-significant [likelihood $\chi^2(1) = 0.5$, $p = 0.491$].

CBT group by views of treatment variables

Comparisons were made between the standard and high contact CBT groups with regard to the views of treatment variables described in the section 'Additional data collected' (p. 31) (i.e. how much they remembered of treatment, how much additional treatment they had received since the original trial, how much they practised anything learnt during treatment, how they were with regard to anxiety now as compared with before the trial, their reason for being better, how helpful they viewed the trial treatment and how hopeful they were of coping in the future).

For those patients who reported being better, significantly more in the standard CBT group gave the trial treatment as their main reason for being better (as opposed to any other reason) than patients in the high contact CBT group [$\chi^2(1) = 5.6$, $p = 0.018$; with 69.6% of the standard CBT group versus 31.3% of the high contact CBT group stating trial treatment]. There were no significant differences between the two groups with regard to any of the other views of treatment variables.

Discussion

The results indicate that there were few differences at long-term follow-up between patients allocated to different intensities of CBT for GAD in the original studies. Patients receiving standard contact CBT generally had more favourable scores at long-term follow-up, but in no case were the results significant. The only significant finding was that patients in the standard contact CBT group who reported being better were more likely to attribute any improvement to the trial treatment than were patients in the high contact CBT group. The results indicate that high contact CBT did not confer any long-term advantage over standard contact CBT in this group of GAD patients.

Intensity of CBT in panic disorder (Trials 4 and 5)

Introduction

Panic disorder with or without agoraphobia is a prevalent condition which results in considerable distress for sufferers. People with panic disorder and agoraphobia are significant consumers of healthcare resources, particularly in the primary care sector. A large research literature exists attesting to the efficacy of CBT as a treatment in the short term. However, there are few data on the endurance of short-term treatment gains over a longer period of time. Such follow-up studies which have been conducted tend to be either naturalistic follow-ups of cohorts of patients which, of course, can tell us little about the efficacy of specific treatments, or they report on the follow-up of specific treatment over relatively short durations post-treatment such as 6 months to 1 year. There is a need, therefore, for longer term follow-up studies of research-controlled CBT treatments.

Although there is considerable evidence attesting to the short-term efficacy of CBT as a treatment for panic disorder and agoraphobia, this treatment is not widely available clinically owing to a shortage of suitably trained therapists. This observation gave rise to a research literature investigating the possibility of increasing the efficiency of delivery of CBT in an effort to increase the clinical availability of the treatment. Studies have been conducted on reduced therapist contact formulations of CBT and on group treatment formulations. There have been no investigations to date of the long-term efficacy of any CBT delivered at differing intensities.

This section examines the effectiveness of CBT for panic disorder, 5–7 years after initial treatment, at two different intensities: low contact CBT and standard contact CBT.

Method

Outcome measures

Long-term outcome measures used in the analysis for this section are the HAM-A, CGS, CGI, SRT, BSI-GSI, PANAS, SF-36 and FQ-Agora. These are described fully in the section 'Main outcome measures' (p. 27).

Treatment groups

Trials 4 and 5 were selected for this analysis as they both delivered different intensities of CBT according to the same protocol, conducted by the same therapist (DS) and using the same outcome measures at pretreatment and long-term follow-up. The treatment conditions in the original trials were allocated to combined treatment groups as follows.

Low contact CBT

This group includes minimum contact CBT from Trial 4, bibliotherapy from Trial 4, and group CBT from Trial 5. Treatment was delivered over an average of 4.4 (SD 2.5) sessions, with an average total therapist contact time of 1.3 (SD 0.6) hours including assessments.

Standard contact CBT

This group includes the maximum contact CBT group from Trial 4 and those from the WL

TABLE 37 *Means (SD) of main outcome measures at long-term follow-up by intensity of CBT, and results of independent t-tests for Trials 4 and 5*

	Treatment group				
	Low contact CBT		Standard CBT		
Long-term follow-up	*n*	Mean (SD)	*n*	Mean (SD)	Independent t-test 95% CI for difference
CGS[a]	50	3.30 (1.72)	25	3.44 (1.50)	–0.9 to 0.7, $p = 0.730$
CGI[a]	50	2.18 (1.60)	25	2.68 (1.97)	–1.3 to 0.3, $p = 0.240$
BSI-GSI[a]	53	1.12 (0.76)	32	1.21 (0.83)	–0.4 to 0.3, $p = 0.619$
PA[a]	48	29.38 (8.09)	24	30.38 (7.11)	–4.9 to 2.9, $p = 0.609$
NA[a]	48	23.23 (8.19)	24	24.67 (8.81)	–5.6 to 2.8, $p = 0.499$
SF-36 PC[a]	48	41.92 (13.25)	31	40.67 (13.61)	–4.9 to 7.4 $p = 0.686$
SF-36 MC[a]	48	38.70 (11.44)	31	40.91 (9.24)	–7.1 to 2.7, $p = 0.371$
HAM-A	48	12.81 (7.99)	25	13.04 (6.25)	–3.9 to 3.4, $p = 0.902$
SRT	54	29.02 (19.33)	32	27.91 (18.50)	–7.3 to 9.6, $p = 0.794$
FQ-Agora	53	12.42 (12.47)	32	11.97 (12.15)	–5.1 to 6.0, $p = 0.872$

[a] Outcome measures common across all eight anxiety disorder studies.

condition in Trial 5 who went on to have individual CBT. Treatment was delivered over an average of 7.0 (SD 1.8) sessions, with an average total therapist contact time of 5.8 (SD 1.9) hours including assessments.

Results

All of the following results refer to patients entered into Trials 4 and 5. Examination of the pretreatment composite scores revealed a significant difference between combined treatment groups on the depression subscale of the patient-rated SRT, with the low contact CBT group [mean = 11.2 (SD 4.8)] having significantly worse scores than the Standard CBT group [mean = 9.4 (SD 4.8)] (95% CI for difference 0.3 to 3.3, $p = 0.023$). There were no differences between the two groups on the clinician-rated measure of depression (i.e. the MADRS), nor were there any differences on any of the other common pretreatment measures.

Independent *t*-tests were used to compare scores on the common outcome measures at long-term follow-up by intensity of CBT treatment. These included the seven outcome measures common to all eight anxiety studies at long-term follow-up, plus the outcome measures used in the original panic disorder trials (i.e. HAM-A, SRT and FQ-Agora). The results are shown in *Table 37*.

There were no significant differences between the low and standard contact CBT group on any of the outcome measures at long-term follow-up. Further, a MANOVA by intensity of CBT on the seven outcome measures which were common across all anxiety disorder studies (see *Table 37*) was not significant [$F(7,58) = 0.9$, $p = 0.543$], nor was a MANOVA on the three common outcome measures used in the original analysis (i.e. HAM-A, SRT and FQ-Agora) [$F(3,68) = 0.1$, $p = 0.854$]. An independent *t*-test between the low and standard contact CBT groups on the long-term outcome factor was also non-significant (means: low contact CBT = 0.0, standard CBT = 0.10; 95% CI for difference –0.6 to 0.4, $p = 0.694$).

Intensity of CBT group by clinical status at long-term follow-up is shown in *Table 38*. There were no statistically significant differences between the low contact CBT group and the standard contact CBT group on any measure.

CBT group by response to trial treatment

Next, we examined whether there were any differences with regard to Jacobson clinically significant change at long-term follow-up for those with post-treatment data between those who had and had not achieved clinically significant change at post-treatment on the HAM-A. The results are shown in *Table 39*.

It can be seen from *Table 39* that, in the low contact CBT group, a greater number of patients who had achieved clinically significant change on the HAM-A post-treatment also achieved clinically significant change at long-term follow-up, and the χ^2 test within the low contact CBT group was significant. In the standard CBT group, almost half of those who had achieved clinically significant change post-treatment no longer reached criteria for clinically significant change at

TABLE 38 *Clinical status by intensity of CBT at long-term follow-up for Trials 4 and 5*

	Low contact CBT		Standard CBT			
	n	%	*n*	%	$\chi^2(1)$	*p*
No clinical diagnosis	51	54.9	25	44.0	0.8	0.372
Below BSI-GSI cut-off[a]	53	52.8	32	40.6	1.2	0.275
Achieved Jacobson on:						
HAM-A[b]	48	62.5	25	48.0	1.4	0.234
SRT[c]	53	18.5	32	12.5	0.5	0.465
FQ-Agora[c]	54	54.7	32	53.1	0.0	0.887

[a] Cut-off [Jacobson criterion (c) without reliable change] calculated using all available pretreatment data, applied universally across all trials at long-term follow-up.
[b] Jacobson criterion (a) as used in the original studies.
[c] Jacobson criterion (c) without reliable change as used in the original studies.

TABLE 39 *Percentages[a] within CBT group by clinically significant change at post-treatment and at long-term follow-up for Trials 4 and 5*

	Achieved Jacobson criterion (a) on HAM-A post-treatment[b]			
	Low contact CBT group (*n* = 39)		Standard CBT group (*n* = 20)	
	Yes	No	Yes	No
Achieved Jacobson at long-term follow-up[b]	35.9	33.3	40.0	10.0
Did not achieve Jacobson at long-term follow-up	2.6	28.2	35.0	15.0
	$\chi^2(1) = 6.6$, $p = 0.010$		$\chi^2(1) = 0.3$, $p = 0.606$	

[a] Percentages are within CBT group.
[b] Jacobson criterion (a) on HAM-A using original calculations.

long-term follow-up; in the low contact group, over half of patients who had not achieved clinically significant change post-treatment, went on to achieve this status at long-term follow-up.

In the low contact CBT group, for patients achieving clinically significant change post-treatment the relative risk of maintaining that status to long-term follow-up (compared with those unrecovered post-treatment) was 6.2 (95% CI 0.9 to 42.1), and for those not achieving clinically significant change post-treatment the relative risk of not achieving clinically significant change at long-term follow-up (compared with those recovered post-treatment) was 1.9 (95% CI 1.2 to 2.9). In the standard CBT group, for patients achieving clinically significant change post-treatment the relative risk of maintaining that status to long-term follow-up (compared with those unrecovered post-treatment) was 1.5 (95% CI 0.3 to 7.1), and for those not achieving clinically significant change post-treatment the relative risk of not achieving clinically significant change at long-term follow-up (compared with those recovered post-treatment) was 1.1 (95% CI 0.7 to 1.9). Using a log-linear model, a three-way test of association between CBT group, responder post-treatment and responder at long-term follow-up was non-significant [likelihood $\chi^2(1) = 1.7$, $p = 0.192$].

CBT group by views of treatment variables

Comparisons were made between the low contact and standard CBT groups with regard to the views of treatment variables described in the section 'Additional data collected' (p. 31) (i.e. how much they remembered of treatment, how much additional treatment they had received since the original trial, how much they practised anything learnt during treatment, how they were with regard to anxiety now as compared with before the trial, their reason for being better, how helpful

they viewed the trial treatment and how hopeful they were of coping in the future). When asked how they were with regard to anxiety at long-term follow-up compared with before the original trial, significantly more in the standard CBT group reported being worse (as opposed to no change or better) than patients in the low contact CBT group [$\chi^2(1) = 4.0$, $p = 0.045$; 19.2% of the standard CBT group versus 3.8% of the low contact CBT group]. There were no significant differences between the two groups with regard to any of the other views of treatment variables.

Discussion

A strength of the results presented above is that these two groups of patients were very similar in presenting problems and all treatment conditions were delivered to the same protocol by the same therapist. The results show that there were virtually no differences between the low contact CBT and standard CBT groups with regard to outcome at long-term follow-up in this group of patients originally treated for panic disorder 5–7 years earlier. In fact, the low contact CBT group may have had a slightly more favourable outcome, with most of those considered recovered post-treatment in this treatment group remaining so at long-term follow-up; and more than half of those considered unrecovered post-treatment achieving clinically significant change at long-term follow-up. In contrast, almost half of those in the standard CBT group who were considered recovered post-treatment failed to maintain recovery to long-term follow-up. In addition, significantly more of the standard CBT than the low contact CBT group reported being 'worse' at long-term follow-up. It would appear that there was some fall-off in the post-treatment gains achieved by the standard contact CBT group over long-term follow-up, whereas a proportion of the low contact CBT group showed continuing improvement over the long-term follow-up period.

Unfortunately, it is difficult to discern the reasons for these differing patterns of treatment response over the long term. There were no differences with regard to the amount of interim treatment reported by the two groups, suggesting that this was not responsible for the different patterns of change in clinical status between the two groups. Indeed, when patients were asked why they thought they were better at long-term follow-up, roughly the same percentage of patients in both groups stated the trial treatment (52% of the low contact CBT group and 56% of the standard CBT group). Further, 77% of those in the low contact CBT group who had changed status from unrecovered at post-treatment to recovered at long-term follow-up gave the trial treatment as the reason for their long-term improvement. Despite the difficulties in interpreting the results of this analysis, it is clear from the current research that, over the longer term, standard CBT may be no more effective than low contact CBT, at least in this group of anxiety patients.

Chapter 9

Predictors of long-term outcome

Introduction

In this chapter we examine predictors of long-term outcome across the eight anxiety disorder studies and the two psychosis studies. The first study examines the long-term predictive power of pretreatment variables and post-treatment clinical status and is based on variables common to all eight anxiety disorder studies. It was not possible to examine the predictive power of the quality of the therapeutic alliance in this investigation as only Trials 2 and 7 had consistently measured this variable. This issue has been examined in a separate study of two prognostic indices of the complexity and severity of presenting problems (CASP index) and the quality of the therapeutic alliance (CAIR index) that has recently been submitted for publication. This study is briefly summarised in the section 'Summary of outcome prediction study for Trials 2 and 7 only' (p. 88). The chapter ends with a study of outcome prediction in the two psychosis studies.

Predictors for the eight anxiety disorder studies (Trials 1–8)

Introduction

The section examines the predictors of long-term outcome across the eight anxiety disorders, using a range of pretreatment patient characteristic and demographic variables and the composite pre- and post-treatment outcome measures.

Method

Measures

The eight anxiety disorder studies had each used a variety of different measures of clinical severity, depression and anxiety symptoms at the time of the original trial. The composite measures described in the section 'Composite measures', (p. 29) were therefore used in the current analysis to maximise sample size. These were a composite measure of anxiety symptoms (CANX); a composite measure of depression symptoms (CDEP); a composite clinical global severity score (COMCGS); and the adjusted HAM-A measure (HAM-AD). The composite measures of social adjustment in the family domain (CFAM) and the social domain (CSOC) were also used in some of the following analyses. Other predictor variables covered demographic information (described in the section 'Additional data collected', p. 31), specifically job status, marital status, gender, age and deprivation score. In addition, three variables relating to status at the time of the original trial, which were common across all eight studies (i.e. the taking of concurrent psychotropic medication, evidence of definite avoidance behaviour and duration of current episode of disorder), were also included in the analysis. Finally, the self-reported amount of interim treatment (between original trial and long-term follow-up) described in the section 'Additional data collected' (p. 31) was also used as a predictor variable.

The outcome measures used were the long-term outcome factor described in the section 'Composite measures' (p. 29), whether patients had achieved clinically significant change on one outcome measure (described in the section 'Jacobson method applied to primary outcome measures' (p. 45) and presence of any clinical diagnosis at long-term follow-up.

Results

The pretreatment predictor variables used in this section are described in *Table 40*. This table shows the mean and SD for quantitative variables and the percentages in each category for categorical variables.

Table 41 describes the additional categorical variables used in the current analysis.

Table 42 describes the outcome measures which have been used in the current analysis at pretreatment, post-treatment and at long-term follow-up.

In addition to the long-term outcome factor, the binary variables 'Achieved Jacobson criteria on main measure' and 'Had any clinical diagnosis at long-term follow-up' were also used as long-term outcome variables in the current analysis.

Table 43 shows the correlations between all the predictor variables with the long-term outcome factor and also with whether they had achieved Jacobson criteria on the main outcome measure

TABLE 40 *Descriptive data on pretreatment variables for Trials 1–8*

Pretreatment	*n*	%
Gender	854	
Male	304	35.6
Female	550	64.4
Concurrent psychotropic medication during trial	850	
Yes	388	45.6
No	462	54.4
Job status	796	
Not employed or not working owing to health	268	33.7
Housewife	98	12.3
Working or retired (age)	430	54.0
Married/cohabiting	819	
Yes	534	65.2
No	285	34.8
Definite avoidance at time of trial	825	
Yes	557	67.5
No	268	32.5
		Mean (SD)
Age at time of original trial	838	37.86 (11.67)
Deprivation score at time of trial	841	0.00 (2.36)
Duration of current episode (category)	826	4.19 (1.05)

TABLE 41 *Descriptive data for additional variables used in the predictor analysis for Trials 1–8*

	n	%
Completed treatment	854	
Yes	634	74.2
No	220	25.8
Cognitive therapy treatment group	850	
CBT	615	72.4
Non-CBT	235	27.6
Long-term follow-up		
Achieved Jacobson criteria on main measure	341	
Yes	144	42.2
No	197	57.8
Any clinical diagnosis	344	
No	166	48.3
Yes	178	51.7

and whether they had any clinical diagnosis at long-term follow-up.

Table 43 shows that the most important demographic variables in predicting the long-term outcome factor are employment status and deprivation score. Excluding the correlations with CFAM and CSOC which are affected by missing data, the pretreatment scores on the outcome measures correlate moderately with the long-term outcome factor (between 0.25 and 0.42), whereas the post-treatment variables correlate with the long-term outcome factor at about the same level but within a much more narrow range (between 0.39 and 0.43). The correlations with the binary variables achieving a clinically significant improvement as measured by the Jacobson criteria and having any diagnosis at long-term follow-up show the same pattern as the correlations with the long-term outcome factor although they are somewhat lower. Being in the CBT group is significantly ($\alpha = 0.05$) related to the long-term outcome factor but not to the binary outcome components.

TABLE 42 *Descriptive data for common outcome measures at pretreatment, post-treatment and long-term follow-up for Trials 1–8*

	Valid *n*	Mean	SD
Pretreatment			
CFAM	723	0.00	1.00
CSOC	703	0.00	1.00
CDEP	822	0.17	1.10
COMCGS	841	4.87	1.10
HAM-AD	844	21.12	5.79
CANX	812	0.03	1.00
Post-treatment			
CDEP	658	0.06	1.05
COMCGS	646	2.68	2.03
HAMA-AD scores	574	10.08	8.39
CANX	648	0.00	1.00
Long-term follow-up			
CGS	343	3.32	1.82
CGI	342	2.29	1.66
BSI-GSI	382	1.21	0.84
PA	347	29.36	8.58
NA	347	24.14	9.67
SF-36 PC	368	40.85	13.56
SF-36 MC	368	40.36	12.08
Amount of interim treatment	384	2.25	1.14

As we had seen in Chapter 5 that there was a large amount of missing data at long-term follow-up, it was decided to examine the possible effects of missing data on the correlations between scores on the long-term outcome factor and the other measures. This was done by comparing different estimates of the correlations. *Table 44* contains correlations using all the available data (pairwise correlations), and correlations where the missing values are estimated by EM imputation on Statistical Package for Social Sciences (SPSS) using only the predictor variables including interim treatment and again using the predictor variables and all the long-term follow-up variables. The last set of correlations might be criticised because, by including the long-term outcome factor together with the long-term outcome variables, the same information is being used twice.

The main effect of missing data appears to be that the correlations between the psychometric measures and the outcome variables have been slightly underestimated if the missing data are ignored. A smaller but important effect is that the missing data may have resulted in the effect of CBT on outcome being overestimated, as the correlation between CBT group and the long-term outcome factor was lower in the EM imputation. However, all of the effects of missing data are small. A further analysis was performed using AMOS[150] to compute direct maximum likelihood estimates of the same correlations. These are not reported because they are virtually identical with the EM estimates.

Regression analyses

Stepwise regression analyses were performed to predict the long-term outcome factor from all the predictor variables available at post-treatment (i.e. all variables listed in *Table 43*). The social functioning scales (which were not used in Trial 1) were not selected in the initial stepwise analyses and so the analyses were rerun without those variables to include data from Trial 1. The regression was then rerun using only the selected variables, which resulted in another 61 cases being added to the analysis, but the solution remained very similar with a multiple correlation of 0.61, meaning that about 38% of the variance was being explained. The multiple r rose to 0.69 when the amount of interim treatment was added to the solution, but this may not be interesting as these data are not available at post-treatment to predict the outcome. Neither of the treatment variables (i.e. completed treatment or CBT group) contributed to the solution. CSOC and CFAM did contribute to the solution when interim treatment was added. The results of the regressions are shown in *Table 45*.

Stepwise regressions predicting the long-term outcome factor were carried out on different

TABLE 43 *Correlations of the predictor variables with extracted long-term outcome factor and whether achieved Jacobson criteria at long-term follow-up for Trials 1–8*[a]

	Long-term outcome factor			Long-term follow-up achieved Jacobson criteria			Any clinical diagnosis		
	r	*p*	*n*	*r*	*p*	*n*	*r*	*p*	*n*
Pretreatment patient characteristics									
Gender (1 = M; 2 = F)	–0.021	0.716	312	0.056	0.305	341	–0.073	0.175	344
Taking concurrent psychotropic medication	0.134	0.018	311	–0.122	0.024	338	0.116	0.032	341
Job status[b]	–0.351	<0.001	304	0.298	0.000	334	–0.295	<0.001	335
Married or cohabiting	–0.154	0.006	311	0.156	0.004	340	–0.126	0.020	343
Definite avoidance	0.099	0.082	307	–0.054	0.321	335	0.073	0.178	338
Age	–0.053	0.349	312	0.059	0.279	341	0.005	0.921	344
Poorer deprivation score	0.256	<0.001	309	–0.212	0.000	338	0.192	<0.001	341
Longer duration of current episode (category)	0.157	0.006	308	–0.100	0.068	335	0.170	0.002	338
Pretreatment composite measures									
CFAM	0.101	0.091	281	–0.102	0.074	306	0.115	0.044	309
CSOC	0.330	<0.001	273	–0.223	<0.001	298	0.320	<0.001	301
CDEP	0.422	<0.001	307	–0.242	<0.001	337	0.323	<0.001	339
COMCGS	0.247	<0.001	310	–0.267	<0.001	338	0.237	<0.001	341
HAMA-AD	0.419	<0.001	311	–0.327	<0.001	339	0.324	<0.001	342
CANX	0.341	<0.001	303	–0.242	<0.001	334	0.254	<0.001	334
Post-treatment									
Completer	–0.204	<0.001	312	–0.052	0.338	341	–0.174	0.001	344
Had CBT	–0.118	0.038	311	0.059	0.278	339	–0.018	0.746	342
CDEP	0.430	<0.001	268	–0.246	<0.001	296	0.328	<0.001	294
COMCGS	0.398	<0.001	266	–0.265	<0.001	294	0.364	<0.001	292
HAMA-AD	0.394	<0.001	233	–0.278	<0.001	258	0.319	<0.001	256
CANX	0.411	<0.001	268	–0.250	<0.001	295	0.309	<0.001	293
More interim treatment	0.485	<0.001	312	–0.375	<0.001	341	0.443	<0.001	343

[a] Where one variable is continuous and the other is binary the correlations are point biserial correlations and where both are binary the correlations are phi coefficients.
[b] 1 = Not employed or not working owing to health; 2 = housewife; 3 = working or retired (age).

subsets of the variables: the patient characteristics at entry, pretreatment scores on the outcome measures (with and without the social functioning scales as they were not used in Trial 1) and post-treatment outcome scores. Each analysis yielded a multiple *r* between 0.4 and 0.5. Analysing the data with only the selected variables made little difference to the solutions and increased the sample size by less than ten cases (see *Table 46*).

When the pretreatment patient characteristics and composite measures were combined to predict the long-term outcome factor, the multiple *r* rose to 0.56 (see *Table 46*). To investigate further the relationship between these groups of predictors and the outcome variable, principal component analyses were performed on the five pretreatment patient characteristics that correlated with the long-term outcome factor (as shown in *Table 43*) and on the four composite pre- and post-treatment outcome measures. The first component accounted for 27% of the variance of the patient characteristics in the first analysis, 60% of the variance of the pretreatment tests in the second and 84% of variance of the post-treatment tests in the third. The correlations between these components and the long-term outcome factor were 0.42, 0.47 and 0.42. This means that these principal components were as good as the stepwise regressions reported in *Table 46* at predicting the long-term outcome factor and that the exact regression weightings are somewhat arbitrary. Predicting the long-term outcome factor from a regression with all three components yielded a multiple *r* of 0.6 and all three components contributed significantly to the regression.

Since it is desirable to be able to predict clinically meaningful results, stepwise logistic regressions were run on meeting the Jacobson criterion and

TABLE 44 *Correlations of extracted long-term outcome factor with remaining variables including estimated correlations for Trials 1–8*

	Estimation method		
	Pairwise (based on all cases with scores on both variables)	**EM estimation using all predictor variables plus interim treatment**	**EM estimation using all predictor and long-term outcome variables plus interim treatment**
Pretreatment			
Gender (1 = M; 2 = F)	–0.021	–0.037	–0.015
Taking concurrent medication	0.134	0.174	0.193
Job status[a]	–0.351	–0.333	–0.328
Married or cohabiting	–0.154	–0.115	–0.104
Definite avoidance	0.099	0.140	0.134
Age	–0.053	–0.090	–0.082
Poorer deprivation score	0.256	0.221	0.211
Longer duration current episode	0.157	0.149	0.138
CFAM	0.101	0.186	0.210
CSOC	0.330	0.367	0.362
CDEP	0.422	0.465	0.449
COMCGS	0.247	0.295	0.292
HAMA-AD	0.419	0.431	0.422
CANX	0.341	0.347	0.351
Post-treatment			
CDEP	0.430	0.508	0.512
COMCGS	0.398	0.448	0.435
HAMA-AD	0.394	0.496	0.505
CANX	0.411	0.455	0.457
Had CBT	–0.118	–0.098	–0.085
Completer	–0.204	–0.233	–0.221
Interim treatment	0.485	0.494	0.511

[a] 1 = Not employed or not working owing to health; 2 = housewife; 3 = working or retired (age).

TABLE 45 *Regressions of long-term outcome factor using all the variables for Trials 1–8*

		Selected from					
All variables entered		**Only variables below entered**		**All variables plus interim treatment**		**Only variables below entered**	
HAM-AD pretreatment	0.307	HAM-AD pretreatment	0.307	More interim treatment	0.351	More interim treatment	0.353
COMCGS post-treatment	0.153	COMCGS post-treatment	0.174	HAM-AD pretreatment	0.261	HAM-AD pretreatment	0.241
Job status[a]	–0.136	Job status[a]	–0.119	CDEP post-treatment	0.225	CDEP post-treatment	0.265
Poorer deprivation score at time of trial	0.159	Poorer deprivation score at time of trial	0.134	Poorer deprivation score at time of trial	0.117	Poorer deprivation score at time of trial	0.097
CDEP post-treatment	0.208	CDEP post-treatment	0.215	CFAM pretreatment	0.171	CFAM pretreatment	0.144
				CSOC pretreatment	–0.151	CSOC pretreatment	–0.107
Multiple *r*	0.619		0.614		0.692		0.683
r^2	0.383		0.377		0.478		0.467
n	193		253		193		233

[a] 1 = Not employed or not working owing to health; 2 = housewife; 3 = working or retired (age).

TABLE 46 *Stepwise regressions predicting long-term outcome factor at long-term follow-up from selected subsets of variables for Trials 1–8*

		Selected from					
Pretreatment patient characteristics		**Pretreatment outcome measures**		**Pretreatment outcome measures less social functioning**		**Post-treatment variables**	
Job status[a]	–0.321	CDEP pretreatment	0.244	CDEP pretreatment	0.266	HAMA-AD post-treatment	0.240
Poorer deprivation score at time of trial	0.218	HAMA-AD pretreatment	0.199	HAMA-AD pretreatment	0.264	CDEP post-treatment	0.196
		CSOC pretreatment	0.173				
Multiple *r*	0.413		0.486		0.467		0.410
r^2	0.170		0.236		0.218		0.168
n	292		268		302		234

[a] 1 = Not employed or not working owing to health; 2 = housewife; 3 = working or retired (age).

TABLE 47 *Stepwise logistic regressions predicting achieved Jacobson criterion for clinically significant change at long-term follow-up for Trials 1–8*

Selected from all variables but interim treatment			**Only variables below entered**			**Variables in column 4 plus interim treatment**		
	B	**Exp(*B*)**		***B***	**Exp(*B*)**		***B***	**Exp(*B*)**
Job status[a]	0.509	1.663	Job status[a]	0.585	1.795	Job status[a]	0.517	1.676
COMCGS pretreatment	–0.667	0.513	COMCGS pretreatment	–0.566	0.568	CSOC pretreatment	–0.716	0.489
COMCGS post-treatment	–0.242	0.785	COMCGS post-treatment	–0.185	4.867	More interim treatment	–0.834	0.434
Constant	2.487	12.026	Constant	1.582	4.867	Constant	3.694	40.221
Cox and Snell r^2	0.196			0.184			0.269	
n	216			288			288	

[a] 1 = Not employed or not working owing to health; 2 = housewife; 3 = working or retired (age).

the presence of any clinical diagnosis variable. The variables predicting the Jacobson criterion were job status and the CGS scales pre- and post-treatment, and together they explained around 20% of the variance (see *Table 47*). This dropped slightly when only these variables were used in order to maximise sample size. When the interim treatment variable was added the predicted variance rose to 26.9% (equivalent to a multiple *r* of 0.52).

Stepwise logistic regressions predicting any diagnosis at long-term follow-up explained somewhat more of the variance (see *Table 48*). When the interim treatment variable was added the predicted variance rose to 34.2%. Since the social functioning variable was used, Trial 1 was excluded from this analysis. When social functioning was dropped from the last regression to include Trial 1 the variance explained dropped from 34.2 to 32.7%.

Summary of outcome prediction study for Trials 2 and 7 only

This study, which has been submitted for publication, tested two hypotheses regarding the short- and long-term predictive validity of two prognostic indices concerning the complexity of presenting problems measured before treatment and quality of therapeutic alliance measured during the initial phase of therapy. The hypotheses were based on the findings of Trial 7 in which outcome at follow-up was better predicted by

TABLE 48 *Stepwise logistic regressions predicting any clinical diagnosis at long-term follow-up for Trials 1–8*

Selected from all variables but interim treatment			Only variables below entered			Variables in column 4 plus interim treatment		
	B	**Exp(*B*)**		***B***	**Exp(*B*)**		***B***	**Exp(*B*)**
Poorer deprivation score pretreatment	0.147	1.158	Poorer deprivation score pretreatment	0.131	1.140	Job status[a]	–0.370	0.691
Job status[a]	–0.614	0.541	Job status[a]	–0.463	0.629	CSOC pretreatment	0.429	1.535
CSOC pretreatment	0.395	1.484	CSOC pretreatment	0.381	1.463	HAMA-AD pretreatment	0.088	1.092
HAMA-AD pretreatment	0.100	1.105	HAMA-AD pretreatment	0.090	1.094	COMCGS post-treatment	0.308	1.361
COMCGS post-treatment	0.300	1.350	COMCGS post-treatment	0.387	1.472	More interim treatment	0.749	2.115
Constant	–1.408	0.245	Constant	–1.951	0.142	Constant	–-3.522	0.030
Cox and Snell r^2	0.280		Cox and Snell r^2	0.282		Cox and Snell r^2	0.342	
n	213		*n*	250		*n*	250	

[a] 1 = Not employed or not working owing to health; 2 = housewife; 3 = working or retired (age).

pretreatment variables than by therapeutic alliance variables, whereas outcome immediately post-treatment was better predicted by therapeutic alliance variables than by pretreatment variables.[116]

This study was made possible by the fact that Trials 2 and 7 of CBT for GAD both used very similar measures. Specifically, both included specific measures of the quality of the therapeutic alliance and initial response to therapy in addition to pretreatment variables relating to the complexity and severity of presenting problems.

The analysis found that poorer post-treatment outcome as measured by CGS was predicted reasonably well by both poorer quality of therapeutic alliance ($r = 0.35$) and greater complexity of problems ($r = 0.36$) but the equivalent analysis predicting CGS at long-term follow-up showed that quality of therapeutic alliance was a poor predictor ($r = –0.05$) whereas complexity of problems was a good one ($r = 0.62$). The same general pattern of results was found when predicting CGI and STAI-T from these variables. Interestingly, the relationship between complexity and severity of problems and outcome variables was stronger at long-term follow-up than at post-treatment.

These results suggest that a close relationship between therapist and patient will not in itself reduce underlying vulnerability if it is not a vehicle for effective change. The CAIR index, therefore, may need to be supplemented with an additional index assessing the validity of the case formulation and, by implication, the competency with which therapy was delivered. There is now a growing body of evidence pointing to differences in therapist efficacy within clinical trials[158] and significant associations between positive outcomes of treatment and ratings of therapist expertise in delivering treatment protocols used in clinical trials.[159] Sustained reductions in vulnerability to common mental health problems may be linked to therapist expertise in delivering specific treatment protocols, especially for more complex and severe problems, whereas short-term improvements in symptom severity and morale may be influenced primarily by a positive therapeutic alliance irrespective of treatment approach. This merits further investigation.

Predictors for the two psychosis studies (Trials 9 and 10)

Introduction

The section examines the predictors of long-term outcome across the two psychosis studies, using pretreatment patient characteristics and pre- and post-treatment outcome measures which were common to both studies, as predictor variables. This was possible as, prior to the initiation of the two studies, the principal investigators (AIG and RCD) had agreed common measures to facilitate future collaboration. The current analysis below, consistent with other studies of CBT,[63] uses symptom severity (total score on the PANSS) as the primary outcome measure.

TABLE 49 *Descriptive data on pretreatment patient characteristics and trial status variables for Trials 9 and 10*

Pretreatment	***n***	**%**
Gender	210	
Male	150	71.4
Female	60	28.6
Completed treatment	210	
Yes	191	91.0
No	19	9.0
Cognitive therapy treatment group	210	
CBT	94	44.8
Non-CBT	116	55.2
		Mean (SD)
Chlorpromazine equivalent level	196	466.7 (366.5)
Age at time of original trial	210	36.2 (10.0)
Deprivation score at time of trial	201	1.8 (3.1)
Duration of illness (months)	199	128.5 (92.3)

Method

Measures

The two psychosis studies used a number of common measures of clinical severity and symptoms at the time of the original trials. Pre- and post-measures examined in the current analysis were the PANSS [i.e. the positive, negative and general subscales plus the total scores), the BSI-GSI and the seven subscales of the social functioning scale [i.e. social engagement/withdrawal (here termed withdrawal), interpersonal communication, independence performance, independence competence, recreation, prosocial and occupation/employment (here termed employment)]. Other predictor variables related to patient characteristics at trial entry; these were gender, age, deprivation score, duration of illness and level of medication (assessed as chlorpromazine equivalent). Only limited demographic data were recorded at pretreatment in Trial 10, so it was not possible to examine the effects of factors such as marital status and employment on outcome in the psychosis patients. Two variables relating to trial status, namely whether or not the patient was allocated to the CBT group and whether or not the patient completed treatment (termed completer), were also included in the analysis. For the psychosis patients, the amount of interim treatment was not used as a predictor variable, as 93% of patients fell into the same category (i.e. 'a lot'), and this variable proved not to be related to outcome. The outcome measure used was the long-term follow-up total score on the PANSS. All of these variables are described fully in the section 'Composite measures' (p. 29).

Results

The pretreatment patient characteristics and trial status predictor variables used in this analysis are described in *Table 49*. This table shows the mean and SD for quantitative variables and the percentages in each category for categorical variables.

Table 50 describes the outcome measures which were been used in the current analysis at pretreatment, post-treatment and long-term follow-up.

Table 51 shows the correlations between all the predictor variables with the PANSS total score at long-term follow-up. With regard to the patient characteristic variables, it can be seen that higher pretreatment levels of medication (chlorpromazine equivalent) and longer duration of illness at trial entry were significantly associated with poorer long-term outcome on the PANSS.

High (i.e. worse) scores on the BSI-GSI and PANSS, taken both at pre- and post-treatment, correlated moderately to highly with poorer long-term follow-up PANSS total scores (between 0.26 and 0.62). With regard to the social functioning subscales, there were significant associations between long-term follow-up total scores on the PANSS and the pretreatment independence competence scale and also with the post-treatment withdrawal, independence competence, recreation and employment/occupation scales; in all cases poorer social functioning was related to worse PANSS scores. The relationship between being in the CBT group and PANSS total scores at long-

TABLE 50 *Descriptive data for common outcome measures at pretreatment, post-treatment and long-term follow-up for Trials 9 and 10*

	Valid *n*	Mean	SD
Pretreatment			
BSI-GSI	186	1.39	0.87
PANSS positive	203	15.12	7.53
PANSS negative	203	17.01	8.43
PANSS general	203	35.31	10.10
PANSS total	203	67.67	24.05
Social functioning			
Withdrawal	202	94.36	12.66
Interpersonal communication	202	113.74	17.60
Independence performance	202	102.29	12.83
Independence competence	201	107.67	13.29
Recreation	202	95.29	14.44
Prosocial	202	99.60	13.09
Employment	200	93.23	11.97
Post-treatment			
BSI-GSI	157	1.12	0.82
PANSS positive	192	13.84	7.68
PANSS negative	192	15.55	8.46
PANSS general	192	32.44	11.20
PANSS total	192	61.96	25.99
Social functioning			
Withdrawal	191	95.73	14.37
Interpersonal communication	191	116.48	18.98
Independence performance	192	103.78	12.80
Independence competence	192	109.30	13.43
Recreation	192	96.28	13.91
Prosocial	191	101.05	14.82
Employment	190	95.17	13.46
Long-term follow-up			
PANSS total	93	68.94	19.77

term follow-up approached significance ($p = 0.054$), with those in the CBT group having lower PANSS scores.

As we had seen that there was a large amount of missing data at long-term follow-up (see Chapter 5), it was decided to examine the possible effects of missing data on the correlations between PANSS total scores at long-term follow-up and the other measures. This was done by comparing estimates of the correlations. *Table 52* contains correlations using all the available data (pairwise correlations), and correlations where the missing values are estimated by EM imputation on SPSS using the predictor variables and PANSS total scores at long-term follow-up.

It can be seen from *Table 52* that the effects of missing data are small, with the pairwise and EM estimated correlations being very similar. However, one important effect is that the missing data may have resulted in the effect of CBT on outcome being overestimated, as the correlation between CBT group and the PANSS total score was lower in the EM imputation.

Regression analyses

Stepwise regression analyses were performed to predict the PANSS total score at long-term follow-up from all the predictor variables available at post-treatment (i.e. all variables listed in *Table 51*) (excluding the PANSS total scores as these are duplicated by the PANSS subscales). The regression was then rerun using only the selected variables, which resulted in another 26 cases being added to the analysis, but the solution remained very similar with a multiple correlation of 0.64, meaning that about 38% of the variance was being explained. Neither of the treatment variables (i.e. completed treatment or CBT group) contributed

TABLE 51 *Correlations of the predictor variables with the PANSS total score at long-term follow-up for Trials 9 and 10*[a]

	Long-term follow-up PANSS total		
	r	***p***	***n***
Pretreatment			
Age	–0.106	0.313	93
Gender	–0.031	0.766	93
Chlorpromazine equivalent	0.234	0.025	91
Deprivation score	–0.111	0.296	91
Duration of illness (months)	0.224	0.032	92
BSI-GSI	0.344	0.001	85
PANSS positive	0.602	<0.001	92
PANSS negative	0.475	<0.001	92
PANSS general	0.578	<0.001	92
PANSS total	0.605	<0.001	92
Social functioning			
Withdrawal	–0.384	<0.001	92
Interpersonal communication	–0.062	0.558	92
Independence performance	–0.029	0.786	92
Independence competence	–0.220	0.035	92
Recreation	–0.034	0.746	92
Prosocial	0.044	0.675	92
Employment	–0.091	0.391	92
Post-treatment			
Completer	–0.087	0.406	93
CBT group	–0.201	0.054	93
BSI-GSI	0.261	0.028	71
PANSS positive	0.612	<0.001	91
PANSS negative	0.562	<0.001	91
PANSS general	0.595	<0.001	91
PANSS total	0.620	<0.001	91
Social functioning			
Withdrawal	–0.353	0.001	91
Interpersonal communication	–0.142	0.180	91
Independence performance	–0.112	0.293	91
Independence competence	–0.263	0.012	91
Recreation	–0.206	0.051	91
Prosocial	–0.139	0.192	91
Employment	–0.215	0.041	91

[a] Where one variable is continuous and the other is binary the correlations are point biserial correlations and where both are binary the correlations are phi coefficients.

to the solution. The results of the regressions are shown in *Table 53*.

Stepwise regressions predicting the PANSS total score at long-term follow-up were carried out on different subsets of the variables, namely the patient characteristics at entry, pretreatment scores on the outcome measures and post-treatment outcome scores. Each analysis yielded a multiple *r* between 0.27 and 0.59 (see *Table 54*).

Analysing the data with only the selected variables made little difference to the first two solutions and increased the sample size by less than six cases. However, in the third case (post-treatment variables only), running the analysis with only the PANSS positive subscale score increased the sample size by 21 and the multiple *r* rose from 0.42 to 0.61 (explaining around 37% of the variance). The difference in the sample size was largely due to missing data on the post-treatment BSI-GSI; removing this variable from the stepwise analysis reported in *Table 54* resulted in a very similar solution to that when the post-treatment PANSS positive subscale score was entered as the only variable.

TABLE 52 *Correlations of long-term outcome variables with remaining variables including estimated correlations for Trials 9 and 10*

	Long-term follow-up PANSS Total	
	Pairwise[a]	**Estimation[b]**
Pretreatment		
Age	–0.106	–0.150
Gender (female)	–0.031	–0.062
Chlorpromazine equivalent	0.234	0.238
Deprivation score	–0.111	–0.085
Duration of illness (months)	0.224	0.201
BSI-GSI	0.344	0.419
PANSS positive	0.602	0.608
PANSS negative	0.475	0.463
PANSS general	0.578	0.566
PANSS total	0.605	0.588
Social functioning		
Withdrawal	–0.384	–0.366
Interpersonal communication	–0.062	0.020
Independence performance	–0.029	–0.130
Independence competence	–0.220	–0.190
Recreation	–0.034	–0.107
Prosocial	0.044	0.049
Employment	–0.091	–0.080
Post-treatment		
Completer	–0.087	–0.201
CBT group	–0.201	–0.162
BSI-GSI	0.261	0.413
PANSS positive	0.612	0.610
PANSS negative	0.562	0.560
PANSS general	0.595	0.594
PANSS total	0.620	0.618
Social functioning		
Withdrawal	–0.353	–0.321
Interpersonal communication	–0.142	–0.116
Independence performance	–0.112	–0.142
Independence competence	–0.263	–0.258
Recreation	–0.206	–0.225
Prosocial	–0.139	–0.257
Employment	–0.215	–0.181

[a] Pairwise correlations based on all cases with scores on both variables.
[b] EM estimation using all predictor variables plus PANSS total score at long-term follow-up.

TABLE 53 *Regressions of long-term follow-up PANSS total score using all the variables for Trials 9 and 10*

Selected from			
All variables entered		**Only variables below entered**	
Age	–0.386	Age	–0.127
Longer duration of illness	0.262	Longer duration of illness	0.129
PANSS general subscale post-treatment	0.535	PANSS general subscale post-treatment	0.614
Social functioning – recreation pretreatment	–0.242	Social functioning – recreation pretreatment	–0.127
Multiple *r*	0.652		0.636
r^2	0.424		0.404
n	63		90

TABLE 54 *Stepwise regressions predicting PANSS total score at long-term follow-up from selected subsets of variables for Trials 9 and 10*

Selected from					
Pretreatment patient characteristics		**Pretreatment outcome measures**		**Post-treatment variables**	
Chlorpromazine equivalent	0.265	PANSS positive subscale pretreatment	0.592	PANSS positive subscale post-treatment	0.420
Multiple *r*	*0*.265		0.592		0.420
r^2	0.070		0.351		0.176
n	88		85		70

When the pretreatment patient characteristics and outcome measures were combined to predict the PANSS total score at long-term follow-up, the multiple *r* rose to 0.65 (see *Table 53*). To investigate further the relationship between these groups of predictors and the outcome variable, principal component analyses were performed on the two pretreatment patient characteristics, the six pretreatment outcome measures and the seven post-treatment outcome measures that correlated with the PANSS total score at long-term follow-up (as shown in *Table 51*, excluding the pre- and post-treatment PANSS total scores as these duplicate information held in the subscales). The first component accounted for 59% of the variance of the patient characteristics in the first analysis, 62% of the variance of the pretreatment tests in the second and 53% of variance of the post-treatment tests in the third. The correlations between these components and the PANSS total score at long-term follow-up were 0.31, 0.55 and 0.42. This means that these principal components were as good as the stepwise regressions reported in *Table 54* at predicting the PANSS total score at long-term follow-up and that the exact regression weightings are somewhat arbitrary. Predicting the long-term outcome factor from a regression analysis with all three components yielded a multiple *r* of 0.43, but only the post-treatment component contributed significantly to the regression.

Discussion

Higher scores on measures of symptom severity prior to randomisation and following treatment were moderately (BSI-GSI) to highly (PANSS) correlated with poorer outcome at longer-term follow-up. In addition, measures of social functioning and integration were also moderately associated with symptomatic outcome, although these effects were clearer and more consistent for post-treatment measures. In all cases poorer social functioning was associated with worse PANSS scores. The relationship between being in the CBT group and PANSS total scores at long-term follow-up approached significance ($p = 0.054$), with those in the CBT group having lower PANSS scores. The effects of missing data were small with the pairwise and EM estimated correlations being very similar. However, one important effect is that the missing data may have resulted in the effect of CBT on outcome being overestimated, as the correlation between CBT group and the PANSS total score was lower in the EM imputation.

We were able to account for a large proportion of the variance in long-term outcome ($r^2 = 0.42$). We found that younger age, longer duration of illness, higher PANSS general psychopathology scores post-treatment and lower recreational social functioning at prerandomisation were associated with worse PANSS scores at follow-up. Hence patients who do poorly over the longer term are more likely to be younger, have a longer illness duration and therefore earlier onset, experience more distress despite treatment and be less socially integrated. These findings are consistent with those of Tarrier and colleagues,[63] who found that at 3-month follow-up longer duration of illness and greater symptom severity were associated with being less likely to achieve clinically important improvements in positive symptom severity. However, unlike Tarrier and colleagues, we did not find that receipt of CBT was associated with better or worse longer term outcome.

Not only severity of positive symptoms but also worse scores on measures of social functioning were associated with poor outcome at long-term follow-up. One could argue that social functioning and integration should also be included as important targets for future studies of CBT, especially if CBT seeks to secure more long-lasting and enduring outcomes.

Chapter 10

Health economic analyses

Introduction

This chapter sets out the results from the cost-effectiveness analyses separately for the eight anxiety disorder studies (Trials 1–8) and the two psychosis studies (Trials 9 and 10).

Method

Creation of the cost variable

As noted in the section 'Identification and measurement of resource use' (p. 10), the primary costs of interest with respect to the economic evaluation are those incurred by patients following treatment. Added to this is the actual cost of the intervention received in the original study (mean cost in Trials 1–8 = £83.26; mean cost in Trials 9 and 10 was unavailable from the original trial data). Consequently, the costs in the 2-year period prior to follow-up were added to the cost of the original intervention. Obviously this approach assumes that this 2-year period is, on average, representative across the whole time period. For 17 individuals in Trials 1–8, the cost of the original intervention was unavailable whereas resource use was available for the 2 years prior to follow-up. Here the average cost of intervention for the relevant study arm and trial was applied. In three instances, the treatment arm was unavailable and consequently the average for this study was applied. As a result, full costs were then available for 361 individuals in Trials 1–8. Two of these individuals were dropped from the cost-effectiveness analysis owing to lack of availability of study arm from the original trial. Full costs were available for 94 individuals in the psychosis studies (Trials 9 and 10).

Creation of the benefit variable

The benefit variables for the cost-effectiveness analysis, namely the SF-36 physical and mental health summary component scores (SF-36 PC and SF-36 MC), were available for 342 individuals of the 359 detailed above for Trials 1–8. The 17 missing cases occurred in Trials 1–5 and 8. The approach here was to impute a value using the mean summary scores for both the physical and mental health components from the relevant treatment arm and study.

SF-36 PC and MC scores were available for 80 of the above 94 individuals in Trials 9 and 10. Partial SF-36 data were, however, available for six individuals where cost data were present. Here, SF-36 summary scores were imputed on the basis of the available SF-36 data for each individual. The approach adopted was to regress the physical and mental health summary score for the whole sample on the component items available for each individual and use predicted values to impute their data. In two cases this imputation was based on seven of the eight domains, and in no case was it based on less than three of these.

Cost-effectiveness analysis

In order to assess the relative cost-effectiveness of CBT compared with other treatments, it is necessary to compare the costs experienced by those who did and did not receive CBT with the benefits that accrued to them. The formal expression of this comparison is the incremental cost-effectiveness ratio[160] (ICER), expressed as

$$\frac{\text{Cost treatment A} - \text{cost treatment B}}{\text{Benefit treatment A} - \text{benefit treatment B}}$$

Table 55 shows the point mean estimates for costs and benefits in both treatment groups for Trials 1–8 inclusive.

Using the data presented in *Table 55* to calculate the ICER gives the following results:

- the difference in cost is £5, where CBT confers this slightly higher cost.
- the difference in benefit is 1.41 on the physical summary score and 2.55 on the mental summary score, where those who received CBT have this slightly higher health status.
- the ICER is therefore £3.55 per one point gain on the physical summary score and £2.76 per one point gain on the mental summary score.

Table 56 shows the point mean estimates for costs and benefits in both treatment groups for Trials 9 and 10.

Using the data presented in *Table 56* to calculate the ICER gives the following results:

TABLE 55 *Mean costs and benefits – CBT and non-CBT groups for Trials 1–8*

CBT group	Mean cost (£)	Mean SF-36 PC	Mean SF-36 MC
No	1469	39.45	39.21
Yes	1474	40.86	41.02

TABLE 56 *Mean costs and benefits – CBT and non-CBT groups for Trials 9 and 10*

CBT group	Mean cost (£)	Mean SF-36 PC	Mean SF-36 MC
No	8,362	43.19	42.69
Yes	8,535	42.37	38.40

- The difference in cost is £173, where CBT confers this slightly higher cost.
- The difference in benefit is 0.82 on the physical summary score and 4.29 on the mental summary score, where those who did not receive CBT have this higher health status.

However, there clearly exists some uncertainty around both sets of these point estimates. Alongside this uncertainty is the need to explore the significance of estimates, whilst retaining the mean as the summary statistic given the need to refer back to potential budgetary implications of any advocated change in policy. Methods recently developed within health economics use non-parametric approaches that recognise the skewed nature of the data presented in the section 'Clinical profiles on Brief Symptom Inventory' (p. 48) whilst retaining the mean as the summary statistic. This approach is known as bootstrapping.[161,162] These methods effectively generate estimates of uncertainty by taking repeated samples of individual level observations of the variables of interest (total cost and SF-36 summary scores). It is customary to draw 1000 bootstrapped samples. One thousand estimates of the cost differences and benefit differences are therefore generated and can be used to construct ranges within which true values will lie on 95% of occasions.

Figure 21 shows the 1000 replications for the differences in cost and differences in the SF-36 physical summary component score for Trials 1–8. Also shown are the proportion of replications that lie within each quadrant, such that in 40.1% of cases CBT treatment is associated with higher levels of cost and higher levels of benefit. *Figure 22* shows the 1000 replications for the differences in cost and differences in the SF-36 physical summary component scores for Trials 9 and 10. Also shown are the proportion of replications that lie within each quadrant, such that in 19.1% of cases CBT treatment is associated with higher levels of cost and higher levels of benefit.

Similarly, *Figure 23* shows the 1000 replications for the differences in cost and differences in the SF-36 mental summary component score for Trials 1–8. Also shown are the proportion of replications that lie within each quadrant, such that in 47.2% of cases CBT treatment is associated with higher levels of cost and higher levels of benefit.

Figure 24 shows the 1000 replications for the differences in cost and differences in the SF-36 mental summary component scores for Trials 9 and 10. Also shown are the proportion of replications that lie within each quadrant, such that in 1.4% of cases CBT treatment is associated with higher levels of cost and higher levels of benefit.

In the case of Trials 1–8, these results are clearly non-significant with respect to the overall level of cost-effectiveness when using either measure of benefit. We can, however, conclude that in 89.2% (42.0 + 47.2) of cases CBT does confer benefit with respect to mental health summary scores. In terms of cost-effectiveness, only one quadrant can be unequivocally interpreted as non cost-effective and this is where an intervention offers less benefit at greater cost. In the case of the mental health aspect of the SF-36, 93.5% (42.0 + 47.2 + 4.3) of observations do not lie within this quadrant. Similarly based conclusions can also be drawn from *Figure 21*. This does not indicate, however, whether CBT **is** cost effective. Given the non-significant differences in outcomes outlined in Chapter 8 and the very minimal contribution of CBT costs to the overall costs of these subjects, analysis has not been undertaken to explore

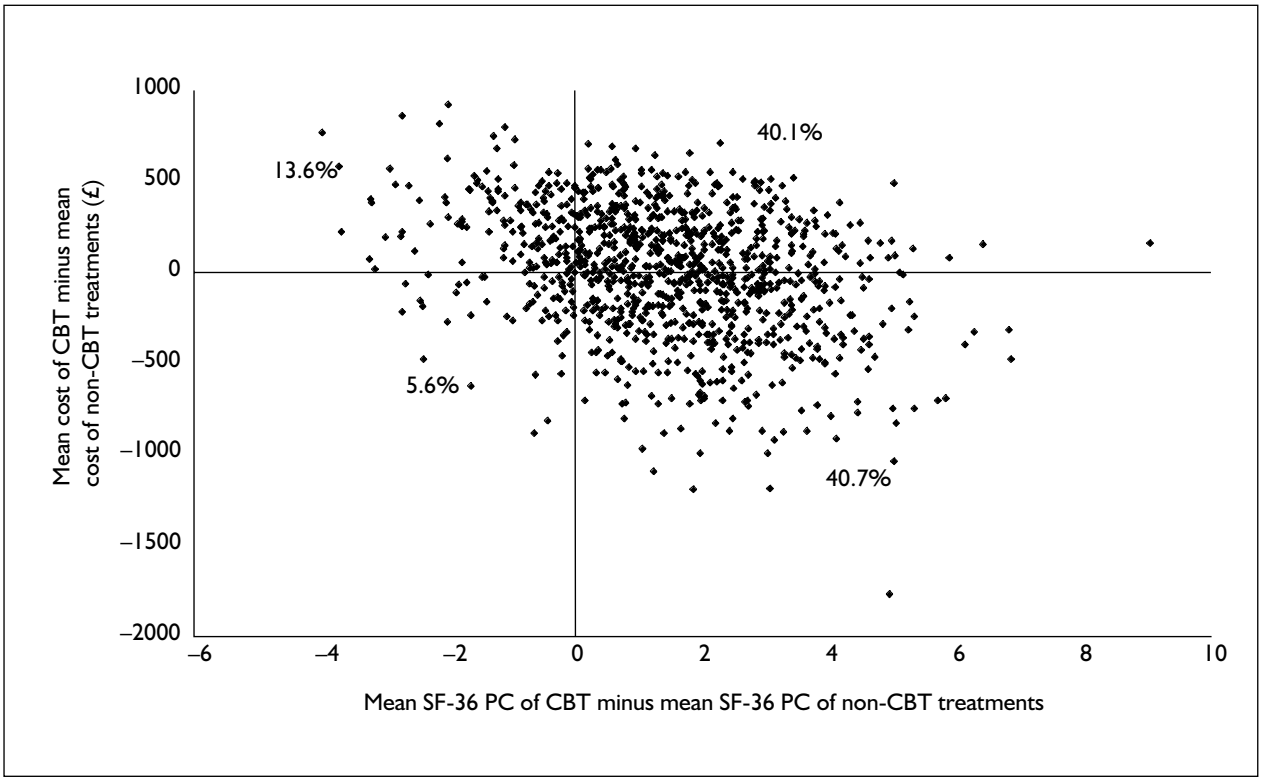


FIGURE 21 *Cost-effectiveness analysis for Trials 1–8 – bootstrapped replications* (n = *1000). Mean cost per point difference on SF-36 PC = £3.55*

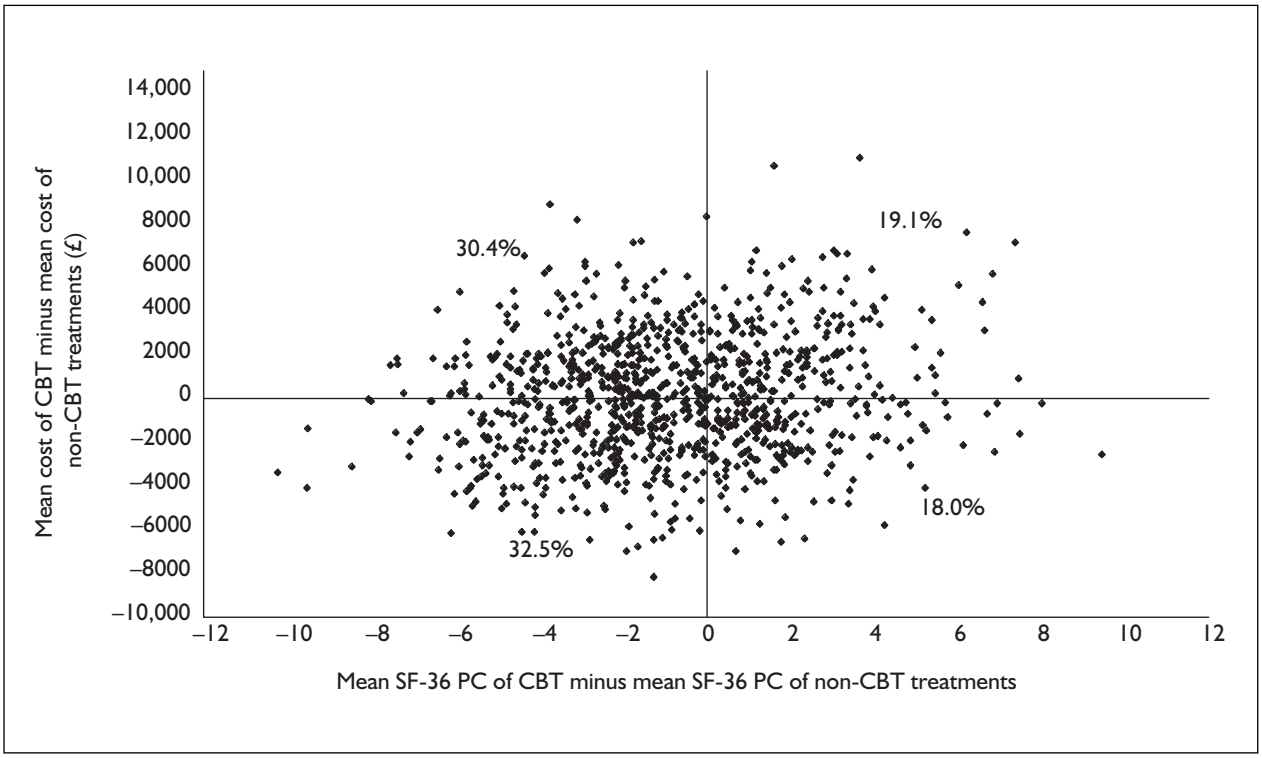


FIGURE 22 *Cost-effectiveness analysis for Trials 9 and 10 for SF-36 PC – bootstrapped replications* (n = *1000)*

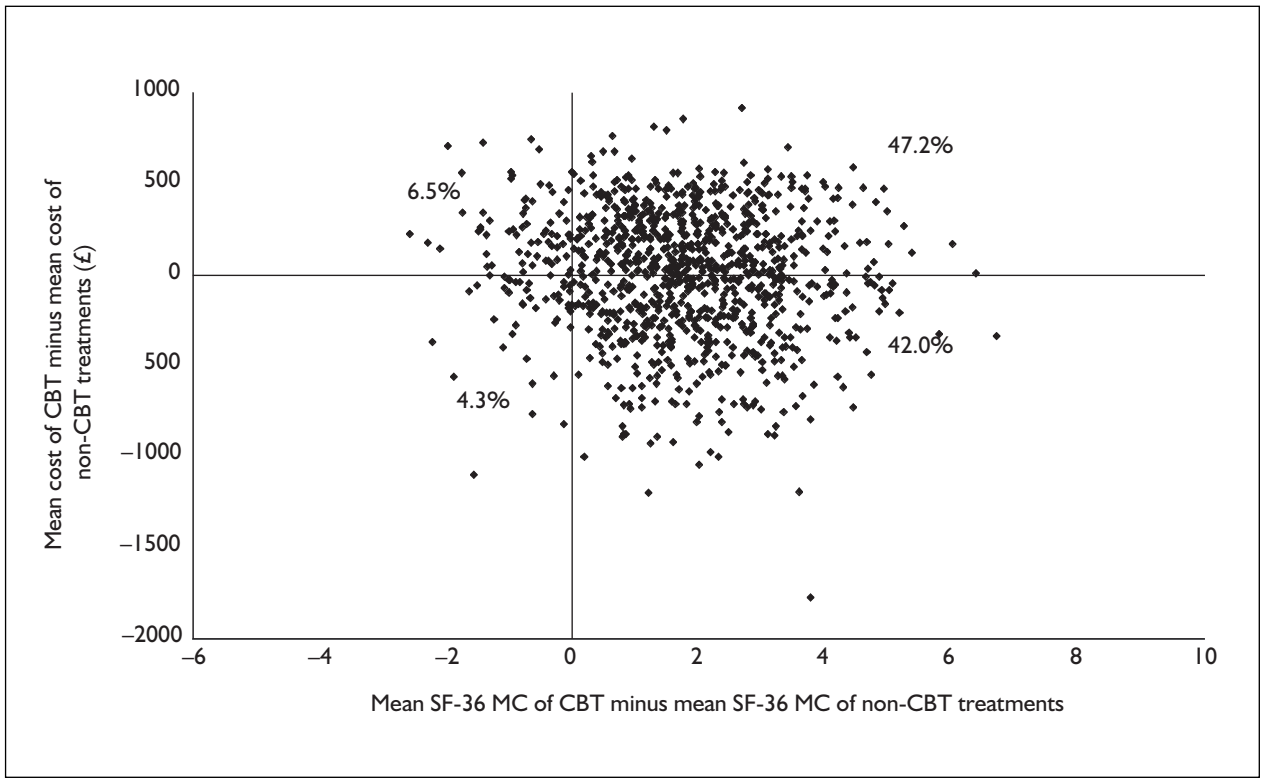


FIGURE 23 *Cost-effectiveness analysis for Trials 1–8 – bootstrapped replications* (n = *1000). Mean cost per point difference on SF-36 MC = £2.76.*

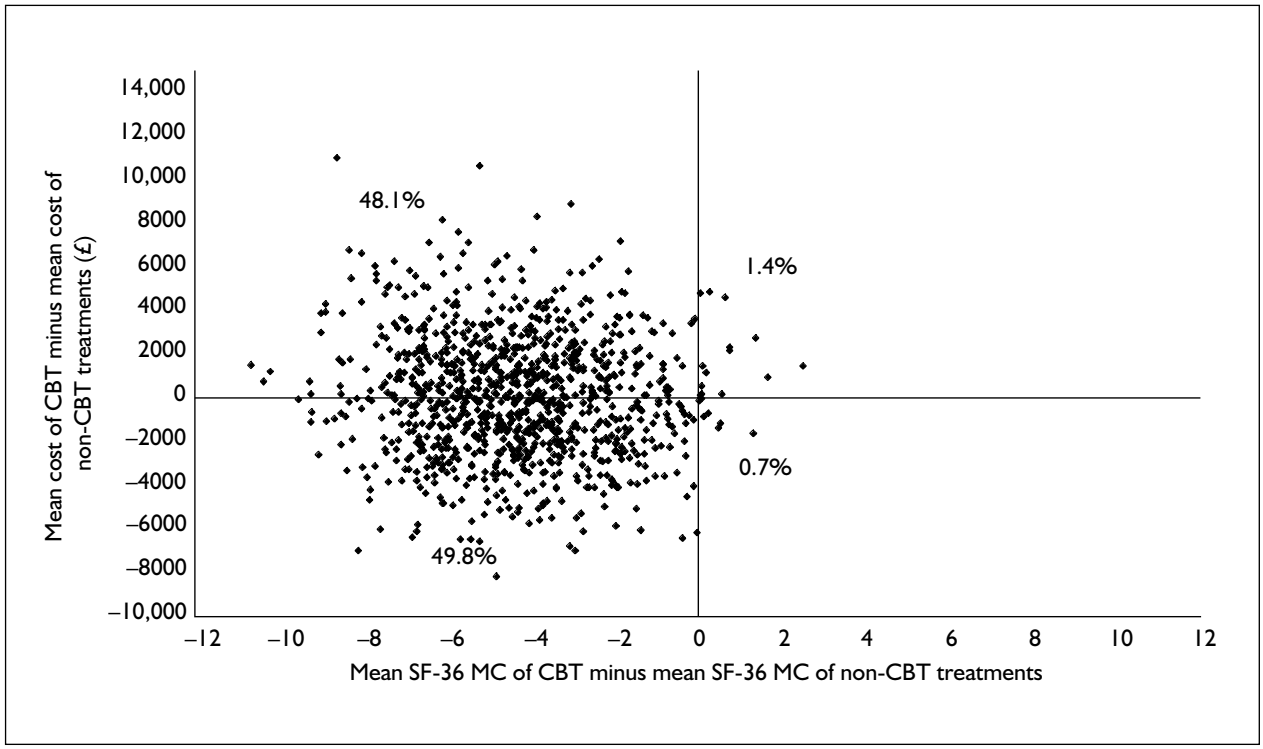


FIGURE 24 *Cost-effectiveness analysis Trials 9 and 10 for SF-36 MC – bootstrapped replications* (n = *1000)*

TABLE 57 *Mean costs and benefits – CBT completers and others for Trials 1–8*

Group	Mean cost (£)	Mean SF-36 PC	Mean SF-36 MC
Others	1502	39.87	38.97
CBT	1457	40.90	41.52

differences associated with the level of intensity of CBT.

Considering Trials 9 and 10, these results are also non-significant with respect to the overall levels of cost-effectiveness when using either measure of benefit. We can conclude, however, that in 97.9% of cases (48.1 + 49.8) of cases, CBT does not confer benefit with respect to mental health summary scores. In terms of cost-effectiveness, only one quadrant can be unequivocally interpreted as non-cost-effective and this is where an intervention offers less benefit at a greater cost. In the case of the mental health aspect of the SF-36, 51.9% (49.8 + 0.7 +1.4) of observations do **not** lie within this quadrant. Similarly based conclusions can be drawn with respect to the physical health results.

In addition to the above intention-to-treat analysis, an exploration of the cost-effectiveness was also carried out comparing those who completed CBT with an aggregate group of those who were not allocated to CBT, those who dropped out of CBT and those who did not commence CBT at all. This analysis was carried out solely with respect to Trials 1–8, given the higher level of uncertainty associated with these results.

Using the data presented in *Table 57* to calculate the ICER gives the following results:

- The difference in cost is £45, where the 'others' group incurs this slightly higher cost.
- The difference in benefit is 1.03 on the physical summary score and 2.55 on the mental summary score, where those who received CBT have this higher level of health status.

Therefore the result is that, in the anxiety disorder studies (Trials 1–8), CBT completers have lower costs and higher levels of benefit and that CBT should therefore be implemented for those who will complete treatment. However, again there is considerable uncertainty around these estimates and bootstrapping was therefore pursued to estimate this level of uncertainty.

As with the intention-to-treat analysis, *Figure 25* shows the 1000 replications for the differences in costs and the differences in the physical summary component score, and also the proportion of replications that lie within each quadrant for Trials 1–8. *Figure 26* shows the same data for the mental health summary score. A general observation regarding both scatterplots is that the data are less distributed across the plane. Moreover, these results are more favourable to CBT than the intention-to-treat analysis, particularly in the case of mental health status.

A further technique is available for summarising the cost-effectiveness planes presented in *Figures 21–26*. Cost-effectiveness acceptability curves (CEACs) show the probability that a particular intervention is cost-effective, over a range of possible monetary values decision-makers may be willing to pay for a range of improvements in health outcomes.[163,164] They are therefore a simple summary of the incremental costs and benefits and the level of uncertainty associated with these point estimates. The curve cuts the *y*-axis at the point where decision-makers are unwilling to pay for health benefits and represents the proportion of replications where the costs in the CBT arm are less than those in the non-CBT arm. Where the decision-maker is willing to pay large amounts for health benefits, the CEAC plateaus to a value representing the proportion of replications in which the CBT arm has higher benefits than the non-CBT arm.

Figure 27 shows the CEACs for Trials 1–8. All of the curves increase monotonically, indicating that, in the majority of replications, the CBT arm offers higher costs and higher benefits. The evidence is most favourable to CBT for the SF-36 MC in treatment completers: here if a decision-maker was willing to pay £50 for a one unit increase in this health benefit measure, the probability of CBT being cost-effective is 0.706. If the decision-maker was prepared to pay £350, the probability would increase to 0.959.

Figure 28 shows the CEAC for Trials 9 and 10. Here, the SF-36 MC curve declines monotonically,

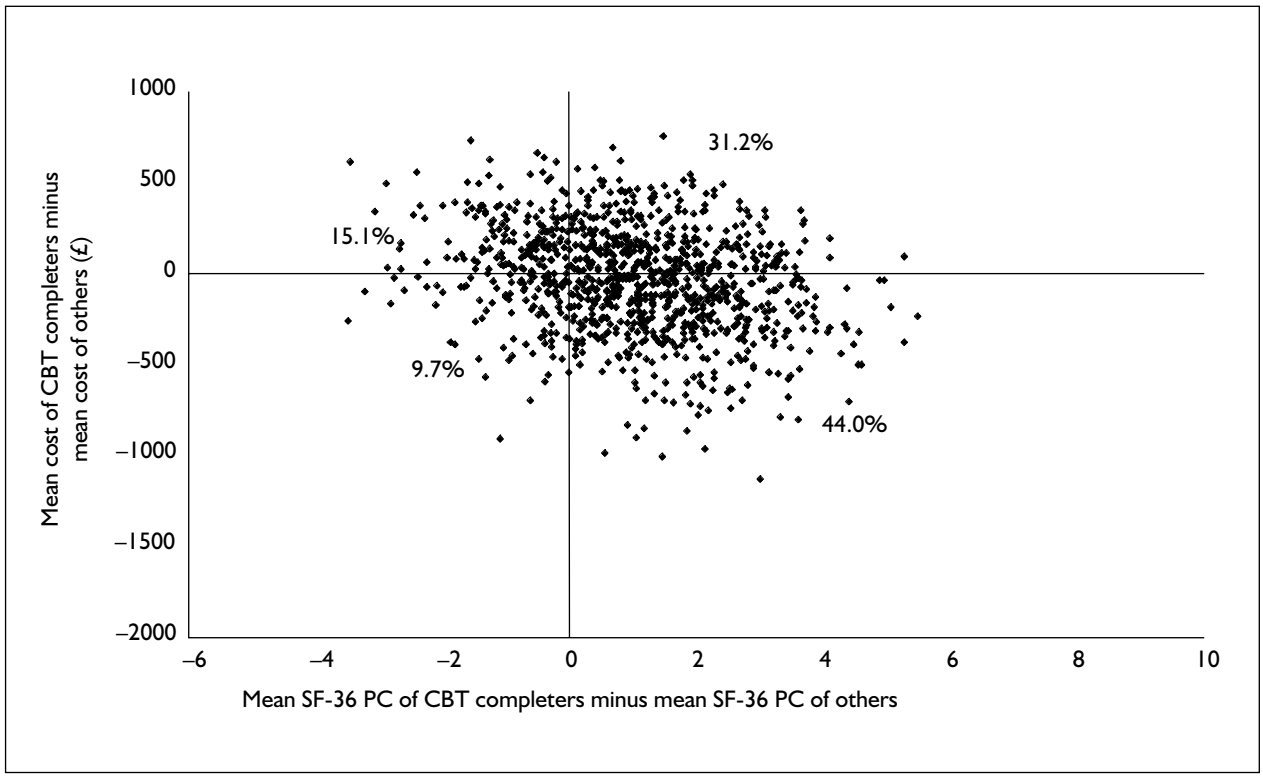


FIGURE 25 *Cost-effectiveness analysis for Trials 1–8 – bootstrapped replications* (n = *1000). Mean cost per point difference on SF-36 PC (completers vs others)*

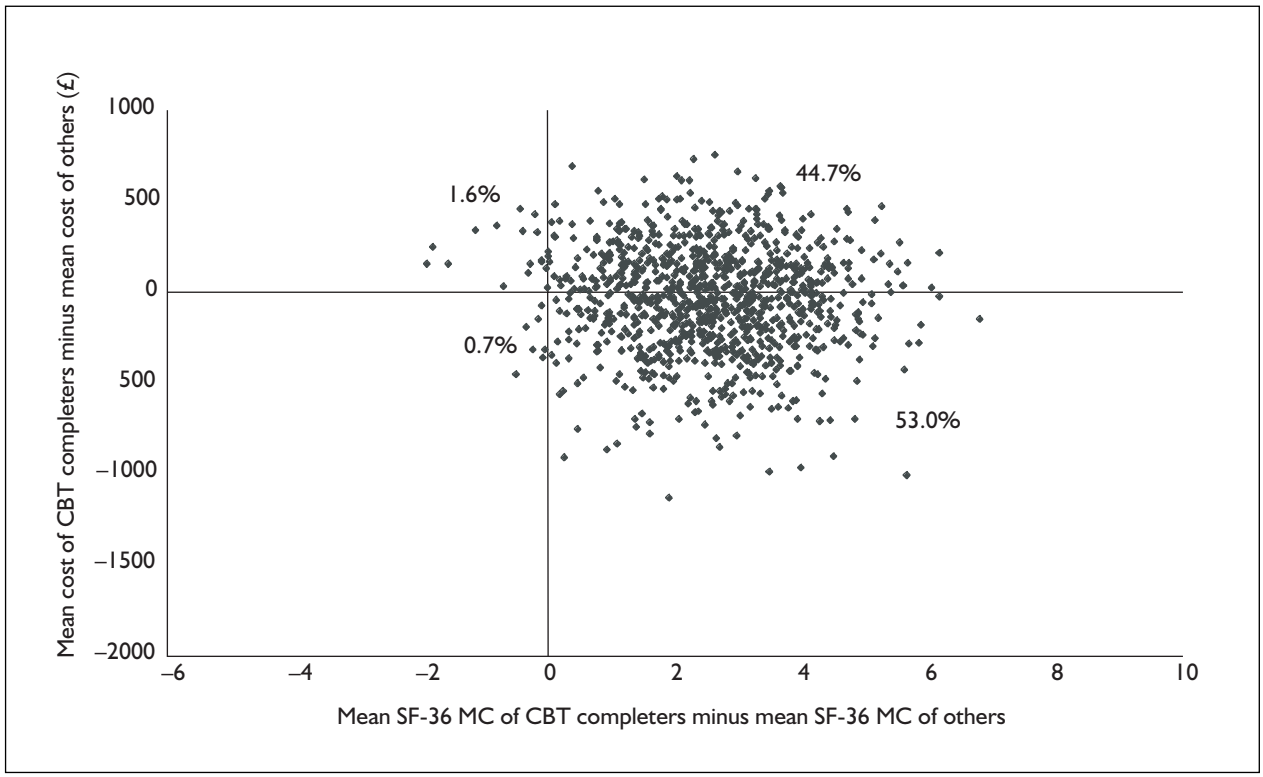


FIGURE 26 *Cost-effectiveness analysis for Trials 1–8 – bootstrapped replications* (n = *1000). Mean cost per point difference on SF-36 MC (completers vs others)*

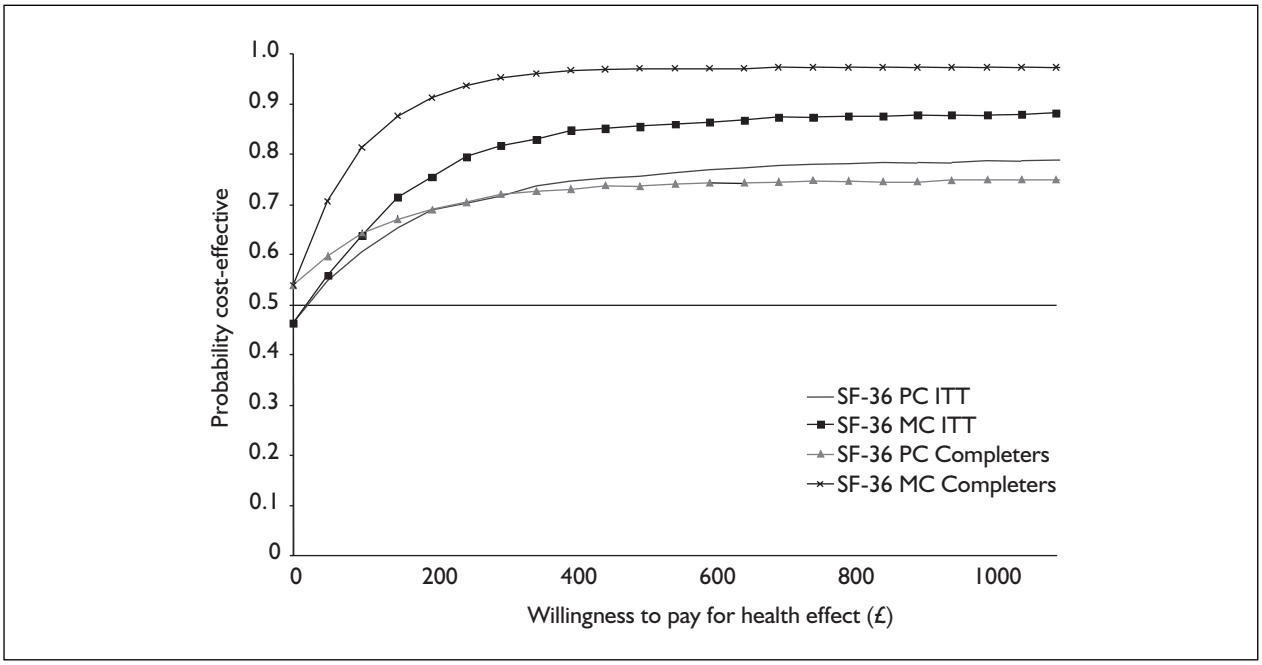


FIGURE 27 *CEACs – Trials 1–8, SF-36 PC and SF-36 MC [intention-to-treat (ITT) and completers vs non-completers analysis]*

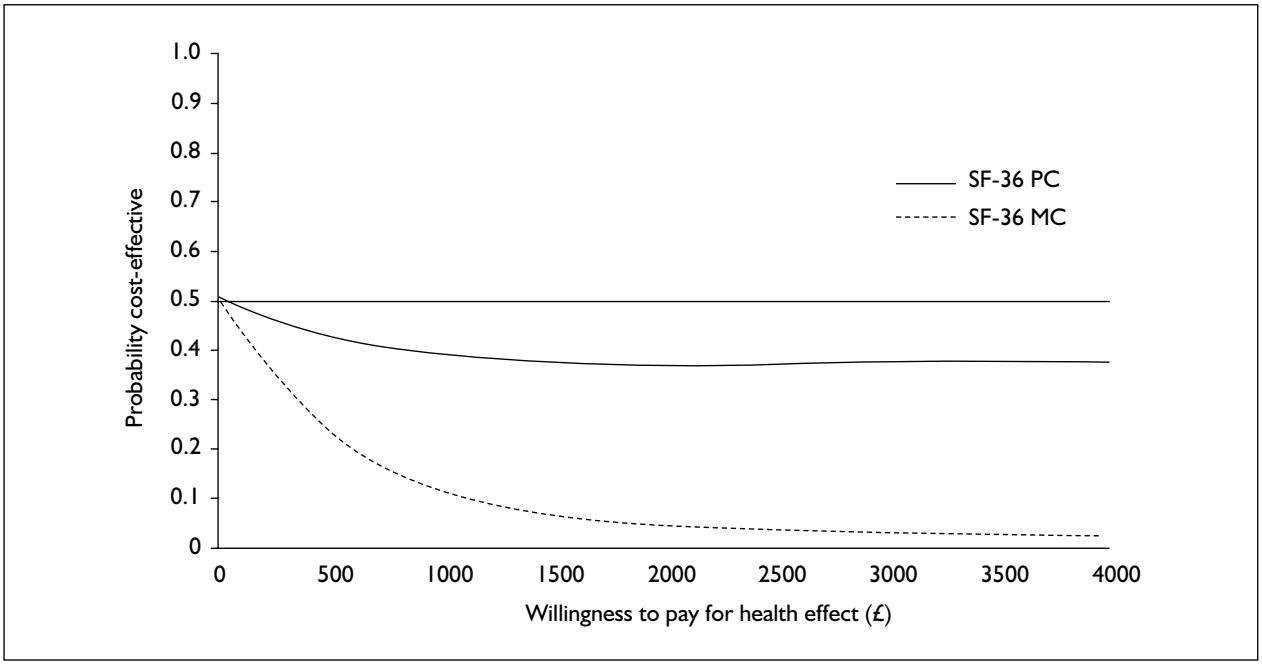


FIGURE 28 *CEACs – Trials 9 and 10, SF-36 PC and SF-36 MC (intention-to-treat analysis)*

indicating that in the majority of replications, the CBT arm offers lower levels of benefit. For SF-36 PC, the results are more equivocal, but are generally unfavourable.

Conclusions

The cost-effectiveness analysis associated with Trials 1–8, although not producing significant results in terms of decision-making around the availability of CBT, has complemented the outcome analysis in a number of ways. First, CBT confers higher levels of benefit, as measured by the SF-36 PC score, in 81% of cases. Moreover, when combined with measures of resource use, in 41% of cases using this benefit measure, there is a higher level of benefit and a lower level of resource use. When considering the evidence with respect to the SF-36 MC score, the results are,

unsurprisingly, stronger. In 89% CBT confers a higher level of benefit and in 42% of cases this is a higher level of benefit with lower resource use. As noted earlier, however, this does not provide an indication that CBT is cost-effective.

The analysis of Trials 1–8 was further explored by comparing those who completed CBT with all others within the trial – a *post hoc* construction of a group who received other forms of treatment. In this analysis, CBT was associated with higher levels of physical health in 75% of cases. In 44% of cases the result with respect to cost-effectiveness was unequivocally favourable, namely higher benefit and lower cost. With respect to mental health, the CBT completer group was associated with higher levels of benefit in 98% of cases and this result was a combination of higher benefit and lower cost in 53% of cases. Further analysis using CEACs indicates that CBT would be a cost-effective intervention for this selected group with a >95% probability if decision-makers are willing to pay more than £300 for a one-unit increase in the SF-36 MC.

There are two possible confounding factors related to this analysis, whose impact merits further exploration. The first of these is the extent to which the original randomisation procedure was valid. As noted in the section 'Comparison of aggregate CBT on non-CBT groups on pretreatment characteristics for anxiety disorder studies (Trials 1–8)' (p. 36), a reanalysis of baseline characteristics revealed very few significant differences between the CBT and non-CBT groups. One notable significant result was that those who received CBT tended to be more affluent. Given the well-established relationship between morbidity and response to treatment, the impact of this bias would be to overestimate the cost-effectiveness of CBT. A further issue is the length of follow-up and whether this differs systematically across CBT and non-CBT groups. As noted in the section referred to above, the length of time since follow-up in the CBT group was significantly shorter (6.5 versus 8.0 years) than the non-CBT group. Given the well-established relationship between ageing and healthcare resource use, again this would overestimate the cost-effectiveness of CBT.

The analysis in Trials 9 and 10 was less favourable. The analysis demonstrated that CBT was associated with higher levels of physical benefit in 37% of cases. In 18% of instances this was a higher level of benefit and lower levels of resource use. The results were even less favourable when considering mental health, with only 2% of instances demonstrating higher levels of benefit in the CBT group. A fraction of replications, 1%, had higher benefit and lower cost on this benefit measure.

Clearly, the results associated with Trials 9 and 10 were based on much smaller numbers; however, it is also worth repeating why this group were analysed separately. In addition to being a distinct diagnostic group, their levels of resource use were significantly higher –£1379 in Trials 1–8 and £9725 in Trials 9 and 10 for the 2-year pre-follow-up period. This difference in cost is overwhelmingly driven by inpatient stays – a relatively common feature of care amongst participants in Trials 9 and 10 and relatively rare for Trials 1–8. Possibly unsurprisingly, it is extremely difficult within the framework of cost-effectiveness analysis to assess the impact of a relatively brief intervention when considered from a long-term perspective; this applies across all 10 trials. Assuming that all other differences are adequately captured by randomisation procedures is perhaps somewhat heroic. More detailed analysis, including factors that we know to influence levels of resource use such as clinical severity, poverty, proximity to services, changes in technology (e.g. the introduction of atypical antipsychotics) and age, may in the future provide a more meaningful assessment of CBT interventions in a range of diagnostic groupings. Additionally, the breadth of data captured across a cohort of individuals experiencing a range of mental health diagnoses over such a long time period offers the potential for use as a panel study to investigate a range of issues associated with health and health care provision amongst these client groups.

Chapter 11

Summary of findings

Chapters 2–5: Methodology

Anxiety disorder studies (Trials 1–8)

- A total of 861 people entered the original trials, of whom 640 completed treatment and 221 either dropped out or failed to attend. An attempt was made to follow-up all 861 original entrants, of whom 85 were unavailable to contact. A total of 396 people agreed to participate in the long-term follow-up study (51% of available participants, 46% of original trial entrants). Length of follow-up ranged from 2 to 14 years post-treatment.
- In comparison with the census data, a higher proportion of the long-term follow-up sample were female, married and not working owing to ill-health, and a smaller proportion were single, in gainful employment, from the professional socioeconomic group and educated to degree level.
- In terms of the primary aim of the study, which was to detect differences in outcome between those patients who had received CBT (n = 230) and those who had not (n = 81), the power of a two-sample t-test to detect large, medium and small effects (0.8 to 0.5 and 0.2 of an SD) is virtually 1, 0.97 and 0.34, respectively. Hence a result which failed to reach the 0.05 level of significance is inconsistent with large and medium CBT effects but does not rule out the possibility of a small, undetected effect of CBT on the long-term outcome measures.
- The main clinician-rated outcome measures administered at long-term follow-up were ADIS-IV to obtain diagnostic status and co-morbidity, CGS (0–8) and HAM-A. These were administered by a research psychologist who was blind to all information on participants' original clinical status or allocation to treatment conditions. The main patient-rated measures administered at long-term follow-up were BSI, SF-36 II measure of health status, CGI (1–7), PANAS and the trait version of the STAI-T.

Psychosis studies (Trials 9 and10)

- A total of 210 people entered the original trials, of whom 191 completed treatment and 19 either dropped out or failed to attend. An attempt was made to follow-up all 210 original entrants, of whom 40 were unavailable to contact. Ninety-three people agreed to participate in the long-term follow-up study (55% of available participants, 43% of original trial entrants). Length of follow-up was between 2 and 6 years post-treatment.
- In comparison with the census data, a higher proportion of the long-term follow-up sample were male and single and a smaller proportion were married.
- In terms of the primary aim of the study, which was to detect differences in outcome between those patients who had received CBT (n = 44) and those who had not (n = 49), the power of a two-sample t-test to detect large, medium and small effects (0.8, 0.5 and 0.2 of an SD) is 0.97, 0.66 and 0.16, respectively. The data set for the two psychosis studies would give 80% power of detecting an effect size of 0.59.
- The main clinician-rated outcome measures administered at long-term follow-up were the PANSS, to obtain a picture of overall symptomatology, and CGS (0–8). These were administered by researchers who were blind to information on participants' original clinical status or allocation to treatment conditions. The main patient-rated measures administered at long-term follow-up were BSI, SF-36 II measure of health status and CGI (1–7).

Chapter 6: Overall outcome

Anxiety disorder studies (Trials 1–8)

Diagnostic status

- Over half of the participants (52%) still had at least one clinical diagnosis at long-term follow-up [this proportion was similar across GAD studies (52%) and panic disorder studies (48%) but higher in the PTSD trial (74%)].
- Average levels of co-morbidity for those participants with at least one diagnosis were relatively high (2.7, SD 1.7).
- For the whole sample the most frequently occurring diagnoses were GAD (27%), agoraphobia (26%), panic disorder (23%) and depression (22%).

- GAD participants were most likely to have GAD at follow-up (34%) and PTSD participants were most likely to have PTSD at follow-up (55%). Panic disorder participants, however, were marginally more likely to have agoraphobia (29%) than panic disorder (26%), probably because of extensive use of avoidance as a coping mechanism.

Clinical global severity

- Mean (SD) CGS scores were 3.3 (1.8) for the whole sample where a score of 4, the threshold for diagnosis, is defined as definitely disturbing/disabling.
- Very few participants (18%) had none or only mild symptoms and a significant proportion (31%) had subthreshold symptoms of at least moderate severity.

Recovery status

- The Jacobson method for calculating clinically significant change was applied to the HAM-A, used by all but one of the studies at short and long-term follow-up. This gave an overall recovery rate of 80% at short-term follow-up and 52% at long-term follow-up, the latter figure being made up of 49% who had maintained recovery and only 3% who had achieved it during the follow-up period.
- Jacobson criteria for clinically significant change were used to calculate a cut-off point for the BSI-GSI and the STAI-T. This gave an overall recovery rate at long-term follow-up of 57% on the BSI-GSI and 44% on the STAI-T.

Clinical profiles on BSI subscales

- For the participants as a whole, self-reported symptoms on the BSI subscales are markedly more severe than those reported by normative samples.
- For those participants with at least one diagnosis the severity of self-reported symptoms is generally high and, for most of the eight subscales, is well above the cut-off points for caseness.
- Although raw scores for male participants are generally lower than those for female participants, they reported higher levels of symptom severity when compared with their peers, so that their *T*-scores tend to be higher than those of female participants.

Health status as measured by the SF-36

- For the participants as a whole, the severity of health status scores were equivalent to those scored by the lowest 25% of the general population.
- For those participants with at least one diagnosis, the subscale scores (with the exception of pain) were equivalent to those scored by the lowest 10% of the general population and the emotional role subscale corresponded to the lowest 2.1%.
- There were no significant differences on any of the health status subscales for the GAD and panic disorder patients who had mean scores equivalent to the worst 27% of the general population. PTSD participants, however, had significantly worse scores than either GAD or panic disorder patients with mean scores equivalent to the worst 14% of the general population on all subscales and the worst 4% for physical and emotional role.

Patient ratings of overall improvement

- Despite elevated symptom levels and generally poor health across the follow-up sample, the majority of participants (80%) felt that they had improved to some degree since the original trial (48% reporting marked improvement, 32% reporting some improvement, 10% reporting no change and only 10% stating that they were worse).

Amount of interim treatment reported during follow-up period

- Only 36% of participants reported receiving no interim treatment for psychosis over the follow-up period. The remaining 64% reported receiving varying degrees of further professional help with 19% receiving almost constant treatment. Participants with at least one diagnosis were more likely to have had more treatment than those with no diagnosis.

Use of healthcare resources from case note review

- Data on healthcare usage for the 2-year period prior to entering the original trials and the 2-year period prior to the follow-up interview were collected on 366 (92.4%) of the 396 participants in the follow-up study.
- Total costs per patient over the 4-year period ranged from nearly £40 to nearly £44,000.
- A breakdown of cost by diagnostic category revealed particularly high costs in the PTSD sample.
- There was a 40% real increase in the value of healthcare costs over the two periods, over half of which can be attributed to an increase in prescribing.
- Whereas 14 individuals experienced a reduction in resource use of £2000 or more, a total of 39 individuals experienced a rise in resource use of a similar amount.

Psychosis studies (Trials 9 and 10) and comparison with anxiety disorder studies

Clinical global severity

- Mean (SD) CGS scores were 4.0 (1.7) for the whole sample where a score of 4, the threshold for diagnosis, is defined as definitely disturbing/disabling.
- Very few participants (19%) had none or only mild symptoms and around one-fifth (22%) had symptoms which were severely distressing or disabling.
- CGS scores for the psychosis patients were significantly worse than those of the anxiety patients.

Recovery status

- Jacobson criterion (c) for clinically significant change was used to calculate a cut-off point for the BSI-GSI. Nearly two-fifths (38%) of the psychosis patients had achieved this cut-off at long-term follow-up. However, only 12% of patients had also improved by an amount greater than or equal to the reliable change index.
- When the same Jacobson criteria were applied to the psychosis and anxiety patients, there were no significant differences between groups.
- Only 10% of patients achieved a 25% reduction in total PANSS scores from pretreatment to long-term follow-up, and only 1% achieved a 50% reduction.

Clinical profiles on BSI subscales

- For both male and female patients, self-reported symptoms on the BSI subscales were markedly more severe than those reported by normative samples, and were above the cut-off points for caseness on all subscales except hostility.
- Although raw scores of male participants were generally lower than those of female participants, their *T*-scores in comparison to population norms tended to be higher than those of female participants.
- BSI scores for the male psychosis patients were significantly lower (i.e. better) than scores for the anxiety patients as a whole on the somatisation and hostility subscales. There were no other significant differences between the anxiety and psychosis patients on the BSI.

Health status as measured by the SF-36

- With the exception of the pain subscale (lowest 32%), the psychosis patients as a whole had health status scores which were equivalent to those scored by the lowest 20% of the general population.
- There were no significant differences on any of the health status subscales between the psychosis patients and either the GAD or panic disorder patients. The psychosis patients had significantly better scores than the PTSD patients on the physical role, social functioning and pain subscales of the SF-36.

Patient ratings of overall improvement

- Despite elevated symptom levels and generally poor health across the follow-up sample, the majority of participants (81%) felt that they had improved to some degree since the original trial (with 27% reporting marked improvement, 54% reporting some improvement, 9% reporting no change and only 10% stating that they were worse).
- Overall, CGI scores for the psychosis patients were significantly worse than those for the anxiety patients, with the main difference being in reports of marked improvement (27% of psychosis patients versus 47% of anxiety patients).

Amount of interim treatment reported during follow-up period

- Only 1% of participants reported receiving no interim treatment for psychosis over the follow-up period. Nearly all (93%) reported almost constant treatment, and the remaining 6% reporting receiving some interim treatment.
- Levels of interim treatment reported by the psychosis patients were significantly higher than the anxiety disorder patients.

Use of healthcare resources from case note review

- Data on healthcare usage for the 2-year period prior to entering the original trials and the 2-year period prior to the follow-up interview were collected on 94 study participants.
- Total costs per patient over the 4-year period ranged from £664 to over £250,000.
- There was a non-significant fall in the value of healthcare costs over the two periods.
- The majority of costs in this client group was accounted for by inpatient stays. Prescribing costs increased by 102% over the time periods.
- Fifty-six individuals experienced a reduction in resource use whereas 38 experienced a rise in resource use.

Chapter 7: Efficacy of CBT versus non-CBT

Anxiety disorder studies (Trials 1–8)

- A comparison of the seven outcome measures common to all studies at long-term follow-up was made between CBT and non-CBT

conditions in the original trials. A MANOVA by CBT versus non-CBT treatment on the seven outcome measures showed a significant effect of CBT group with those receiving CBT having more favourable scores on all measures at long-term follow-up.
- In general, differences between the CBT and non-CBT groups were much smaller at long-term follow-up than at post-treatment.
- No significant differences for the whole sample were found between the CBT and non-CBT groups with regard to diagnostic status or clinically significant change, although a significant difference was found in respect of the Jacobson cut-off point for clinically significant change.
- Patients who had achieved clinically significant change following participation in the original trials, regardless of treatment modality, were more likely to maintain clinically significant change from post-treatment to long-term follow-up.
- Patients who failed to respond to initial treatment, whether CBT or non-CBT, were very unlikely to have achieved clinically significant change at long-term follow-up.
- Patients in the CBT group reported treatment as being significantly more helpful to them over the long term and were significantly more likely to have used what they had learnt and to attribute any improvement to the treatment received.
- No significant differences were found between CBT and non-CBT groups with respect to the amount of interim treatment received since the original trial or degree of hopefulness of coping in the future.

Psychosis studies (Trials 9 and 10)

- A comparison of the eight outcome measures common to both studies at long-term follow-up was made between the CBT and non-CBT conditions in the original trials. There were individual effects on the PANSS negative symptom subscale and the SF-36 MC, with those in the CBT group having lower scores on both measures (i.e. more favourable PANSS scores and less favourable SF-36 scores). A MANOVA by CBT versus non-CBT treatment on the eight outcome measures just failed to reach significance.
- A MANOVA on the three PANSS subscales at long-term follow-up reached significance, with the CBT group having more favourable scores than the non-CBT group. However, pretreatment differences on the PANSS may be partly accountable for this finding.
- No significant differences were found between the CBT and non-CBT groups with regard to diagnostic status or clinically significant change.
- Very few patients maintained a 25% reduction in PANSS scores from post-treatment to long-term follow-up, regardless of treatment modality.
- Patients in the CBT group were significantly more likely to have a good memory of treatment than those in the non-CBT group. There were no differences between the CBT and non-CBT groups with respect to the other views of treatment variables.

Chapter 8: Efficacy of different intensities of CBT for GAD and panic disorder

Standard versus high contact CBT in GAD (Trials 2 and 7)

- Standard contact CBT conditions (n = 34) consisted of an average of 8.2 therapy sessions and was compared with high contact CBT conditions (n = 22), which consisted of an average of 12.9 sessions. Outcome measures used were the HAM-A, CGS, CGI, BSI-GSI, PANAS and SF-36.
- No significant differences were found between standard and high contact CBT on any of the outcome measures at long-term follow-up.
- A comparison of clinical status (diagnosis, clinically significant change) between standard and high contact CBT found that standard CBT had consistently more favourable results than high contact CBT but none of the differences were significant.
- For those patients who reported being better, significantly more in the Standard CBT group gave the trial treatment as their main reason for being better than patients in the high contact CBT group. There were no significant differences between the two groups with regard to any of the other 'views of treatment' variables.

Low contact versus standard contact CBT in panic disorder (Trials 4 and 5)

- Low contact CBT conditions (n = 50) consisted of an average of 4.4 therapy sessions and were compared with standard contact CBT conditions (n = 25), which consisted of an average of 7.0 sessions. Outcome measures used were the HAM-A, CGS, CGI, SRT, BSI-GSI, PANAS, SF-36 and FQ-Agora.
- No significant differences were found between low contact and standard contact CBT on any of the outcome measures at long-term follow-up.
- A comparison of clinical status (diagnosis, clinically significant change) between low contact and standard contact CBT found that

low contact CBT had consistently more favourable results than standard contact CBT but none of the differences were significant.
- When asked how they were with regard to anxiety at long-term follow-up, significantly more in the standard contact CBT group reported being worse than patients in the low contact CBT. There were no significant differences between the two groups with regard to any of the other 'views of treatment' variables.

Chapter 9: Predictors of long-term outcome

Anxiety disorder studies (Trials 1–8)

Predictor variables

- The eight anxiety disorder studies used a variety of different measures of clinical severity at the time of the original trial and composite measures of anxiety symptoms, depression symptoms and clinical global severity were constructed in order to maximise sample size in the prediction equations. These variables were available at pretreatment and post-treatment. Composite measures of social adjustment were constructed for the same reason.
- Additional predictor variables covered demographic information (job status, marital status, gender, age and level of social deprivation), clinical status at the time of the original trial (concurrent psychotropic medication, definite avoidance behaviour and duration of disorder), completion of treatment, receipt of CBT and amount of self-reported treatment over the follow-up period.

Dependent variables at long-term outcome

- Three dependent measures were used to define outcome at long-term follow-up: a composite long-term outcome factor constructed from the main outcome measures, a binary measure of clinically significant change using Jacobson criteria and a binary measure of receiving any clinical diagnosis.

Significant correlations of patient characteristics with long-term outcome

- Significant correlations with all three dependent measures were found for concurrent psychotropic medication, marital status and social deprivation, of which the most important predictors were employment status (–0.29 to –0.35) and social deprivation scores (0.19 to 0.26). Duration of current episode showed a significant correlation with the long-term outcome factor and diagnostic status but not with the Jacobson criterion for clinically significant change.

Significant correlations of composite measures with long-term outcome

- At pretreatment the composite measures of social adjustment, self-rated depression and anxiety, Hamilton-rated anxiety and clinical global severity all showed significant correlations with the long-term outcome factor (0.25 to 0.42), the Jacobson criterion of clinically significant change (–0.22 to –0.33) and diagnostic status (0.24 to 0.32).

Significant correlations of post-treatment measures with long-term outcome

- Composite measures of self-rated depression and anxiety, Hamilton-rated anxiety and clinical global severity at post-treatment all showed significant correlations with the long-term outcome factor (0.39 to 0.43), the Jacobson criterion of clinically significant change (–0.25 to –0.28) and diagnostic status (0.31 to 0.36) with the same general pattern as at pretreatment.
- Completion of treatment, irrespective of modality, was significantly correlated with the long-term outcome factor (–0.20) and diagnostic status (–0.17). Receipt of treatment with CBT correlated significantly with the long-term outcome factor (–0.12) but not with the Jacobson criterion for clinically significant change or with diagnostic status.
- Amount of self-reported interim treatment, rated retrospectively at long-term follow-up, showed the highest correlations of all with the long-term outcome factor (0.48), Jacobson criterion for clinically significant change (–0.37) and diagnostic status (0.44). The more treatment received over the follow-up period the worse was the outcome.

Analyses of the effects of missing data

- Possible effects of missing data on the correlations between scores on the long-term outcome factor and other measures were examined by comparing different estimates of the correlations. In general, the effects of missing data were small but may have led to a slight underestimate of the size of correlations between psychometric measures and long-term outcome and an overestimate of the effects of CBT on outcome.

Regression analyses

- Stepwise regressions predicting the long-term outcome factor from all predictor variables available at post-treatment gave a multiple r of

0.61 explaining about 38% of the variance. Pretreatment Hamilton-rated anxiety, post-treatment employment status, post-treatment composite depression scores, post-treatment CGS and pretreatment social deprivation scores all contributed to the prediction. When amount of interim treatment was added to the solution the multiple *r* rose to 0.69 and the composite social adjustment scores also contributed to the solution.

- Stepwise logistic regressions predicting Jacobson criteria for clinically significant change from pre- and post-treatment variables explained 20% of the variance with employment status and clinical global severity pre- and post-treatment contributing to the solution. When the interim treatment variable was added the predicted variance rose to 27%, equivalent to a multiple *r* of 0.52, with employment status and pretreatment composite social adjustment contributing to the solution.
- Stepwise logistic regressions predicting any diagnosis at long-term follow-up from pre- and post-treatment variables explained 28% of the variance with pretreatment social deprivation, employment status, pretreatment social adjustment, pretreatment Hamilton-rated anxiety and post-treatment clinical global severity all contributing to the solution. When interim treatment was added the predicted variance rose to 34%.

GAD studies using quality of therapeutic alliance as a predictor (Trials 2 and 7)

- Trials 2 and 7 both used the same measures of the complexity and severity of presenting problems at pretreatment and also the quality of the therapeutic alliance and initial response to treatment. This permitted a test of two hypotheses derived from Trial 7 in which outcome at post-treatment was better predicted by therapeutic alliance variables than by pretreatment complexity and severity of problems and outcome at 6-month follow-up was better predicted by pretreatment complexity and severity of problems than by therapeutic alliance variables.
- Both therapeutic alliance and complexity and severity of problems were moderately related to outcome at post-treatment, whereas only complexity and severity of problems was significantly related to long-term outcome.

Psychosis studies (Trials 9–10)

Predictor variables

- Level of symptomatology and functioning at pre- and post-treatment, as measured by the PANSS, BSI-GSI and social functioning scales, were used as predictor variables.
- Additional predictor variables covered demographic information (gender, age and level of social deprivation), clinical status at the time of the original trial [concurrent medication (chlorpromazine equivalent) and duration of disorder], completion of treatment and receipt of CBT.

Dependent variables at long-term outcome

- The dependent measure used to define outcome at long-term follow-up was the PANSS total score.

Significant correlations of patient characteristics with long-term outcome

- Significant correlations with the PANSS total score were found for level of concurrent medication (chlorpromazine equivalent) (0.23) and duration of illness (0.22)

Significant correlations of pretreatment outcome measures with long-term outcome

- At pretreatment the BSI-GSI and PANSS subscales all showed a significant association with the long-term follow-up PANSS total score (0.34 to 0.61). With regard to the pretreatment social functioning subscales, withdrawal and independence competence were significantly associated with the PANSS total score (–0.38 to –0.22).

Significant correlations of post-treatment outcome measures with long-term outcome

- At post-treatment the BSI-GSI and PANSS subscales all showed a significant association with the long-term follow-up PANSS total score (0.26 to 0.62). With regard to the post-treatment social functioning subscales, withdrawal, independence competence and occupation/employment were significantly associated with the PANSS total score (–0.35 to –0.22).
- Completion of treatment was not significantly correlated with the PANSS total score at long-term follow-up. Receipt of treatment with CBT approached significance with the PANSS total score (–0.20).

Analyses of the effects of missing data

- Possible effects of missing data on the correlations between scores on the PANSS total score at long-term follow-up and other measures were examined by comparing different estimates of the correlations. In general, the effects of missing data were small

but may have led to a slight overestimate of the association of CBT with outcome.

Regression analyses

- Stepwise regressions predicting the PANSS total score at long-term follow-up from all predictor variables available at post-treatment gave a multiple r of 0.65 explaining about 38% of the variance. Age, duration of illness, pretreatment recreation subscale of the social functioning scale and post-treatment PANSS general symptoms scores all contributed to the prediction.

Chapter 10: Health economic analyses

Anxiety disorder studies (Trials 1–8)

Method

- The cost variable consisted of the healthcare costs in the 2-year period prior to follow-up plus the cost of the original intervention. Full costs were available for 359 individuals. The benefit variable consisted of the SF-36 PC and MC scores. This was available for 342 individuals with missing data for the remaining 17 cases imputed using the mean summary scores for the SF-36 from the relevant treatment arm and trial.

Cost-effectiveness analysis using intention-to-treat sample of CBT versus non-CBT

- The relative cost-effectiveness of CBT compared with other treatments was assessed with the ICER expressed as the cost of CBT minus the cost of non-CBT divided by the benefit of CBT minus the benefit of non-CBT. This was done for the whole sample as an intention-to-treat analysis. CBT was associated with a slightly higher cost of £5 (£1474 – £1469) and a slightly higher benefit of 1.41 on the physical summary score (40.86 – 39.45) and 2.55 on the mental health summary score (41.02 – 39.21). The ICER was, therefore, £3.55 per one point gain on the physical summary score and £2.76 per one point gain on the mental summary score.
- Bootstrapping with 1000 samples was used to generate estimates of uncertainty around the above figures. It was concluded that in 80–90% of cases CBT does confer benefit with respect to physical and mental health summary scores and in only a very small proportion of cases is CBT unequivocally non-cost-effective and associated with less benefit at greater cost.

Cost-effectiveness analysis using CBT completers versus the remainder

- The ICER was recalculated so as to compare costs and benefits for CBT completers versus an aggregate group of patients who were not allocated to CBT, who dropped out of CBT and who failed to commence CBT. CBT completers were associated with a lower cost of £45 and a slightly higher benefit of 1.03 on the physical summary score and 2.55 on the mental health summary score.
- Bootstrapping with 1000 samples was again used to generate estimates of uncertainty around the above figures with results that were generally more favourable to CBT than the intention-to-treat analysis.

Psychosis studies (Trials 9 and 10)

Method

- The cost variable consisted of the healthcare costs in the 2-year period prior to follow-up. Full costs were available for 94 individuals. The benefit variable consisted of the SF-36 PC and MC scores. This was available for 80 individuals with partial data for six cases that were imputed using regression.

Cost-effectiveness analysis using intention-to-treat sample of CBT versus non-CBT

- The relative cost-effectiveness of CBT compared with other treatments was assessed with the ICER expressed as the cost of CBT minus the cost of non-CBT divided by the benefit of CBT minus the benefit of non-CBT. This was done for the whole sample as an intention-to-treat analysis. CBT was associated with a higher cost of £173 and lower levels of benefit of 0.82 on the physical summary score and 4.29 on the mental health summary score.
- Bootstrapping with 1000 replications was used to generate estimates of uncertainty around the above figures. It was concluded that in 37% of cases CBT does confer benefit with respect to physical summary score and in only a very small proportion of cases (around 2%) with respect to the mental health summary score. In around 48% of cases CBT was associated with lower levels of benefit and higher levels of cost for the mental health summary score. In around 30% of cases this was the result for the physical component summary score.

Chapter 12

Dissemination

Publications

Chambers JA, Power KG, Durham RC. Parental styles and long-term outcome following treatment for anxiety disorders. *Clin Psychol Psychother* 2004;**11**:187–98.

Chambers JA, Power KG, Durham RC. The relationship between trait vulnerability and anxiety and depressive diagnoses at long-term follow-up of generalised anxiety disorder. *J Anxiety Disord* 2004;**18**:587–607.

Durham RC, Chambers JA, Macdonald RR, Power KG, Major K. Does cognitive behavioural therapy influence the long-term outcome of generalised anxiety disorder? An 8–14 year follow-up of two clinical trials. *Psychol Med* 2003;**33**:499–509.

Submitted

Chambers JA, Power KG, Durham RC, Sharp DM. Perceived parenting and co-morbidity at long-term follow-up of treatment for anxiety disorders.

Durham RC, Macdonald RR, Chambers JA, Fisher PL, Power KG, Dow MGT, *et al*. Predictive validity of two prognostic indices in generalised anxiety disorder: complexity of problems and quality of therapeutic alliance.

Conference presentations

2001 Symposium (Chair, RC Durham; Discussant, Prof. Derek Johnston; Contributors, JA Chambers, KA Major, DM Sharp, KG Power, RC Durham). Title: 'Do the effects of CBT endure? Long-term outcome of clinical trials for GAD and panic disorder'. Venue: British Association for Behavioural and Cognitive Psychotherapies (BABCP) Annual Conference, University of Strathclyde, 23 June.

2004 Symposium (Chair, KG Power; Discussant, Prof. Kate Davidson; Contributors: JA Chambers, KA Major, RR Macdonald, RC Durham, DM Sharp). Title: 'A decade on: long-term outcome of eight clinical trials of CBT for anxiety disorders'. Venue: European Association for Behavioural and Cognitive Therapy (EABCT) Annual Conference, University of Manchester, 7 September.

Future plans

There are a number of more detailed analyses of our findings that we intend to submit for publication following the end of the research project.

Chapter 13
Conclusions

Implications for healthcare

Anxiety disorder studies (Trials 1–8)

Overall outcome

1. *Findings on the overall outcome of Trials 1–8 confirm previous research suggesting that, irrespective of medical or psychological treatment, anxiety disorders tend to follow a chronic course with a significant minority doing very poorly indeed. Treatment services for these relatively common conditions need to recognise this fact and include resources for establishing diagnosis during initial screening and monitoring long-term outcome. Continued support and booster sessions may need to be available to help vulnerable individuals to minimise levels of disability and reduce frequency of relapse. This requires further investigation.*
   (a) The majority of participants in this study are clearly not the 'worried well'. Overall health status scores are comparable to the lowest 25% of the general population. One half of the participants have at least one diagnosis at long-term follow-up with significant levels of co-morbidity and health status scores comparable to the lowest 10% of the general population. Over 60% of participants report varying degrees of professional help for anxiety over the follow-up period with 19% receiving almost constant treatment. These findings suggest that anxiety disorders should be managed as a 'chronic disease' much like diabetes. Arguments for managing depression in this way have been advanced in recent years[165] and it may well be that anxiety disorders should also be managed in this way.
   (b) In considering the implications for healthcare, it is important to remember that all participants in the clinical trials met diagnostic criteria for specific anxiety disorders. Our findings are only generalisable to clinical services in which the severity and duration criteria of individual diagnoses are taken seriously and systematically assessed. Diagnostic rigour is important not just in distinguishing between anxiety disorders and adjustment disorders, which are less severe and generally short-lived, but also in assessing co-morbidity with other clinical and personality disorders. In particular, co-morbid anxiety and depression is known to have a significantly worse outcome than anxiety or depression alone[6] and co-morbid anxiety and personality disorder is also associated with a generally worse outcome than anxiety alone.[166,167]
   (c) GAD and panic disorder were found to have generally similar outcomes but the outcome of PTSD appeared to be significantly worse. If this finding is generally confirmed it may well be that PTSD requires specialist clinical services.

Efficacy of CBT in changing the long-term course of anxiety disorders

2. *Treatment with CBT was found to have a better long-term outcome than non-CBT in terms of overall symptom severity but not in regard to diagnostic status. The positive effects of CBT found in the original clinical trials are eroded over longer time periods. Hence, although CBT may be the most efficacious psychological treatment for anxiety disorders,[1] there is clearly room for improvement in the power of CBT to bring about enduring change. The present study does not have any specific implications for how these improvements might occur but changes in the type of treatment, the training of therapists, or the manner in which treatment is delivered, may all be required. This needs further investigation.*
   (a) A positive response to CBT for anxiety disorders over the short term is associated with a greater likelihood of a positive outcome over the long term, as it is with other treatments, but this should not be taken as evidence that a positive response will be sustained over the longer term. A negative response is associated with a poor outcome. In general, treatment protocols for GAD, panic disorder and PTSD may need to place greater emphasis on relapse prevention, booster sessions and the consolidation of coping skills. Psychological therapy for anxiety disorders may need to be recast in terms of episodes of treatment over extended periods of time. This will require a significant shift in the expectations given to patients about the likely outcome of therapy and the need for

extended contact to review progress. This in turn may have implications for the delivery of services so that greater resources are allocated to follow-up and additional therapy for previously treated cases.

(b) No evidence in the present study was found for an association between more intensive therapy and more enduring effects of CBT. The dose–effect relationship examined in the present study was restricted but within the parameters of normal practice and typical costs.[168] It may be that only multiple episodes of therapy, or an intensity of therapy well outside the normal range, will add significantly to the efficacy of CBT in bringing about enduring change. Until further research investigates this possibility, as advocated for example by Aveline,[169] it seems reasonable to conclude that clinicians who go beyond standard treatment protocols of about 10 sessions during a 6-month period are unlikely to bring about greater improvement.

(c) In considering the implications of the present study it should be remembered that the clinical effectiveness of CBT in routine practice may, in fact, be less than that found in RCTs. Treatment services may need to include some of the same methodological rigour found in clinical research (i.e. clinical supervision and assessment of therapist competence in delivering treatment protocols) in order to achieve the same results.[170] This is likely to be especially important in some of the recent developments in therapy technology, such as metacognitive models of CBT,[171] which appear to be more demanding of therapist skill than standard treatments.

Predictors of long-term outcome in the anxiety disorders

3. *Long-term outcome of the anxiety disorders was found to be most strongly predicted by the complexity and severity of presenting problems at the time of referral, completion of treatment, whether CBT or non-CBT, and the receipt of further treatment during the follow-up period. Systematic assessment of these variables should be included in clinical services in order to address the needs of cases where the prognosis is likely to be poor.*

   (a) Significant long-term outcome predictors assessed at a screening interview included demographic variables (employment status, level of social deprivation), measures of the severity of symptomatology and quality of social adjustment. These variables were found to be moderately related to long-term outcome and underline the association between social inequality, severity of symptomatology and poor mental health.[172] Broadly similar results have been found in other studies of outcome prediction.[173] These findings suggest that pretreatment variables of the kind that can be assessed at a screening interview are likely to be of value in identifying patients who are most vulnerable to poor long-term outcome and who may require more intensive therapy and longer term follow-up. Prognostic indices based on the complexity and severity of presenting problems, such as the prognostic index used in Trial 7, merit further investigation as a clinical tool in services for anxiety disorders.

   (b) Treatment completers did better at long-term outcome regardless of treatment modality. This underlines the importance of engagement and treatment compliance.

   (c) The most powerful overall predictor, although rated retrospectively and open to the possibility of some bias, was amount of interim treatment for anxiety over the follow-up period, with more interim treatment being related to poorer outcome. Poor initial outcome leads to further treatment but not to better long-term outcome, although it may be the case that additional treatment prevents further deterioration.

   (d) Finally, the quality of the therapeutic alliance, which is conventionally thought to be a valid indicator of treatment outcome, was found to be moderately related to short-term outcome but unrelated to long-term outcome. Positive engagement in therapy is no guarantee of good outcomes over the long term. Our current treatment technology produces very worthwhile improvements over the short and medium term (6–12 months following a course of therapy) but does not have the power to change the overall course of the disorder. Mental health professionals need to be aware of this fact when designing clinical services.

Health economic analyses for the anxiety disorder studies

4. *Decisions on the cost-effectiveness of providing CBT for anxiety disorders in the NHS should take account of the fact that the provision of CBT makes only a*

minor contribution to the overall costs of healthcare for this patient group. Provided that CBT can be properly targeted on those in most need and has some effect on improving health it is likely to be a worthwhile investment.

(a) Although the cost-effectiveness analysis did not produce a significant result, it is important to note the relatively minor contribution that the original intervention made to the overall costs of treatment for this patient population – a mere 6.4% of the 2-year total costs.

(b) Given appropriate selection and patient choice, the primary decision for the NHS with respect to CBT is not likely to be an economic one given the very minor relative levels of resource commitment.

(c) The health economic analysis makes it clear that people suffering from anxiety disorders consume very significant levels of healthcare resources relative to the general population. This confirms previous research findings that the economic costs of anxiety disorders are high.[174–176]

(d) From a health economic perspective, effort should be concentrated on the development of initiatives to help anxiety disorder patients become less dependent on NHS resources. The significant development, both in number and form, of self-administered CBT programmes is entirely consistent with such aims. Indeed, refining treatment delivery to the point where effective self-care becomes possible has been considered the most sophisticated stage of a therapeutic science.[177]

(e) Possibly unsurprisingly, it is extremely difficult within the framework of cost-effectiveness analysis to assess the impact of a relatively brief intervention when considered from a long-term perspective; this applies across all 10 trials. Assuming that all other differences are adequately captured by randomisation procedures is perhaps somewhat heroic. More detailed analysis including factors we know to influence levels of resource use, such as clinical severity, poverty, proximity to services, changes in technology (such as the introduction of atypical antipsychotics) and age, may in the future provide a more meaningful assessment of CBT interventions in a range of diagnostic groupings.

5. *Chronic anxiety disorder is associated with both poor physical health and poor mental health and the combination of the two results in very poor quality of life and significant disability. Clinical services for chronic anxiety disorders need to recognise this close association and provide more integrated care across medical and psychiatric services.*

(a) The close relationship between poor mental and physical health in those patients suffering from chronic anxiety disorder underlines the complexity of the healthcare needs of this group and potential importance of more integrated care across psychiatric, psychological and medical services.

(b) The nature of the causal interaction between poor physical and mental health was not a part of the present study but evidence is accumulating that points to the adverse effects of chronic anxiety on physical health.[178] This underlines the potential importance of identifying vulnerable individuals at an early stage of the disorder and providing more intensive therapy.

Psychosis studies (Trials 9 and 10)

6. *No evidence was found in the present study that the beneficial effects of CBT found in the short-term in Trials 9 and 10 were maintained over the longer term. The relative gains of CBT are much greater in anxiety disorders than in psychosis.*

(a) In general, patients with psychosis were rated as having higher clinical global severity than that of anxiety disorder patients, were in receipt of virtually continuous treatment since the original trials, incurred very significantly higher healthcare costs and reported less overall improvement. It should be noted, however, that their scores on self-report measures of symptomatology were broadly comparable to those of GAD and panic disorder and were significantly better than the PTSD patients. Patients with psychosis are not necessarily more distressed or disabled than patients with anxiety disorders and some anxiety disorder patients appear significantly more distressed and disabled than patients with psychosis.

(b) Poor long-term outcomes were associated with being younger, having a longer duration of illness, having higher scores on general psychopathology at post-treatment and poorer social adjustment at pretreatment.

(c) The cost-effectiveness analysis for psychosis was much less favourable to CBT than that for anxiety disorders. Their levels of resource use were very significantly higher

than for anxiety disorders largely as a result of much more frequent episodes of inpatient care.

Recommendations for future research

Anxiety disorders studies (Trials 1–8)

Overall outcome

7. *In general, longitudinal research designs over extended periods of time (2–5 years or more) are required to investigate the determinants of good and poor outcome within the anxiety disorders. Worthwhile investigations with sufficient power require a large number of participants (500+) and this is likely to be best achieved with naturalistic cohorts of patients recruited within primary care. Representative samples of patients with both acute and chronic disorder are essential in developing a balanced picture of overall outcome.[173] A better understanding of the psychobiological mechanisms underlying vulnerability and resilience in stress-related psychopathology needs to draw on expertise in the clinical, biological and social sciences. Some specific research questions arising from the present study are as follows.*
   (a) How important are specific diagnoses as a determinant of long-term outcome? There was little evidence of differences in outcome between GAD and panic disorder patients in the present study, which is consistent with the findings of other researchers.[34] PTSD patients, however, had a noticeable worse outcome and this merits further investigation.
   (b) Average levels of co-morbidity were high in the present study and, in a separate investigation not described in this report, degree of co-morbidity was found to be closely related to the severity of positive and negative affect.[179] The mechanisms whereby poor treatment responders become increasingly disabled by multiple disorders need to be elucidated.
   (c) A significant proportion of participants in the present study were found to have a very poor outcome despite consuming large amounts of a diverse range of healthcare resources during the follow-up period. Why is treatment so unsuccessful for this group and how can existing treatments be tailored to address chronic populations more effectively?
   (d) Improvements in the management and prevention of chronic anxiety disorder are likely to come ultimately from a better understanding of the factors underlying chronicity. Unlike chronic depression, where considerable progress has been made in recent years in defining the condition, chronic anxiety is less well defined. Progress needs to be made in specifying operational criteria and elucidating the essential characteristics of the condition. How does it differ from acute anxiety? How does chronic anxiety differ from chronic depression?
   (e) The broad determinants of chronic anxiety are unlikely to be fundamentally different from those of chronic depression where developmental factors (childhood adversity, early trauma), neuroticism, psychosocial stressors, biological factors (especially immunological response to stress) and cognitive factors are all of potential importance.[180] Prospective, longitudinal research needs to investigate these factors in anxiety disorders every bit as much as in depression.

Efficacy of CBT in changing the long-term course of anxiety disorders

8. *RCTs have clearly established the efficacy of CBT with anxiety studies over short-term follow-up periods and the task now is to examine issues of clinical utility or effectiveness in routine clinical practice.[181] The erosion of the effects of CBT over time suggests two important avenues of research: at what point does relapse occur and how can psychological therapy be improved in order to produce more enduring change? These questions are likely to be best answered by grafting specific experimental investigations on to naturalistic follow-up studies of the kind indicated above.*
   (a) Six-monthly reviews of the progress of treated patients are needed to establish the time periods over which treatment effects are sustained.
   (b) Experimental investigations of different approaches to relapse prevention and the maintenance of change are needed to develop more robust forms of CBT. Without minimising the value of conventional RCTs, clinical effectiveness may be further enhanced by placing greater weight than hitherto on comprehensive screening interviews and regular reviews of progress into routine clinical practice and the randomisation of therapists to different therapeutic strategies for addressing relapse and maintenance of change. Following Baer and colleagues'[182] dictum almost 40 years ago, 'generalisation

should be programmed, rather than expected or lamented'.

9. *The importance of patient characteristics as a significant determinant of long-term outcome clearly merits further investigation. The development and refinement of a prognostic index that can be used in everyday practice will enable a much more searching examination of the clinical effectiveness of CBT as delivered to patients with differing degrees of complexity and severity by therapists of differing degrees of clinical expertise.*
   (a) How does therapist competence interact with treatment outcome in patients of differing degrees of complexity and severity? Is it the case that therapists with high competency ratings on established measures of clinical skill produce better long-term outcomes?
   (b) What is the relative efficacy of CBT and medication in patients with a poor prognosis?
   (c) How can prognostic indices best be used to identify patients who, on the one hand, are most likely to respond to brief treatment and self-help and who, on the other hand, are most likely to require more intensive therapy with extended contact?

Psychosis studies (Trials 9 and10)

10. *Clinical research on CBT for psychosis has produced some promising findings but no reliable and substantive evidence base for clinically significant benefits from CBT beyond the active treatment phase. The present results suggest that future clinical trials should evaluate the efficacy of courses of CBT that are more intensive and that build into the protocol some degree of continuing care to ensure maintenance of treatment gains. As with anxiety disorders, the influence of therapist expertise in delivering these treatment protocols and the quality of the therapeutic alliance have yet to be determined.*

Health economic analyses

11. *The most important next step in the economic evaluation presented in this report is to conduct more detailed analyses of the current data set in order to address the following issues.*
   (a) What forms of missing data analysis are of most value in studies of this kind?
   (b) What proportion of the resources used by the different diagnostic groups can be attributed to physical health needs and what proportion can be attributed to mental health needs?
   (c) What is the source of drivers for the cost increases identified over the two time periods where resource use was collected?
   (d) What influence do variables other than original treatment group have on the relative cost-effectiveness of interventions. In particular, what is the influence of social deprivation, education level, employment status and age?

Acknowledgements

The preparation of this report was made possible by a grant from the HTA Programme, whose support is gratefully acknowledged. The views and opinions expressed in this report do not necessarily reflect those of the HTA Programme. The senior author is grateful to Professor Ian Reid for drawing his attention to the initial call for research proposals to investigate the long-term effects of CBT and to both Ian Reid and Professor Keith Matthews for much encouragement and support over the period of this project. The authors are grateful to Jen Petrie for her invaluable assistance in coordinating all phases of this project, to Susan Schooling, Alison Tennant and Alana Gibb for their steadfast efforts in conducting the interviews and collecting case note data for the two psychosis studies, and to Cathryn Hau for her initial contribution as a member of the long-term follow-up research team. For their help with the Health Economics, the authors wish to thank Matthew Sutton, Senior Research Fellow, University of Glasgow, for econometric advice and assistance with the analysis; Marjorie Marshall, Health Economist, NHS Ayrshire and Arran, for detailed micro costings throughout the studies; Susan Callaghan and Debbie Lawson, Administrative Assistants, NHS Ayrshire and Arran, for designing data recording systems and for significant amounts of data recording; and Fiona Brennan and Nadine Kane, Administrative Assistants, NHS Ayrshire and Arran, for data recording. Finally, our appreciation must also go to the many GP practices and their patients who gave their time and help during the course of this study.

Contribution of the authors

RC Durham (Senior Lecturer in Clinical Psychology), KG Power (Head of Service, Clinical Psychology) and DM Sharp (Senior Lecturer in Behavioural Oncology) had primary responsibility for the conception and design, with contributions from MGT Dow (Senior Clinical Tutor), AI Gumley (Senior Lecturer in Clinical Psychology) and KA Major (Director of Strategic Planning and Performance). JA Chambers (Research Psychologist) and RR Macdonald (Lecturer in Psychology) and all the authors contributed to the analysis, interpretation and writing of the report.

References

1. Roth A, Fonagy P. *What works for whom? A critical review of psychotherapy research*. London: Guilford Press; 1996.
2. Department of Health. Doh.gov.uk [homepage on the Internet]. Treatment choice in psychological therapies and counselling. *Evidence based clinical practice guideline*. 2001. URL: http://www.doh.gov.uk/mentalhealth/treatmentguideline. Accessed 2 December 2003.
3. Cormac I, Jones C, Campbell C. Cognitive behaviour therapy for schizophrenia (Cochrane Review). *Cochrane Library*. Issue 4. Oxford: Update Software; 2002.
4. Cowley D, Flick S, Roy-Byrne P. Long-term course and outcome in panic disorder: a naturalistic follow-up study. *Anxiety* 1996;**2**:13–21.
5. Rickels K, Schweizer E. The clinical course and long-term management of generalised anxiety disorder. *J Clin Psychopharmacol* 1990; **10** Suppl 3:101–10.
6. Emmanuel J, Simmonds S, Tyrer P. Systematic review of the outcome of anxiety and depressive disorders. *Br J Psychiatry* 1998;**173** Suppl 34:35–41.
7. Parry G. *Review of strategic policy on NHS psychotherapy services in England*. London: NHSE; 1996.
8. Van Balkom AJLM, Bakker A, Spinhoven P, Blaauw BMJW, Smeenk S, Ruesink B. A meta-analysis of the treatment of panic disorder with or without agoraphobia: a comparison of psychopharmacological, cognitive-behavioral, and combination treatments. *J Nerv Ment Dis* 1997; **185**:510–16.
9. Arntz A, van der Hout M. Psychological treatments of panic disorder without agoraphobia: cognitive therapy versus applied relaxation. *Behav Res Ther* 1996;**34**:113–21.
10. Barlow DH, Gorman JM, Shear KM, Woods SW. Cognitive-behavioral therapy, imipramine, or their combination for panic disorder: a randomized controlled trial. *JAMA* 2000;**283**:2529–36.
11. Bouchard S, Gauthier J, Laberge B, French D, Pelletier MH, Godbout C. Exposure versus cognitive restructuring in the treatment of panic disorder with agoraphobia. *Behav Res Ther* 1996; **34**:213–24.
12. Craske MG, DeCola JP, Sachs AD, Pontillo DC. Panic control treatment for agoraphobia. *J Anxiety Disord* 2003;**17**:321–33.
13. Hecker JE, Losee MC, Fritzler BK, Fink CM. Self-directed versus therapist-directed cognitive behavioral treatment for panic disorder. *J Anxiety Disord* 1996;**10**:253–65.
14. Kenardy JA, Dow MGT, Johnston DW, Newman MG, Thomson A, Taylor CB. A comparison of delivery methods of cognitive behavioural therapy for panic disorder: an international multi-centre trial. *J Consult Clin Psychol* 2003;**71**:1068–75.
15. Schmidt NB, McCreary BT, Trakowski JJ, Santiago HT, Woolaway BK, Ialongo N. Effects of cognitive behavioral treatment on physical health in patients with panic disorder. *Behav Ther* 2003; **34**:49–63.
16. Hegel MT, Ravaris CL, Ahles TA. Combined cognitive-behavioral and time-limited alprazolam treatment of panic disorder. *Behav Ther* 1994; **25**:183–95.
17. Biondi M, Picardi A. Increased probability of remaining in remission from panic disorder with agoraphobia after drug treatment in patients who received concurrent cognitive-behavioural therapy: a follow-up study. *Psychother Psychosom* 2003; **72**:34–42.
18. Clark DM, Salkovskis PM, Hackmann A, Wells A, Ludgate J, Gelder M. Brief cognitive therapy for panic disorder: A randomized controlled trial. *J Consult Clin Psychol* 1999;**67**:583–9.
19. Clark DM, Salkovskis PM, Hackmann A, Middleton H, Anastasiades P, Gelder M. A comparison of cognitive therapy, applied relaxation and imipramine in the treatment of panic disorder. *Br J Psychiatry* 1994;**164**:759–69.
20. Chambless DL, Gillis MM. Cognitive therapy of anxiety disorders. *J Consult Clin Psychol* 1993; **61**:248–60.
21. Borkovec TD, Whisman MA. Psychosocial treatment for generalized anxiety disorder. In Mavissakalian M, Prien RF, editors. *Long-term treatment of anxiety disorders*. Washington, DC: American Psychological Association; 1996. pp. 171–99.
22. Gould R, Otto M, Pollack M, Yap L. Cognitive behavioural and pharmacological treatment of generalized anxiety disorder: a preliminary meta-analysis. *Behav Ther* 1997;**28**:285–305.
23. Borkovec TD, Costello E. Efficacy of applied relaxation and cognitive-behavioural therapy in the treatment of generalised anxiety disorder. *J Consult Clin Psychol* 1993;**61**:611–19.

24. Power KG, Simpson RJ, Swanson V, Wallace L, Feistner AJG, Sharp DM. A controlled comparison of cognitive behavioural therapy, diazepam and placebo alone and in combination in the treatment of generalised anxiety disorder. *J Anxiety Disord* 1990;**4**:267–92.

25. Fisher PL, Durham RC. Recovery rates in generalised anxiety disorder following psychological therapy: An analysis of clinically significant change in STAI-T across outcome studies since 1990. *Psychol Med* 1999;**29**:1425–34.

26. Durham RC, Fisher PL, Treliving L, Hau CM, Richard K, Stewart JB. One year follow-up of cognitive therapy, analytic psychotherapy and anxiety management training for generalised anxiety disorder: symptom change, medication usage and attitudes to treatment. *Behav Cogn Psychother* 1999;**27**:19–35.

27. Ladouceur R, Dugas MJ, Freeston MH, Leger E, Gagnon F, Thibodeau N. Efficacy of a cognitive-behavioral treatment for generalized anxiety disorder: evaluation in a controlled clinical trial. *J Consult Clin Psychol* 2000;**68**:957–64.

28. Borkovec TD, Newman M, Pincus A, Lytle R. A component analysis of cognitive behavioural therapy for generalized anxiety disorder and the role of interpersonal problems. *J Consult Clin Psychol* 2002;**70**:288–98.

29. Öst LG, Breitholtz E. Applied relaxation vs. cognitive therapy in the treatment of generalized anxiety disorder. *Behav Res Ther* 2000;**38**:777–90.

30. Butler G, Fennell M, Robson P, Gelder M. Comparison of behaviour therapy and cognitive-behaviour therapy in the treatment of generalised anxiety disorder. *J Consult Clin Psychol* 1991; **59**:167–75.

31. Barlow DH, Rapee RM, Brown TA. Behavioral treatment of generalized anxiety disorder. *Behav Ther* 1992;**23**:551–70.

32. Dugas MJ, Ladouceur R, Leger E, Freeston MH, Langlois F, Provencher MD, *et al*. Group cognitive-behavioral therapy for generalized anxiety disorder: treatment outcome and long-term follow-up. *J Consult Clin Psychol* 2003;**71**:821–5.

33. White J. "Stress control" large group therapy for generalized anxiety disorder: two year follow-up. *Behav Cogn Psychother* 1998;**26**:237–45.

34. Seivewright H, Tyrer P, Johnson T. Prediction of outcome in neurotic disorder: a 5 year prospective study. *Psychol Med* 1998;**28**:1149–57.

35. Blanchard EB, Hickling EJ, Devineni T, Veazey CH, Galovski TE, Mundy E, *et al*. A controlled evaluation of cognitive behavioral therapy for posttraumatic stress in motor vehicle accident survivors. *Behav Res Ther* 2003; **41**:79–96.

36. Ehlers A, Clark DM, Hackmann A, McManus F, Fennell M, Herbert C, *et al*. A randomized controlled trial of cognitive therapy, a self-help booklet, and repeated assessments as early interventions for posttraumatic stress disorder. *Arch Gen Psychiatry* 2003;**60**:1024–32.

37. Fecteau G, Nicki R. Cognitive behavioural treatment of post traumatic stress disorder after motor vehicle accident. *Behav Cogn Psychother* 1999;**27**:201–14.

38. Marks I, Lovell K, Noshirvani H, Livanou M, Thrasher S. Treatment of post-traumatic stress disorder by exposure and/or cognitive restructuring. A controlled study. *Arch Gen Psychiatry* 1998;**55**:317–25.

39. Paunovic N, Öst LG. Cognitive-behavior therapy vs exposure therapy in the treatment of PTSD in refugees. *Behav Res Ther* 2001;**39**:1183–97.

40. Resick PM, Nishith P, Weaver TL, Astin MC, Feuer CA. A comparison of cognitive-processing therapy with prolonged exposure and a waiting condition for the treatment of chronic posttraumatic stress disorder in female rape victims. *J Consult Clin Psychol* 2002;**70**:867–79.

41. Tarrier N, Pilgrim H, Sommerfield C, Faragher B, Reynolds M, Graham E, *et al*. A randomized trial of cognitive therapy and imaginal exposure in the treatment of chronic posttraumatic stress disorder. *J Consult Clin Psychol* 1999;**67**:13–18.

42. Deblinger E, Steer RA, Lippmann J. Two-year follow-up study of cognitive behavioural therapy for sexually abused children suffering post-traumatic stress symptoms. *Child Abuse Neglect* 1999;**23**:1371–8.

43. Basoglu M, Livanou M, Salcioglu E, Kalender D. A brief behavioural treatment of chronic post-traumatic stress disorder in earthquake survivors: results from an open clinical trial. *Psychol Med* 2003;**33**:647–54.

44. Wilson SA, Becker LA, Tinker RH. Fifteen-month follow-up of eye movement desensitisation and reprocessing (EMDR) treatment for posttraumatic stress disorder and psychological trauma. *J Consult Clin Psychol* 1997;**65**:1047–56.

45. Bryant RA, Moulds ML, Nixon RVD. Cognitive behaviour therapy of acute stress disorder: a four-year follow-up. *Behav Res Ther* 2003;**41**:489–94.

46. Chemtob CM, Novaco RW, Hamada RS, Gross DM. Cognitive-behavioural treatment for severe anger in posttraumatic stress disorder. *J Consult Clin Psychol* 1997;**65**:184–9.

47. Humphreys L, Westerink J, Giarratano L, Brooks R. An intensive treatment program for chronic posttraumatic stress disorder: 2-year outcome data. *Aust N Z J Psychiatry* 1999; **33**:848–54.

48. Richards DA, Lovell K, Marks IM. Post-traumatic stress disorder: evaluation of a behavioural treatment program. *J Trauma Stress* 1994;**7**:669–80.

49. Tarrier N, Sommerfield C, Pilgrim H, Humphreys L. Cognitive therapy or imaginal exposure in the treatment of post-traumatic stress disorder. *Br J Psychiatry* 1999;**175**:571–5.

50. Power KG, McGoldrick T, Brown K, Buchanan R, Sharp DM, Swanson V, *et al*. A controlled comparison of eye movement desensitisation and reprocessing versus exposure plus cognitive restructuring versus waiting list in the treatment of post-traumatic stress disorder. *Clin Psychol Psychother* 2002;**9**:299–318.

51. Bisson JI, Shepherd JP, Joy D, Probert R, Newcombe RG. Early cognitive-behavioural therapy for post-traumatic stress symptoms after physical injury: randomised controlled trial. *Br J Psychiatry* 2004;**184**:63–9.

52. Foa EB, Dancu CV, Hembree EA, Jaycox LH, Meadows EA, Street GP. A comparison of exposure therapy, stress inoculation training and their combination for reducing post-traumatic stress disorder in female assault victims. *J Consult Clin Psychol* 1999;**67**:194–200.

53. Schnurr P, Friedman MJ, Foy D, Shea MT, Hsieh FY, Lavori PW, *et al*. Randomized trial of trauma-focused group therapy for post-traumatic stress disorder: Results from a Department of Veterans Affairs Cooperative Study. *Arch Gen Psychiatry* 2003;**60**:481–9.

54. Horowitz M, Wilner N, Alvarez W. Impact of Event Scale: A measure of subjective stress. *Psychosom Med* 1979;**41**:209–18.

55. Blake D, Weathers F, Nagy L, Kaloupek D, Klauminzer G, Charney D, *et al*. *Clinician Administered PTSD Scale (CAPS)*. Boston, MA: National Center for Post-traumatic Stress Disorder; 1990.

56. Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. *Clin Psychol Rev* 1988;**8**:77–100.

57. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand* 1983; **67**:361–70.

58. Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA. *Manual for the State–Trait Anxiety Inventory (Form Y Self-evaluation Questionnaire)*. Palo Alto, CA: Consulting Psychologists Press; 1983.

59. Tarrier N, Beckett R, Harwood S, Baker A, Yusupoff L, Ugarteburu I. A trial of two cognitive-behavioural methods of treating drug-resistant residual psychotic symptoms in schizophrenic patients: 1. Outcome. *Br J Psychiatry* 1993; **162**:524–32.

60. Drury V, Birchwood M, Cochrane R, MacMillan F. Cognitive therapy and recovery from acute psychosis: a controlled trial. 1. Impact on psychotic symptoms. *Br J Psychiatry* 1996; **169**:593–601.

61. Drury V, Birchwood M, Cochrane R, MacMillan F. Cognitive therapy and recovery from acute psychosis: a controlled trial: II. Impact on recovery time. *Br J Psychiatry* 1996;**169**:602–7.

62. Kuipers E, Garety P, Fowler D, Dunn G, Bebbington P, Freeman D, *et al*. London–East Anglia randomised controlled trial of cognitive-behavioural therapy for psychosis 1: effects of the treatment phase. *Br J Psychiatry* 1997;**171**:319–27.

63. Tarrier N, Yusupoff L, Kinney C, McCarthy E, Gledhill A, Haddock G, *et al*. Randomised controlled trial of intensive cognitive behaviour therapy for patients with chronic schizophrenia. *BMJ* 1998;**317**:303–7.

64. Pinto A, La Pia S, Mennuella R, Giorgio D, DeSimone L. Cognitive-behavioral therapy and clozapine for clients with treatment-refractory schizophrenia. *Psychiatr Serv* 1999;**50**:901–4.

65. Sensky T, Turkington D, Kingdon D, Scott JL, Scott J, Siddle R, *et al*. A randomized controlled trial of cognitive-behavioural therapy for persistent symptoms in schizophrenia resistant to medication. *Arch Gen Psychiatry* 2000;**57**:165–72.

66. Durham RC, Guthrie M, Morton RV, Reid DA, Treliving L, Fowler D, *et al*. Tayside–Fife clinical trial of cognitive-behaviour therapy for medication-resistant psychotic symptoms: results to three-month follow-up. *Br J Psychiatry* 2003; **182**:303–11.

67. Gumley AI, O'Grady M, McNay L, Reilly J, Power KG, Norrie J. Early intervention for relapse in schizophrenia: results of a 12-month randomised controlled trial of cognitive behavioural therapy. *Psychol Med* 2003;**33**:419–31.

68. Barrowclough C, Haddock G, Tarrier N, Lewis SW, Moring J, O'Brien R, *et al*. Randomized controlled trial of motivational interviewing, cognitive behavior therapy, and family intervention for patients with comorbid schizophrenia and substance use disorders. *Am J Psychiatry* 2001; **158**:1706–13.

69. Haddock G, Barrowclough C, Tarrier N, Moring J, O'Brien R, Schofield N, *et al*. Cognitive-behavioural therapy and motivational intervention for schizophrenia and substance misuse. 18-month outcomes of a randomised controlled trial. *Br J Psychiatry* 2003;**183**:418–26.

70. Haddock G, Tarrier N, Morrison AP, Hopkins R, Drake R, Lewis S. A pilot study evaluating the effectiveness of individual inpatient cognitive-behavioural therapy in early psychosis. *Soc Psychiatry Psychiatr Epidemiol* 1999;**34**:254–8.

71. Lewis S, Tarrier N, Haddock G, Bentall R, Kinderman P, Kingdon D, *et al*. Randomised controlled trial of cognitive-behavioural therapy in early schizophrenia: acute-phase outcomes. *Br J Psychiatry* 2002;**181**(S43):S91–7.

72. Startup M, Jackson MC, Bendix S. North Wales randomized controlled trial of cognitive behaviour therapy for acute schizophrenia spectrum disorders: outcomes at 6 and 12 months. *Psychol Med* 2004;**34**:413–22.

73. Garety P, Fowler D, Kuipers E, Freeman D, Dunn G, Bebbington P, *et al*. London–East Anglia randomised controlled trial of cognitive behavioural therapy for psychosis: II. Predictors of outcome. *Br J Psychiatry* 1997;**171**:420–6.

74. Tarrier N, Wittkowski A, Kinney C, McCarthy E, Morris J, Humphreys L. Durability of the effects of cognitive-behavioral therapy in the treatment of chronic schizophrenia: 12-month follow-up. *J Consult Clin Psychol* 1999;**174**:500–4.

75. Bechdolf A, Knost B, Kuntermann C, Schiller S, Klosterkotter J, Hambrecht M, *et al*. A randomized comparison of group cognitive-behavioural therapy and group psychoeducation in patients with schizophrenia. *Acta Psychiatr Scand* 2004; **110**:21–8.

76. Buchkremer G, Klingberg S, Holle R, Schulze Monking H, Hornung W. Psychoeducational psychotherapy for schizophrenic patients and their key relatives or care-givers: results of a two year follow-up. *Acta Psychiatr Scand* 1997;**96**:483–91.

77. Hogarty G, Cornblith S, Greenwald D, DiBarry A, Cooley S, Ulrich R, *et al*. Three-year trials of personal therapy among schizophrenic patients living with or independent of family. I: Description of study and effects on relapse rates. *Am J Psychiatry* 1997;**154**:1504–13.

78. Kemp R, Hayward P, Applewhaite G, Everitt B, David A. Compliance therapy in psychotic patients: a randomised controlled trial. *BMJ* 1996; **312**:345–9.

79. Overall JE, Gorham DR. The Brief Psychiatric Rating Scale. *Psychol Rep* 1962;**10**:799–812.

80. Montgomery SA, Taylor P, Montgomery D. Development of a schizophrenia scale sensitive to change. *Neuropharmacology* 1978;**17**:1053–71.

81. Kay S, Fiszbein A, Opler L. The Positive and Negative Syndrome Scale (PANSS) *Schizophr Bull* 1987;**13**:261–73.

82. Krawiecka M, Goldberg D, Vaughn M. Standardised psychiatric assessment scale for chronic psychiatric patients. *Acta Psychiatr Scand* 1977;**36**:25–31.

83. David AS. Insight and psychosis. *Br J Psychiatry* 1990;**156**:798–808.

84. Kuipers E, Fowler D, Garety P, Chisholm D, Freeman D, Dunn G, *et al*. London–East Anglia randomised controlled trial of cognitive-behavioural therapy for psychosis: III. Follow-up and economic evaluation at 18 months. *Br J Psychiatry* 1998;**173**:61–8.

85. Kemp R, Kirov G, Everitt B, Hayward P, David A. Randomised controlled trial of compliance therapy. *Br J Psychiatry* 1998;**172**:413–19.

86. Tarrier N, Lewis S, Haddock G, Bentall R, Drake R, Kinderman P, *et al*. Cognitive-behavioural therapy in first-episode and early schizophrenia. *Br J Psychiatry* 2004;**184**:231–9.

87. Tarrier N, Kinney C, McCarthy E, Humphreys L, Wittkowski A, Morris J. Two-year follow-up of cognitive-behavioral therapy and supportive counseling in the treatment of persistent symptoms in chronic schizophrenia. *J Consult Clin Psychol* 2000;**68**:917–22.

88. Klingberg S, Buchkremer G, Holle R, Monking H, Hornung W. Differential therapy effects of psychoeducational psychotherapy for schizophrenic patients – results of 2-year follow-up. *Eur Arch Psychiatry Clin Neurosci* 1999;**249**:66–72.

89. Hogarty G, Greenwald D, Ulrich R, Cornblith S, DiBarry A, Cooley S, *et al*. Three-year trials of personal therapy among schizophrenic patients living with or independent of family. II: Effects on adjustment of patients. *Am J Psychiatry* 1997; **154**:1514–24.

90. Hornung WP, Feldman R, Klingberg S, Buchkremer G, Reker T. Long term effects of psychoeducational psychotherapeutic intervention for schizophrenic outpatients and their key persons: results of a five-year follow-up. *Eur Arch Psychiatry Clin Neurosci* 1999;**249**:162–7.

91. Drury V, Birchwood M, Cochrane R. Cognitive therapy and recovery from acute psychosis: a controlled trial: 3. Five-year follow-up. *Br J Psychiatry* 2000;**177**:8–14.

92. Jones C, Cormac I, Silveira da Mota Neto JI, Campbell C. Cognitive behaviour therapy for schizophrenia. *The Cochrane Database Syst Rev* 2004; Issue 4.

93. Birchwood M. Pathways to emotional dysfunction in first-episode psychosis. *Br J Psychiatry* 2003; **182**:373–5.

94. Trower P, Birchwood M, Meaden A, Byrne S, Nelson A, Ross K. Cognitive therapy for command hallucinations: randomised controlled trial. *Br J Psychiatry* 2004;**184**:312–20.

95. Dillman DA. The design and administration of mail surveys. *Ann Rev Sociol* 1991;**17**:225–49.

96. Edwards P, Roberts I, Clarke M, DiGuiseppi C, Pratap S, Wentz R, *et al*. Increasing response rates to postal questionnaires. *BMJ* 2002;**324**:1183–91.

97. Tyrer P, Seivewright H, Ferguson B, Johnson T. "Cold calling" in psychiatric follow up studies: is it justified? *Psychiatr Ethics* 2003;**29**:238–42.

98. Di Nardo PA, Brown TA, Barlow DH. *Anxiety Disorders Interview Schedule for DSM-IV*. Albany, NY: Graywind Publications; 1994.

99. Jenkinson C, Stewart-Brown S, Petersen S, Paice C. Assessment of the SF-36 Version II in the UK. *J Epidemiol Commun Health* 1999;**53**:46–50.

100. Knapp M, Simon J. Cost implications of schizophrenia in the community. *Psychiatry* 2002; **1**:69–72.

101. Netton A, Dennett J. *Unit costs of health and social care*. Canterbury: University of Kent; 1997.

102. Department of Health. *Hospital community health services pay and prices index*. Leeds: Department of Health; 2002.

103. British Medical Association, Royal Pharmaceutical Society of Great Britain. *British National Formulary*, No. 36. London: BMA; 1998.

104. Information and Statistics Division. *Scottish Health Service costs year ended 31st March 1999*. Edinburgh: Scottish Executive Health Department; 2000.

105. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (3rd edition) (DSM-III)*. Washington DC: APA; 1980.

106. Beck AT, Emery G. *Anxiety disorders and phobias: a cognitive perspective*. New York: Basic Books; 1985.

107. Durham RC, Murphy TJ, Allan T, Richard K, Treliving L, Fenton GW. A comparison of cognitive therapy, analytic psychotherapy and anxiety management training in the treatment of generalised anxiety disorder. *Br J Psychiatry* 1994; **165**:315–23.

108. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (3rd rev. ed.) (DSM-III-R)*. Washington DC: APA; 1987.

109. Sharp DM, Power KG, Simpson RJ, Swanson V, Moodie E, Anstee JA, *et al*. Fluvoxamine, placebo and cognitive behaviour therapy used alone and in combination in the treatment of panic disorder and agoraphobia. *J Anxiety Disord* 1996;**10**:219–42.

110. Barlow DH. *Anxiety and its disorders: the nature and treatment of anxiety and panic*. New York: Guilford Press; 1988.

111. Power KG, Sharp DM, Swanson V, Simpson, RJ. Therapist contact in cognitive behaviour therapy for panic disorder and agoraphobia in primary care. *Clin Psychol Psychother* 2000;**7**:37–46.

112. Barlow DH, Cerny JA. *Psychological treatment of panic*. New York: Guilford Press; 1988.

113. Sharp DM, Power KG, Swanson V. A comparison of the efficacy and acceptability of group versus individual cognitive behavioural therapy in the treatment of panic disorder and agoraphobia in primary care. *Clin Psychol Psychother* 2004;**11**:73–82.

114. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (4th edition) (DSM-IV)*. Washington, DC: APA; 1994.

115. Shapiro F. *Eye-movement desensitization and reprocessing: basic principles, protocols and procedures*. New York: Guilford Press; 1995.

116. Durham RC, Fisher PL, Dow MGT, Sharp DM, Power KG, Swan JS, *et al*. Cognitive behaviour therapy for good and poor prognosis generalised anxiety disorder: a clinical effectiveness study. *Clin Psychol Psychother* 2004;**11**:145–57.

117. Andrews G, Crino R, Hunt C, Lampe L, Page A. *The treatment of anxiety disorders: clinician's guide and patient manuals*. Cambridge: Cambridge University Press; 1994.

118. Barlow DH, Cerny JA, Klosko JS. Behavioral treatment of panic disorder. *Behav Ther* 1989; **20**:261–82.

119. Clark DM. A cognitive approach to panic. *Behav Res Ther* 1986;**24**,461–70.

120. Tarrier N. Management and modification of residual psychotic symptoms. In Birchwood M, Tarrier N, editors. *Innovations in the psychological management of schizophrenia*. Chichester: John Wiley; 1992. pp. 147–69.

121. Kingdon DG, Turkington D. *Cognitive-behavioural therapy of schizophrenia*. Hove: Lawrence Erlbaum Associates; 1994.

122. Gumley AI, Power KG. Is targeting cognitive therapy during early relapse in psychosis feasible? *Behav Cogn Psychother* 2000;**28**:161–74.

123. Carstairs V, Morris R. *Deprivation and health in Scotland*. Aberdeen: Aberdeen University Press; 1991.

124. General Register Office for Scotland. Gro-scotland.gov.uk [homepage on the Internet]. *Scotland's Census 2001 – Reference Volume*. 2003. URL: http://www.gro-scotland.gov.uk/grosweb/grosweb.nsf/pages/scotcenb. Accessed 12 September 2003.

125. Cohen J. A power primer. *Psychol Bull* 1992; **112**:155–9.

126. Di Nardo PA, Moras K, Barlow DH, Rapee RM, Brown TA. Reliability of DSM-III-R anxiety disorder categories: using the Anxiety Disorders Interview Schedule-Revised (ADIS-R). *Arch Gen Psychiatry* 1993;**50**:251–6.

127. Di Nardo PA, Barlow DH. Anxiety Disorders Interview Schedule-Revised (ADIS-R). Albany, NY: Graywind Publications; 1988.

128. Guy W. Patient assessment in clinical trials. *Prog. Neuropsychopharmocol Biol Psychiatry* 1982; **6** Suppl 6:601–6.

129. Brown TA, Barlow DH. Long-term outcome in cognitive-behavioral treatment of panic disorder: clinical predictors and alternative strategies for assessment. *J Consult Clin Psychol* 1995;**63**:754–65.

130. Hamilton M. The assessment of anxiety states by rating. *Br J Med Psychol* 1959;**32**:50–5.

131. Power KG, Jerrom DWA, Simpson RJ, Mitchell M. Controlled study of withdrawal symptoms and rebound anxiety after six weeks course of diazepam for generalised anxiety. *BMJ* 1985; **290**:1246–8.

132. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979;**134**:382–9.

133. Weissman MM, Prusoff BA, Thompson WD, Harding PS, Myers JK. (1978). Social adjustment by self-report in a community sample and in psychiatric outpatients. *J Nerv Ment Disord* 1978; **166**:317–26.

134. Birchwood M, Smith J, Cochrane R, Wetton S, Copestake S. The social functioning scale: the development and validation of a scale of social adjustment for use in family intervention programmes with schizophrenic patients. *Br J Psychiatry* 1990;**157**:853–9.

135. Derogatis LR. *Brief Symptom Inventory (BSI). Administration, scoring and procedures manual*. 3rd ed. Minneapolis, MN: National Computer Systems; 1993.

136. Derogatis LR, Lipman RS, Covi L. SCL-90: an outpatient psychiatric rating scale – preliminary report. *Psychopharmacol Bull* 1973;**9**:13–28.

137. Marks I, Mathews AM. Brief standard self-rating for phobic patients. *Behav Res Ther* 1979;**17**:263–7.

138. Kellner R, Sheffield BF. A self-rating scale of distress. *Psychol Med* 1973;**3**:88–100.

139. Birchwood M, Mason R, Macmillan F. Depression, demoralisation and control over psychotic illness. *Psychol Med* 1993;**23**:387–95.

140. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol* 1988;**54**:1063–70.

141. Rosenberg M. *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press; 1965.

142. O'Malley PM, Bachman JG. Self-esteem and education: sex cohort comparisons among high school seniors. *J Pers Soc Psychol* 1979;**37**:1153–9.

143. Sheehan DV. *The anxiety disease*. New York: Bantam Books; 1986.

144. Ware JE, Snow KK, Kosinski M, Gandek B. *SF-36 health survey manual and interpretation guide*. Boston, MA: Health Institute, New England Medical Center; 1993.

145. Goldberg DP. *Manual of the general health questionnaire*. Windsor: Nelson Publishing; 1978.

146. Chambless DL, Caputo GC, Bright P, Gallagher R. Assessment of fear of fear in agoraphobics: the body sensations questionnaire and the agoraphobic cognitions questionnaire. *J Consult Clin Psychol* 1984;**52**:1090–7.

147. Davidson J, Smith R, Kudler H. Validity and reliability of the DSM-III criteria for posttraumatic stress disorder. Experience with a structured interview. *J Nerv Ment Dis* 1989;**177**:336–41.

148. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 1991;**59**:12–19.

149. Durham R, Chambers J, Macdonald R, Power K, Major K. Does cognitive behavioural therapy influence the course of generalised anxiety disorder? 10–14 year follow-up of two clinical trials. *Psychol Med* 2003;**33**:499–509.

150. Arbuckle JL. *Amos 5.0 [computer program]*. Chicago IL: Smallwaters; 2003.

151. Review Manager (RevMan) [computer program]. Version 4.2 for Windows. Oxford: Cochrane Collaboration; 2002.

152. Marshall M, Lockwood A. *Assertive community treatment for people with severe mental disorders (Cochrane Review)*. Cochrane Library Issue 1. Chichester: John Wiley; 2004.

153. Snaith P. Anxiety control training. *Adv Psychiatr Treat* 1994;**1**:7–61.

154. Lovell K, Richards D. Multiple access points and levels of entry (MAPLE): ensuring choice, accessibility and equity for CBT services. *Behav Cogn Psychother* 2000;**28**:379–91.

155. Shapiro DA, Barkham M, Rees A, Hardy GE, Reynolds S, Startup M. Effects of treatment duration and severity of depression on the effectiveness of cognitive/behavioural and psychodynamic/interpersonal psychotherapy. *J Consult Clin Psychol* 1994;**62**:522–34.

156. Elkin I, Gibbons RD, Shea MT, Sotsky SM, Watkins JT, Pilkonis PA, *et al*. Initial severity and differential treatment outcome in the NIMH treatment of depression collaboration research program. *J Consult Clin Psychol* 1995;**63**:841–7.

157. Lidren DM, Watkins P, Gould RA, Clum GA, Asterino M, Tulloch HL. A comparison of bibliotherapy and group therapy in the treatment of panic disorder. *J Consult Clin Psychol* 1994; **62**:865–9.

158. Huppert JD, Bufka LF, Barlow DH, Gorman JM, Shear MK, Woods SW. Therapists, therapist variables, and cognitive-behavioral therapy outcome in a multicenter trial for panic disorder. *J Consult Clin Psychol* 2001;**69**:747–55.

159. Davidson K, Scott J, Schmidt U, Tata P, Thornton S, Tyrer P. Therapist competence and clinical outcome in the Prevention of Parasuicide by Manual Assisted Cognitive Behaviour Therapy trial: the POPMACT study. *Psychol Med* 2004; **34**:855–63.

160. Drummond MF, O'Brien BJ, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes*. 2nd ed. Oxford: Oxford University Press; 1997.

161. Barber JA, Thompson SG. Analysis of cost data in randomized trials: an application of the non-parametric bootstrap. *Stat Med* 2000;**19**:3219–36.

162. Briggs AH, Wonderling DE, Mooney CZ. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. *Health Econ* 1997;**6**:327–40.

163. Barrett B, Byford S. Acceptable mental health care? A new decision-making tool for cost-effectiveness analysis. *Ment Health Res Rev* 2003; **9**:39–42.

164. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves – facts, fallacies and frequently asked questions. *Health Econ* 2004; **13**:405–15.

165. Andrews G. Should depression be managed as a chronic disease? *BMJ* 2001;**322**:419–21.

166. Reich JH, Vasile RG. Effect of personality disorders on the treatment outcome of Axis I conditions: an update. *J Nerv Ment Dis* 1993; **181**:475–84.

167. Quinton D, Gulliver L, Rutter M. A 15–20 year follow-up of adult psychiatric patients. Psychiatric disorder and social functioning. *Br J Psychiatry* 1995;**167**:315–23.

168. Olfson M, Pincus H. Outpatient psychotherapy in the United States, 2: patterns of utilisation. *Am J Psychiatry* 1994;**151**:1289–94.

169. Aveline M. Brief therapy: benefits and limitations including cognitive and behavioural therapies. *Curr Opin Psychiatry* 1996;**9**:193–7.

170. Durham RC, Swan JS, Fisher PL. Complexity and collaboration in routine practice of CBT: what doesn't work with whom and how might it work better? *J Ment Health* 2000;**9**:429–44.

171. Wells A. A metacognitive model and therapy for generalized anxiety disorder. *Clin Psychol Psychother* 1999;**6**:86–95.

172. Lewis G, Bebbington P, Brugha T, Farrell M, Gill B, Jenkins R, *et al*. Socio-economic status, standard of living, and neurotic disorder. *Int Rev Psychiatry* 2003;**15**:91–6.

173. Ronalds C, Creed F, Stone K, Webb S, Tomenson T. Outcome of anxiety and depressive disorders in primary care. *Br J Psychiatry* 1997; **171**:427–33.

174. Andrews G. The burden of anxiety and depressive disorders. *Curr Opin Psychiatry* 1998;**11**:121–3.

175. Rice DP, Miller LS. Health economics and cost implications of anxiety and other mental disorders in the United States. *Br J Psychiatry* 1998; **173** Suppl 34:4–9.

176. Greenberg P, Sisitsky T, Kessler R, Finkelstein SN, Berndt ER, Davidson JR, *et al*. The economic burden of anxiety disorders in the 1990s. *J Clin Psychiatry* 1999;**60**:427–35.

177. Marks I. The maturing of therapy. *Br J Psychiatry* 2002;**180**:200–4.

178. Bowen R, Senthilselvan A, Barale A. Physical illness as an outcome of chronic anxiety disorders. *Can J Psychiatry* 2000;**45**:459–63.

179. Chambers JA, Power KG, Durham RC. The relationship between trait vulnerability and anxiety and depressive diagnoses at long-term follow-up of generalised anxiety disorder. *J Anxiety Disord* 2004; **18**:587–607.

180. Riso LP, Miyatake RK, Thase ME. The search for determinants of chronic depression: a review of six factors. *J Affect Disord* 2002;**70**:103–16.

181. Barlow D, Hofmann S. Efficacy and dissemination of psychological treatments. In Clark DM, Fairburn C, editors. *Science and practice of cognitive behaviour therapy*. Oxford: Guilford Press; 1997. pp. 95–117.

182. Baer DM, Wolf MM, Risley TR. Some current dimensions of applied behavior analysis. *J Appl Behav Anal* 1968;**1**:91–7.

183. Hammarberg M. Penn inventory for posttraumatic stress disorder: psychometric properties. *Psychol Assess* 1992;**4**:67–76.

184. Foa, E. *Posttraumatic stress diagnostic scale: manual*. Minneapolis, MN: National Computer Systems; 1995.

185. Zinbarg RE, Barlow DH, Brown TA. Hierarchical structure and general factor saturation of the Anxiety Sensitivity Index: evidence and implications. *Psychol Assess* 1997;**9**:277–84.

186. Foa EB, Riggs DS, Dancu CV, Rothbaum BP. Reliability and validity of a brief instrument for assessing post-traumatic stress disorder. *J Trauma Stress* 1993;**6**:459–73.

187. Brett-Jones J, Garety PA, Hemsley D. Measuring delusional experiences: a method and its application. *Br J Clin Psychol* 1987;**163**:257–65.

188. Hogan TP, Awad AG, Eastwood R. A self report scale predictive of drug compliance in schizophrenics: reliability and discriminative validity. *Psychol Med* 1983;**13**:177–83.

189. Endicott J, Spitzer RL, Fleiss JL, Cohen J. The Global Assessment Scale. A procedure for measuring overall severity of psychiatric disturbance. *Arch Gen Psychiatry* 1976;**33**:766–71.

190. Haddock G, McCarron J, Tarrier N, Faragher EB. Scales to measure dimensions of hallucinations and delusions: the psychotic symptom rating scales (PSYRATS). *Psychol Med* 1999;**29**:879–89.

191. Andreasen NC. Scale for the Assessment of Negative Symptoms (SANS). *Br J Psychiatry* 1989; **155**:53–8.

192. Andreasen NC. *Scale for the Assessment of Positive Symptoms (SAPS)*. Iowa City, IA: University of Iowa; 1984.

Appendix 1

Tables of reviewed long-term follow-up studies of randomised treatment trials

TABLE 58 *Long-term follow-up of randomised treatment trials involving cognitive behaviour therapy (CBT) for generalised anxiety disorder (GAD)*

Author, date, country	Design RCT or RT	Diagnostic criteria	Treatment conditions	Main outcome measures	End of study results	Length of follow-up (months)	Follow-up outcome measures	Follow-up results	Comments
Barlow *et al.*, 1992[31] USA	RCT	DSM III-R	Relaxation (1; *n* = 10) CT (2; *n* = 13) Relaxation + CT (3; *n* = 11) WL control (4: *n* = 10)	Composite of several measures	1, 2, 3 > 4	6, 12, 24 (for participants in active treatment conditions only)	Treatment responder status defined as 20% improvement on at least 3 of 4 measures	No differences between active treatments Treatment gains modest but maintained across 2-year follow-up	Small numbers High attrition rates over follow-up period Clinically significant change achieved by 20–40% of participants at 2-year follow-up
Borkovec and Costello, 1993[23] USA	RT	DSM III–R	Nondirective therapy (ND) (1; *n* = 18) Relaxation (2; *n* = 20) CBT (3; *n* = 17)	Treatment responder status defined as 20% improvement on at least 3 of 4 measures	2 = 3 > 1	6, 12	Treatment responder status defined as 20% improvement on at least 3 of 4 measures	2 = 3 > 1 Highest gains in 3	ND failed to maintain gains at follow-up CBT and relaxation treatments maintained treatment gains
Borkovec *et al.*, 2002[28] USA	RT	DSM III–R and DSM-IV	CT (1; *n* = 23) Self-control desensitisation (2; *n* = 23) CBT (3; *n* = 23)	Composite measure of endstate functioning on 6 measures categorised into high and low	1 = 2 = 3 43–56% met criteria for high endstate functioning	6, 12	Composite measure of endstate functioning on 6 measures categorised into high and low	No differences between active treatments 28–52% met criteria for high endstate functioning	Treatment gains maintained over 2-year follow-up No differences found between treatment conditions Interpersonal difficulties associated with poor outcome

continued

TABLE 58 *Long-term follow-up of randomised treatment trials involving cognitive behaviour therapy (CBT) for generalised anxiety disorder (GAD) (cont'd)*

Author, date, country	Design RCT or RT	Diagnostic criteria	Treatment conditions	Main outcome measures	End of study results	Length of follow-up (months)	Follow-up outcome measures	Follow-up results	Comments
Butler *et al.*, 1991[30] UK	RCT	DSM III–R	CBT (1; *n* = 19) Behaviour therapy (2; *n* = 18) WL control (3; *n* = 19)	A variety of measures of anxiety, depression and cognition	1 > 2; 1 > 3 on all anxiety measures, 2 > 3 on one anxiety measure	6, 18	A variety of measures of anxiety, depression and cognition	1 > 2	Treatment gains generally maintained at follow-up particularly with CBT Behaviour therapy outcomes generally poor
Dugas *et al.*, 2003[32] Canada	RCT	DSM-IV	Group CBT (1; *n* = 25) WL control (2; *n* = 27)	A variety of measures of anxiety, depression and cognition	1 > 2	6, 12, 24	A variety of measures of anxiety, depression and cognition	Treatment gains maintained with 95% of treated participants no longer meeting criteria for GAD at follow-up	Significant attrition at follow-up Participants solicited by newspaper advertisement Limited demographic data
Durham *et al.*, 1999[26] UK	RT	DSM III–R	CT (1; *n* = 19) AP (2; *n* = 35) AMT (3; *n* = 16)	A variety of measures of anxiety, depression and cognition	1 > 2,3 3 > 2	6, 12	A variety of measures of anxiety, depression and cognition	1 > 2, but differences less apparent in less intensive treatment condition	Treatment gains generally maintained at follow-up particularly with CT in which two-thirds of participants in more intensive condition were recovered
Ladoucer *et al.*, 2000[27] Canada	RCT	DSM-IV	CBT (1; *n* = 14) WL control (2; *n* = 12)	A variety of measures of anxiety, depression and cognition	1 > 2	6, 12	A variety of measures of anxiety, depression and cognition	Treatment gains maintained with 77% of treated participants no longer meeting criteria for GAD at follow-up	Participants solicited by newspaper advertisement Limited numbers

continued

TABLE 58 *Long-term follow-up of randomised treatment trials involving cognitive behaviour therapy (CBT) for generalised anxiety disorder (GAD) (cont'd)*

Author, date, country	Design RCT or RT	Diagnostic criteria	Treatment conditions	Main outcome measures	End of study results	Length of follow-up (months)	Follow-up outcome measures	Follow-up results	Comments
Ost and Breitholtz, 2000[29] Sweden	RT	DSM-IV	CT (1; *n* = 18) Applied relaxation (2; *n* = 18)	A variety of measures of anxiety, depression and cognition	1 = 2	6, 12	A variety of measures of anxiety, depression and cognition	1 = 2 Treatment gains maintained with 56–67% meeting criteria for clinically significant change at follow-up	Participants recruited via medical referrals and newspaper advertisements
Seivewright *et al.*, 1998[34] UK	RT	DSM-III	CBT (1; *n* = 24) Medication (2; *n* = 22) Self-help (3; *n* =25)	A variety of measures of clinical status and social adjustment	1 = 2 = 3	60	Standardised measure of global outcome using a 5-point scale	1 = 2 = 3	60% of patients had a favourable outcome Initial diagnosis and original treatment of no predictive value
White, 1998[33] UK	RT	DSM-III-R	CT (1; *n* = 31) Behavioural therapy (2; *n* = 31) CBT (3; *n* = 26) Placebo (4; *n* = 10) WL control (5; *n* = 11)	Measures of anxiety, depression and coping	1, 2, 3, 4 > 5	6, 24	A variety of measures of anxiety, depression and cognition	1 = 2 = 3 Treatment gains maintained between 6 months and 2 years	Follow-up results reported for 1, 2 and 3 only Considerable attrition at 2-year follow-up

TABLE 59 *Long-term follow-up of randomised treatment trials involving cognitive behaviour therapy (CBT) for post-traumatic stress disorder (PTSD)*

Author, date, country	Design RCT or RT	Diagnostic criteria	Treatment conditions	Main outcome measures	End of study results	Length of follow-up (months)	Follow-up outcome measures	Follow-up results	Comments
Richards *et al.*, 1994[48] UK	RT	DSM III-R	Imaginal exposure/*in vivo* (1; *n* = 7) *In vivo* exposure/ imaginal (2; *n* = 7)	IES PTSD-SCL BDI FQ	1 = 2 1 = 2 1 > 2 1 = 2	1, 3, 6, 12	IES PTSD-SCL BDI FQ	Group differences not reported in detail	Aim: to assess efficacy of imaginal or *in vivo* exposure in crossover design All follow-up results fail to differentiate between treatment groups
Foa *et al.*, 1999[52] USA	RCT	DSM III–R	Prolonged exposure (PE) (1; *n* = 25) Stress inoculation training (SIT) (2; *n* = 26) Combined treatment (PE–SIT) (3; *n* = 30) WL control (4; *n* = 15)	PSS-I SAS BDI STAI-S	1 > 2, 3 > 4 1, 2, 3 > 4 1 > 2, 3 > 4 1 > 2, 3 > 4	3, 6, 12	PSS-I SAS BDI STAI-S	ns 1 > 2, 3 ns 1 > 2, 3	Aim: to assess efficacy of exposure-based and anxiety management-based treatments in female assault victims Follow-up results presented as separate 4 × 3 mixed ANCOVAs
Schnurr *et al.*, 2003[53] USA	RT	DSM-IV	Trauma-focused group therapy (1; *n* = 180) Present-centred group treatment (2; *n* = 180)	CAPS PTSD-SCL GHQ ASI	1 = 2 1 = 2 1 = 2 1 = 2	7, 12, 18, 24	CAPS PTSD-SCL GHQ ASI	1 = 2 1 = 2 1 = 2 1 = 2	Aim: to compare trauma-focused group psychotherapy vs present centred treatment that avoided trauma focus among Vietnam veterans

continued

TABLE 59 *Long-term follow-up of randomised treatment trials involving cognitive behaviour therapy (CBT) for post-traumatic stress disorder (PTSD) (cont'd)*

Author, date, country	Design RCT or RT	Diagnostic criteria	Treatment conditions	Main outcome measures	End of study results	Length of follow-up (months)	Follow-up outcome measures	Follow-up results	Comments
Tarrier *et al.*, 1999[49] UK	RT	DSM III-R	Imaginal exposure (1; *n* = 35) CT (2; *n* = 37)	CAPS Penn Inventory[183] IES BDI		12	CAPS Penn Inventory IES BDI	1 = 2 1 = 2 1 = 2 1 = 2	Results indicate benefit of two CBT treatments at 12 months
Bisson *et al.*, 2004[51] UK	RCT	DSM-IV	CBT + standard care (1; *n* = 76) Standard care (2; *n* = 76)	IES PTSD Diagnostic Scale[184] HADS	1 = 2 1 = 2 1 = 2	3, 13	IES PTSD Diagnostic Scale HADS	1 > 2 1 = 2 1 = 2	Aim: to examine efficacy of four-session CBT after physical injury
Power *et al.*, 2002[50] UK	RCT	DSM-IV	EMDR (1; *n* = 39) Exposure + cognitive restructuring (2; *n* = 37) WL control (3; *n* = 29)	CAPS IES MADRS HAM-A	1, 2 > 3 1, 2 > 3 1, 2 > 3 1, 2 > 3	15 (average)	IES MADRS HAM-A	1 = 2 1 > 2 1 = 2	WL patients were not included into long-term follow-up Long-term follow-up was an average rather than set point for all patients

ASI, Anxiety Sensitivity Index,[185] BDI, Beck Depression Inventory;[56] CAPS, Clinician-Administered PTSD Scale;[55] CBT, cognitive behavioural therapy; EMDR, eye-movement desensitisation and reprocessing; FQ, Fear Questionnaire;[137] GHQ, General Health Questionnaire;[145] HADS, Hospital Anxiety and Depression Index;[57] HAM-A, Hamilton Anxiety Scale;[130] IES, Impact of Event Scale;[54] MADRS, Montgomery Asberg Depression Rating Scale;[132] PSS-I, PTSD Symptom Scale – Interview;[186] PTSD-SCL, PTSD Symptom Check List;[147] SAS, Social Adjustment Scale;[133] STAI-S, State–Trait Anxiety Scale, state version.[58]

TABLE 60 *Randomised controlled clinical trials of cognitive behaviour therapy (CBT) for schizophrenia or similar*[a]

Author, year, country	Design RCT or RT	Diagnostic criteria	Treatment conditions	Main outcome measures	End of study results	Length of follow-up (months)	Follow-up outcome measures	Follow-up results	Comments
Barrowclough *et al.*, 2001[68] UK	RCT	ICD-10 DSM-IV	CBT + motivational interviewing + family intervention (1; *n* = 18) TAU (2; *n* = 18)	GAF PANSS Relapse	1 > 2 1 = 2 1 > 2	18	*Patients:* GAF PANSS *Carers:* BDI GHQ	 1 > 2 1 = 2 1 = 2 1 = 2	Main outcomes reported for 9 and 12 months following randomisation
Buchkremer, 1997[76] Germany	RCT	DSM-III-R	PMT + L (1; *n* = 32) PMT + L + KC (2; *n* = 35) PMT + CP (3; *n* = 34) PMT + CP + C (4; *n* = 33) L (5; *n* = 57)	Readmission BPRS SANS GAS	1, 2, 3, 4 = 5 4 > 5 1, 2, 3, 4 = 5 1, 2, 3, 4, =5 1, 2, 3, 4 > 5	60	Readmission BPRS SANS GAS	1, 2, 3, 4 = 5 4 > 5 1, 2, 3, 4 = 5 1, 2, 3, 4 = 5 1, 2, 3, 4 = 5	Main outcomes for this study are reported at 2 years post-randomisation
Drury, 1996[61] UK	RCT	WHO	CBT (1; *n* = 20) Activity/ recreation (2; *n* = 20)	PAS *12 weeks:* Positive Disorganisation Negative *9 months:* Positive Disorganisation Negative	 1 > 2 1 = 2 1 = 2 1 > 2 1 = 2 1 = 2	60	Readmission PAS	1 = 2 1 = 2	

continued

TABLE 60 *Randomised controlled clinical trials of cognitive behaviour therapy (CBT) for schizophrenia or similar[a] (cont'd)*

Author, year, country	Design RCT or RT	Diagnostic criteria	Treatment conditions	Main outcome measures	End of study results	Length of follow-up (months)	Follow-up outcome measures	Follow-up results	Comments
Garety, 1997[73] UK	RCT	DSM-III-R	CBT (1; *n* = 28) TAU (2; *n* = 32)	BPRS	1 > 2	18	BPRS Delusional[187] distress Hallucinations[187] frequency	1 > 2 1 > 2 1 > 2	
Haddock, 1999[70] UK	RCT	DSM-IV	CBT (1; *n* = 10) TAU (2; *n* = 11)	Days to discharge BPRS PSYRATS	1 = 2 1 = 2 1 = 2	24	Readmission	1 = 2	
Hogarty, 1997[89] USA	RCT	RDC	Personal therapy (1; *n* = 74) Family therapy (2; *n* = 50) Supportive therapy (3 *n* = 53)	Relapse (psychotic and affective)	Trial 1 1 > 2, 3 Trial 2 1 < 3				Outcomes for patients living with (Trial 1: *n* = 97) and outwith family (Trial 2: *n* = 54) were analysed separately Results are presented at 3 years following randomisation
Kemp, 1998[85] UK	RCT	DSM-III-R	Compliance therapy (1; *n* = 39) TAU (2; *n* = 35)	Compliance Attitudes to treatment[188] Insight[83] GAF	1 > 2 1 > 2 1 > 2 1 > 2	18	Compliance Attitudes to treatment Insight GAF	1 = 2 1 > 2 1 > 2 1 > 2	

continued

TABLE 60 *Randomised controlled clinical trials of cognitive behaviour therapy (CBT) for schizophrenia or similar[a] (cont'd)*

Author, year, country	Design RCT or RT	Diagnostic criteria	Treatment conditions	Main outcome measures	End of study results	Length of follow-up (months)	Follow-up outcome measures	Follow-up results	Comments
Lewis, 2002[71] UK	RCT	DSM-IV	CBT (1; *n* = 101) Supportive counselling (SC) (2; *n* = 106) Routine care (RC) (3; *n* = 102)	PANSS PSYRATS	1 > 2, 3 1 = 2 = 3	18	PANSS PSYRATS Readmission	1, 2 > 3 1, 2 > 3 1 = 2 = 3	CBT was associated with faster improvement compared with SC and RC
Sensky, 2000[65] UK	RCT	ICD-10	CBT (1; *n* = 46) Befriending (2; *n* = 44)	CPRS MADRS SANS	1 = 2 1 = 2 1 = 2	18	CPRS MADRS SANS	1 > 2 1 > 2 1 > 2	Post-treatment outcomes taken at 9 months
Startup, 2004[72] UK	RCT	DSM-IV	CBT (1; *n* = 47) TAU (2; *n* = 43)	SAPS SANS BPRS SFS	1 > 2 1 = 2 1 > 2 1 > 2	12	SAPS SANS BPRS SFS	1 > 2 1 > 2 1 > 2 1 > 2	Post-treatment outcomes taken at 6 months
Tarrier, 1999[74] UK	RCT	DSM-III-R	CBT (1; *n* = 33) SC (2; *n* = 26) RC (3; *n* = 28)	BPRS Readmission	1 > 2, 3 1, 2 < 3	12 24	BPRS Readmission BPRS Readmission	1 > 3 1 = 2 = 3 1, 2 > 3 1 = 2 = 3	Follow-ups took place 12 and 24 months post-treatment. Treatment duration was 3 months

[a] Based on the search strategy conducted by Cochrane Schizophrenia group (2004) supplemented by further search strategy combining SCHIZOPHRENIA with COGNITIV* and / or BEHAVIO* and / or THERAP* in CINAHL, EMBASE, PsycINFO and MEDLINE databases.

BPRS, Brief Psychiatric Rating Scale;[79] CBT, cognitive behavioural therapy; CP, Cognitive psychotherapy; CPRS, Comprehensive Psychiatric Rating Scale;[80] GAF, Global Assessment of Functioning;[114] GAS, Global Assessment Scale;[189] KC, Key-person counselling; L, Leisure group; MADRS, Montgomery Asberg Depression Rating Scale;[132] PANSS, Positive and Negative Syndrome Scale;[81] PAS, Psychiatric Assessment Scale;[82] PMT, psychoeducation; PSYRATS, Psychotic Symptom rating Scales;[190] RC, routine care; SANS, Scale for assessment of negative symptoms;[191] SAPS, Scale for assessment of positive symptoms;[192] SC, supportive counselling; SFS, Social Functioning Scale;[134] TAU, treatment as usual.

Appendix 2

Contact letters, information sheet and consent form

GP letter – example from Trial 3

Our Ref: KGP/JBP

Date: 03/03/00

Dear Dr

Long-term effectiveness of psychological therapy for anxiety disorders

I am writing to ask your permission to contact the following patients of yours to invite each of them to attend a follow-up interview in connection with this research project. The project is funded by the NHS Executive and has been approved by Forth Valley Health Board Ethics of Research Committee.

The following patients participated in a clinical trial of psychological therapy conducted in 1989/92 by the Forth Valley GP Research Group in conjunction with the University of Stirling.

Patients will be interviewed regarding diagnostic status, social adjustment and recent progress. At the end of the interview, we will seek consent from each patient for the Research Psychologist, Ms Julie Chambers, to inspect his/her medical records in order to obtain information on health economic issues.

Please do not hesitate to get in touch if you have any queries. All we require from you is permission to contact the patient(s). If we do not hear from you within two weeks of the date of this letter, we will assume that we may go ahead and contact the patient(s). A reply slip is enclosed.

Thank you for your help.

With kind regards.

Yours sincerely

Professor Kevin Power
Clinical Psychologist

Dr Donald Sharp
Clinical Psychologist

GP letter – example from Trial 3 (*continued*)

To:
Professor Kevin Power
Anxiety & Stress Research Centre
Department of Psychology
University of Stirling
STIRLING FK9 4LA

Long-term effectiveness of psychological therapy for anxiety disorders

With reference to your letter of 03/03/00

re:

I do / do not * give permission for you to contact each of the above patients of mine to seek his/her consent to participate in a follow-up interview in connection with the above research project:

If any of these patients give their permission for an inspection of their medical records in order to obtain information on health service usage, I am / am not * willing for the Research Psychologist, Ms Julie Chambers, to contact the surgery in order to arrange a convenient time for her to undertake this work.

Signed: ..... Date: .....

*** Delete as appropriate**

Note: The above details are correct to the best of our knowledge, but it is possible that there may be errors or omissions. Please amend as necessary and/or delete any patients who are not currently on your list.

Patient information sheet – example from Trial 3

CENTRAL SCOTLAND
LONG-TERM OUTCOME OF TREATMENT FOR PANIC ATTACKS

PATIENT INFORMATION SHEET

We invite you to participate in a research project which we believe to be of considerable importance. However, before you decide whether or not you wish to participate, we need to be sure that you understand, firstly, why we are doing it, and secondly, what it would involve if you agreed. We are therefore providing you with the following information. Please read it carefully and be sure to ask any questions you have, and, if you want, discuss it with outsiders. We will do our best to explain and to provide any further information you may ask for now or later. You do not have to make an immediate decision.

We are trying to find out as much information as possible about the effectiveness of treatments for various forms of emotional disorder including panic. This research project is concerned with the long-term effectiveness of a form of psychological treatment called cognitive-behavioural therapy. We are particularly interested in finding out how well this treatment compares with medication and whether or not it helps people to learn ways of coping with emotional difficulties that are useful once treatment has ended. The research will be of particular value in planning treatment and services for the future.

You participated in a clinical trial of treatment for panic attacks several years ago and your Doctor has given us permission to contact you. The project involves an interview to find out how you are getting on at the moment and to seek your views on the value of the treatment you received. There will also be some questionnaires to fill in. In order that we may determine how successful this treatment is in the long term, it is important that we hear as many views as possible.

We would also like to seek your permission to look at your medical notes in order to obtain information on the treatment you have received from the NHS over the last two years. We are interested in the amount of treatment you have received so that we can compare this with a similar period before you entered the clinical trial. We are not interested in the details of the treatment, just the type of treatment and its length. All the information we collect will be treated confidentially and will be used to prepare research reports in which no mention will be made of individual patients.

Becoming involved in the project will not affect any of your existing healthcare arrangements. You will continue to receive the help you need from your Doctors and other health service staff. Participation in this study is entirely voluntary and you are free to refuse to take part or to withdraw from the study at any time without having to give a reason and without this affecting your future medical care or your relationship with medical staff looking after you. The Forth Valley Committee on Medical Research Ethics that has responsibility for scrutinising all proposals for medical research involving patients in Forth Valley has examined the proposal and has raised no objections from the point of view of medical ethics. As part of this scrutiny your research records may be examined by monitors from the Forth Valley Medical Research Ethics Committee.

Patient letter – example from Trial 3

Our Ref: JC/JBP
Enquiries: 01786 467678

Date: 26th February 2000

Personal and Confidential

«Title» «FirstName» «LastName»
«Address_1»
«Address_2»
«Town»
«Postal_Code»

Dear «Title» «LastName»

Long-term effectiveness of treatment for panic attacks

Your GP, «GP_Name», has given permission for me to invite you to participate in a research project which is seeking to determine the long-term effectiveness of treatment for panic attacks. You may recall that you participated in a clinical trial of such treatment at some time between 1989 and 1993, and we would be very interested to hear your views of the treatment you received and to know how you are getting on now. This normally involves a one-hour interview and the completion of some questionnaires. If you agree to being interviewed, I will be happy to answer any questions you may have at the start of the interview.

We are keen to arrange an appointment at a place and time that will be convenient for you. This would preferably be at your health centre or at the University of Stirling, but if this is too difficult for you, a home visit can be arranged. I enclose a reply slip with a stamped addressed envelope. Thank you for your help.

Yours sincerely

Ms Julie Chambers
Research Psychologist

Patient letter – example from Trial 3 (*continued*)

To:
Ms Julie Chambers
Anxiety & Stress Research Centre
Department of Psychology
University of Stirling
STIRLING FK9 4LA

From:
«Title» «FirstName» «LastName»
«Address_1»
«Address_2»
«Town» «Postal_Code»

Long term effectiveness of treatment for panic attacks

With reference to your letter of 26th February 2000 in connection with the research project:

I would prefer to be seen:

at «GP_Practice» Health Centre / at the University of Stirling

The best time(s) for me are:

Monday / Tuesday / Wednesday / Thursday / Friday *

mornings / afternoons / early evening *

It would be helpful if we could have a daytime telephone number where we could contact you when making appointments. If you are happy for us to ring you, please enter your telephone number below:

.....

*** delete as appropriate**

Patient consent form

TITLE OF PROPOSAL

Long-term outcome of cognitive-behavioural therapy (CBT) clinical trials in central Scotland.

CONSENT FORM

(The patient should complete this form himself/herself)

	PLEASE CROSS OUT AS NECESSARY
Have you read the Patient Information Sheet?	YES/NO
Have you had an opportunity to ask questions and discuss this study?	YES/NO
Have you received satisfactory answers to all of your questions?	YES/NO
Have you received enough information about the study?	YES/NO
Who have you spoken to? Dr/Mr/Ms	
Do you understand that participation is entirely voluntary?	YES/NO
Do you understand that you are free to withdraw from the study: • at any time? • without having to give a reason for withdrawing? • without this affecting your future medical care?	YES/NO
Do you agree to take part in this study?	YES/NO
Do you give permission for us to look at your medical notes for details of NHS treatment that is relevant to the study?	YES/NO

Patient's Signature ..... Date .....

Patient's name in block letters .....

Telephone number where patient can be contacted:

..... (Home) ..... (Work)

Psychologist's Signature ..... Date .....

Appendix 3

Case note collection forms

Case note data collection – anxiety disorder studies (Trials 1–8)

LONG-TERM FOLLOW-UP STUDY
CASE NOTES – PRE-TRIAL

Patient ref: | Study No:
Name: | GP:
Date to trial: | Pre-treatment records – from | to

GP/Surgery visits (inc. nurse etc.)

Date	What for	Practice GP	Practice Other	Home day	Home out hrs
	Total				
	Total GP visits				

Hospital admissions (including A&E visits)

Date	Hospital/dept	Ailment	Length of stay (days)

Total hospital admissions:

Drugs prescribed (Note: Include all repeat prescriptions)

Drug	Dosage	Duration	Date

GP referrals for tests

Date	What test	Date	What test

Referrals to other services

Date	Profession/Name of clinician	What referred for	No. of visits

Case note data collection – psychosis studies (Trials 9 and 10)

LONG-TERM FOLLOW-UP STUDY: MEDICAL NOTES – 2 YEARS PRE-FOLLOW-UP

Sheet No.

Patient Ref:		Name					
Consultant				CMHT:		Key worker:	
Date followed up		Pre-follow-up records		From:		To:	

CMHT VISITS (Including nurses, psychiatrists, OT, Social Worker, Psychologist)

DATE	PROFESSION	GRADE	No. of visits	Date From/to:

DRUG	DOSAGE	DURATION	
		From	To:

Case note data collection – psychosis studies (Trials 9 and 10) (*continued*)

LONG-TERM FOLLOW-UP STUDY: MEDICAL NOTES – 2 YEARS PRE-FOLLOW-UP

Sheet No. ____

HOSPITAL ADMISSIONS (Including day hospital, Mental Health Act)

Patient Ref:		Pre-follow-up records :	From:		To:	

DATE	HOSPITAL	DAY (D)/ INPATIENT (I)	MENTAL HEALTH ACT (Y/N)	SECTION TYPE	DURATION		No. of days/visits /nights
					From	To	

Appendix 4

Flow diagrams for Trials 1–10

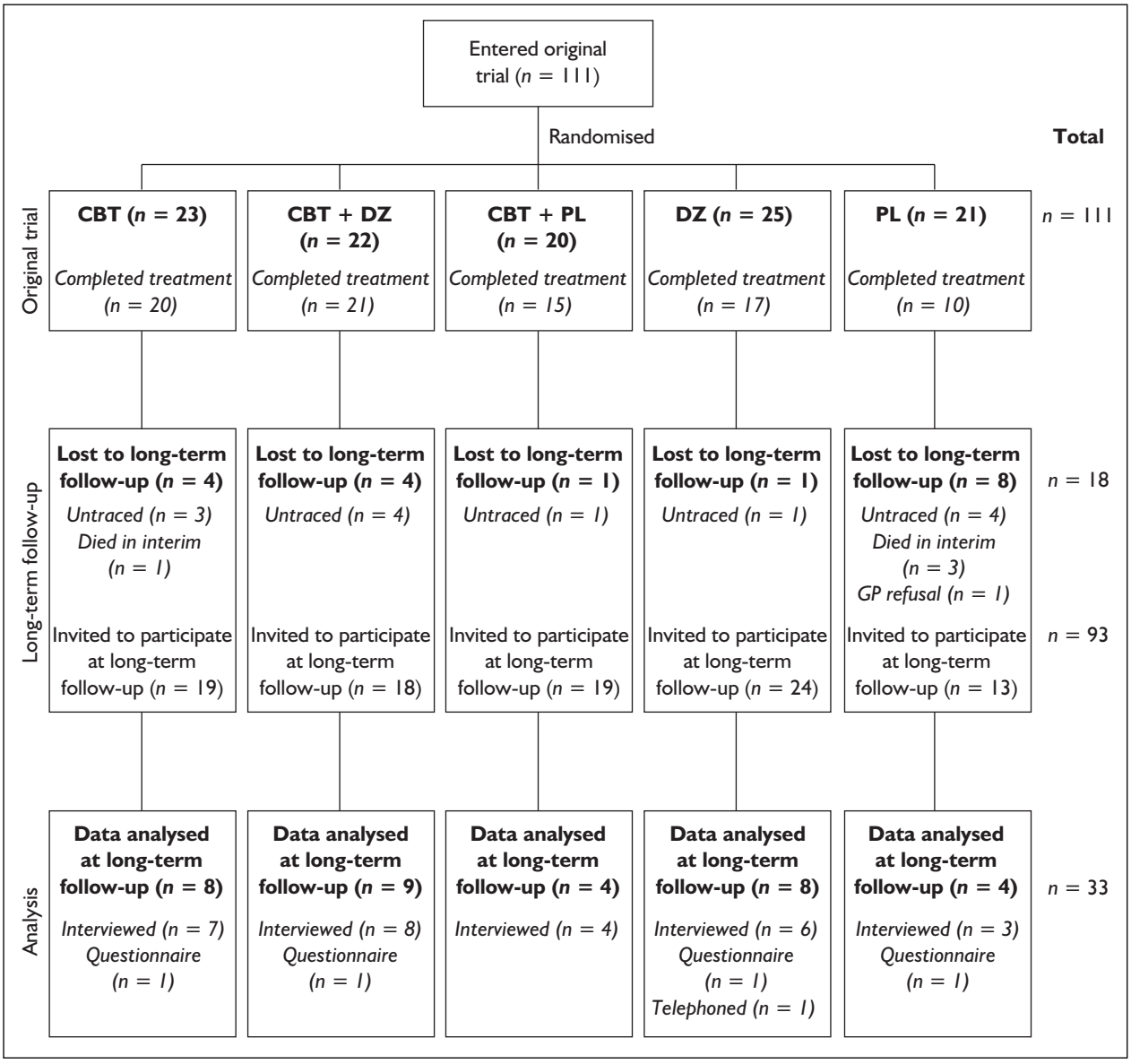


FIGURE 29 *Study 1: flow diagram of participants from original trial to long-term follow-up*

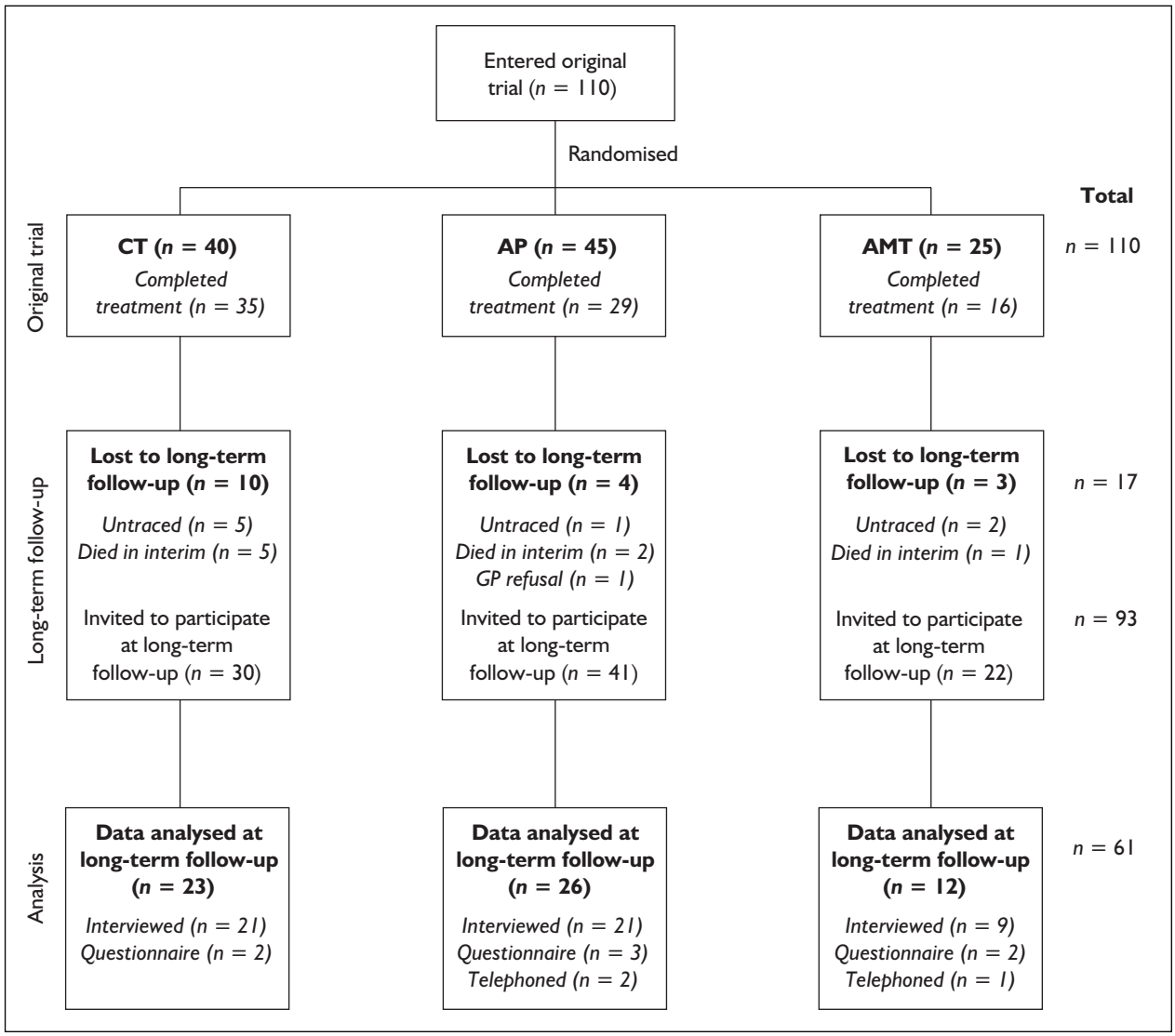


FIGURE 30 *Study 2: flow diagram of participants from original trial to long-term follow-up*

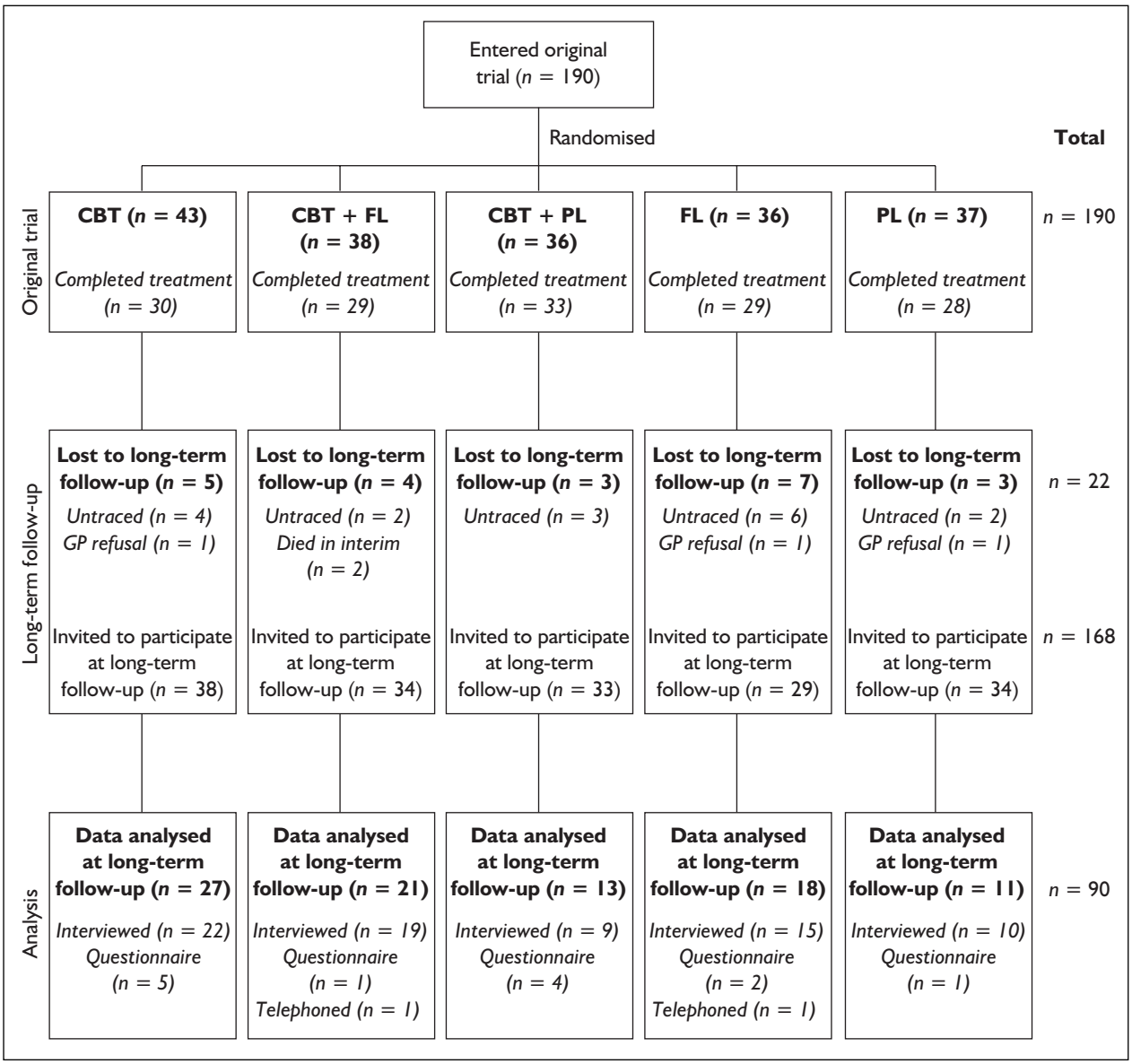


FIGURE 31 *Study 3: flow diagram of participants from original trial to long-term follow-up*

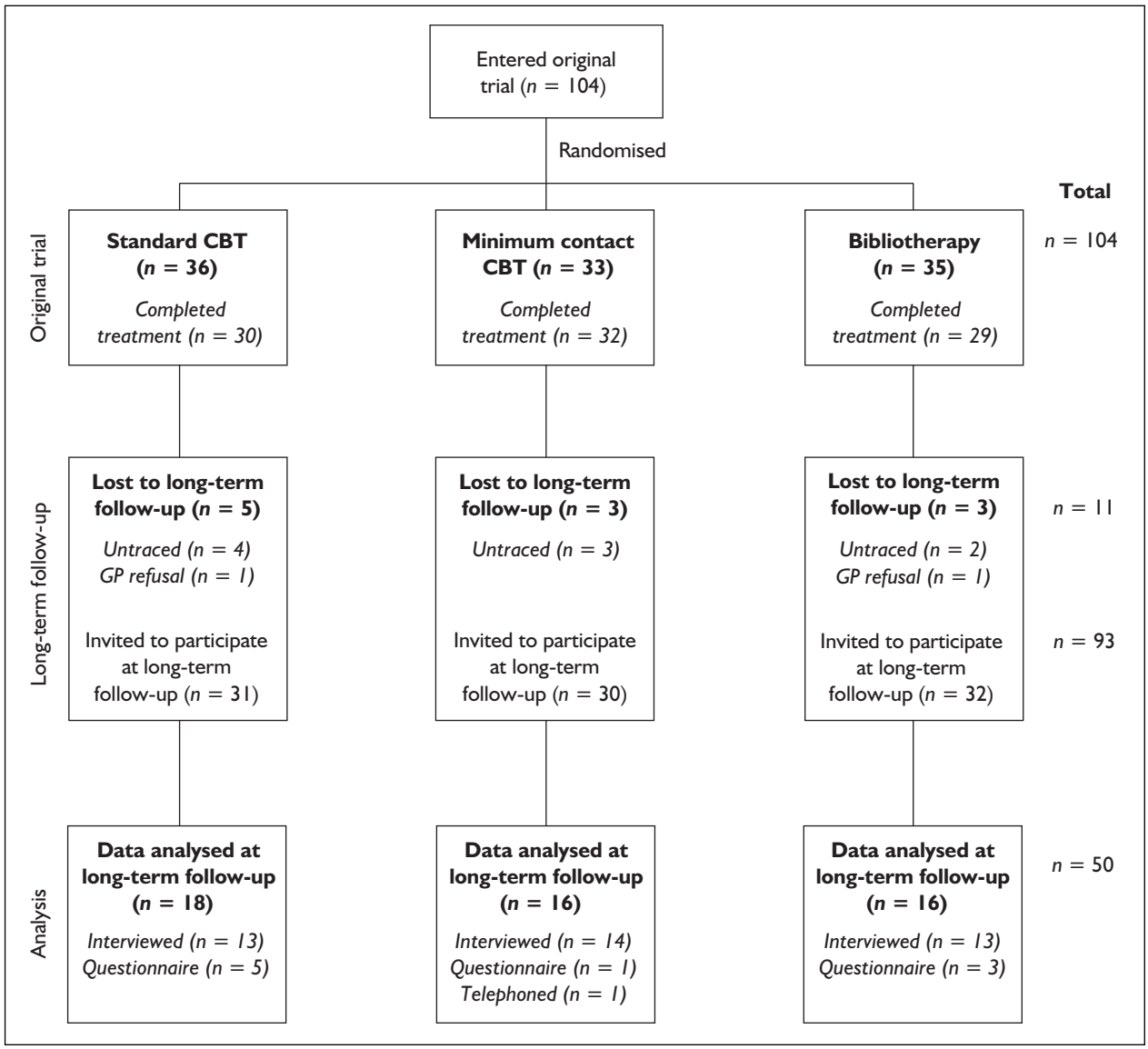


FIGURE 32 *Study 4: flow diagram of participants from original trial to long-term follow-up*

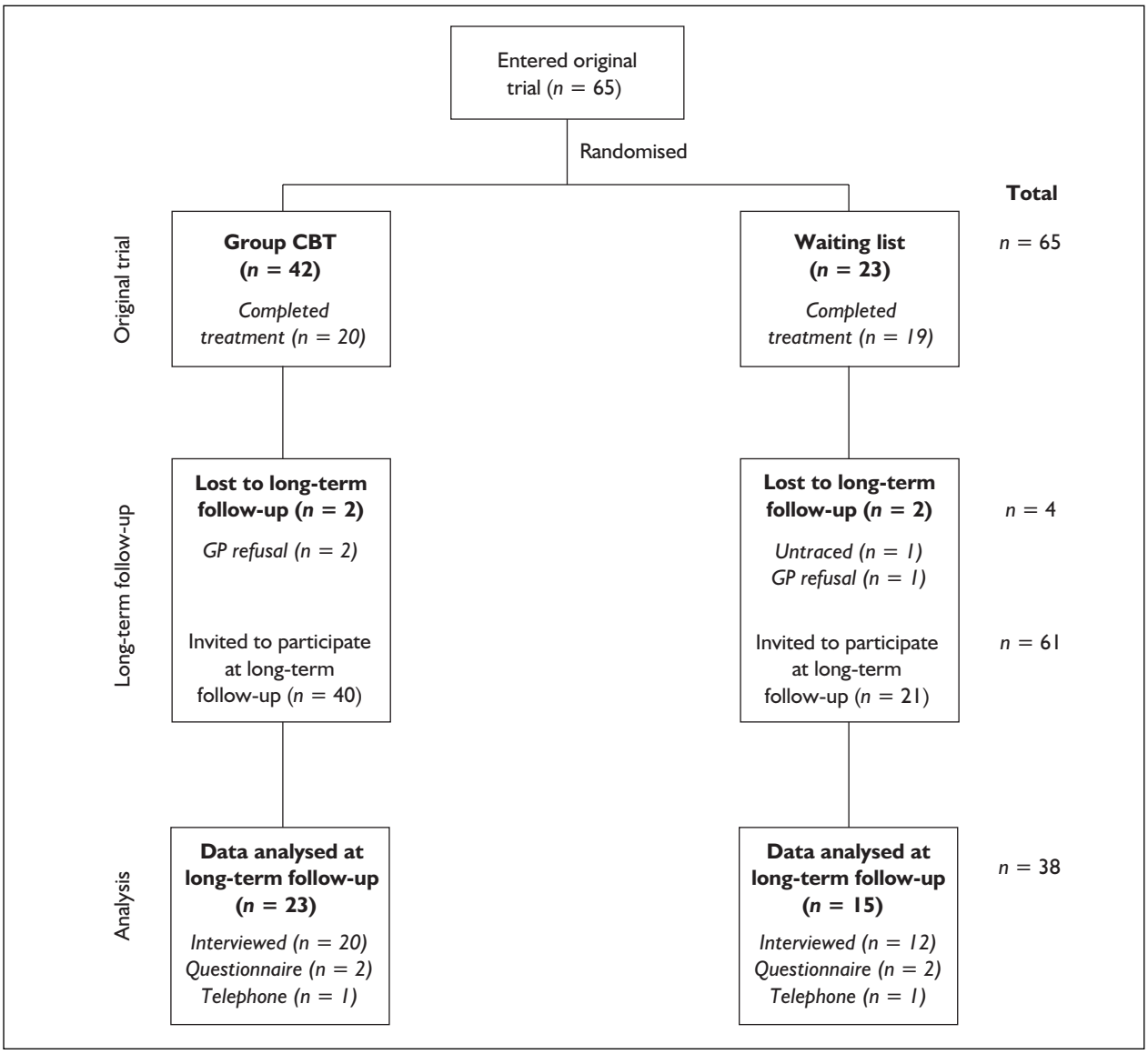


FIGURE 33 *Study 5: flow diagram of participants from original trial to long-term follow-up*

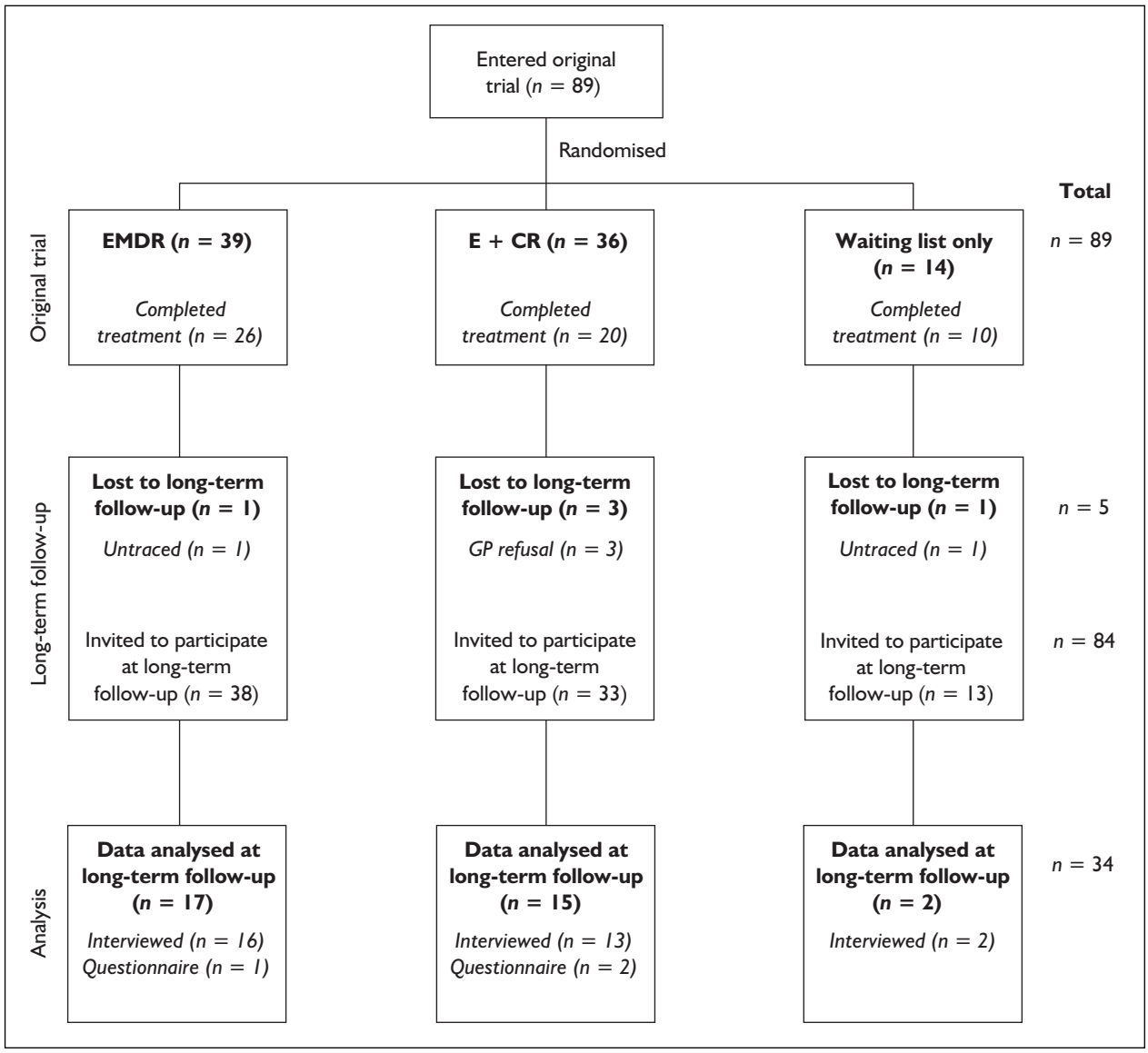


FIGURE 34 *Study 6: flow diagram of participants from original trial to long-term follow-up*

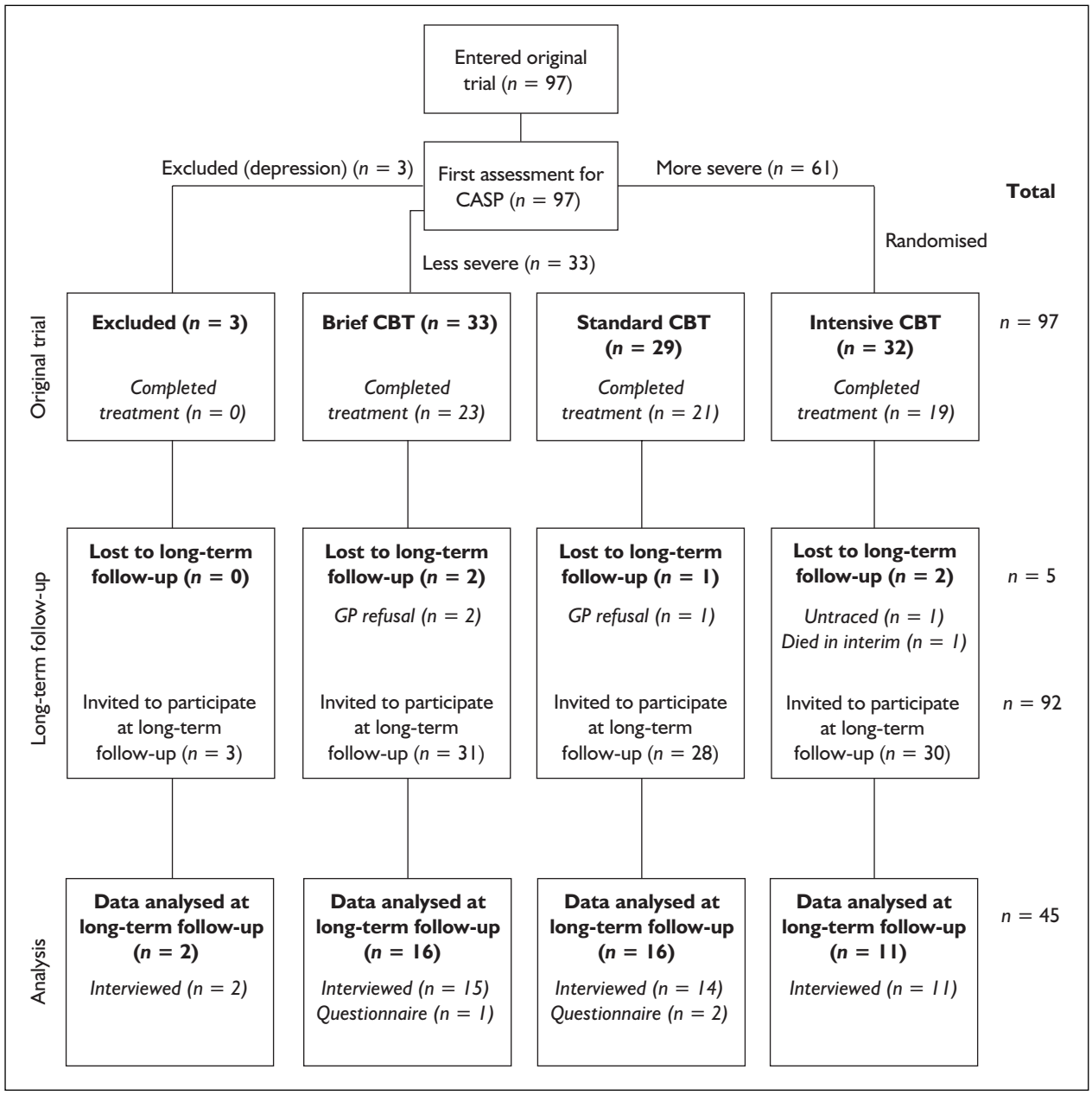


FIGURE 35 *Study 7: flow diagram of participants from original trial to long-term follow-up*

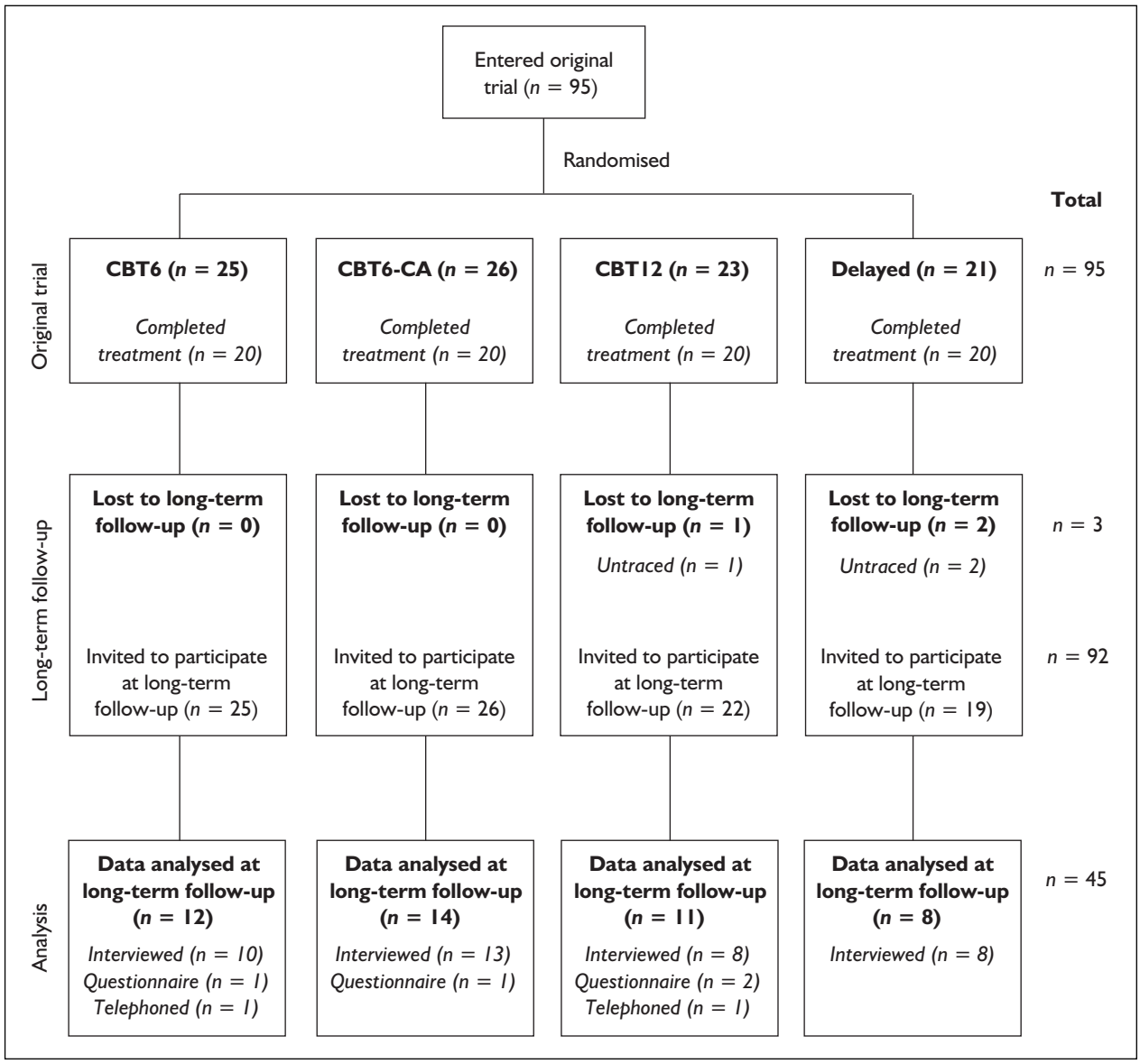


FIGURE 36 *Study 8: flow diagram of participants from original trial to long-term follow-up*

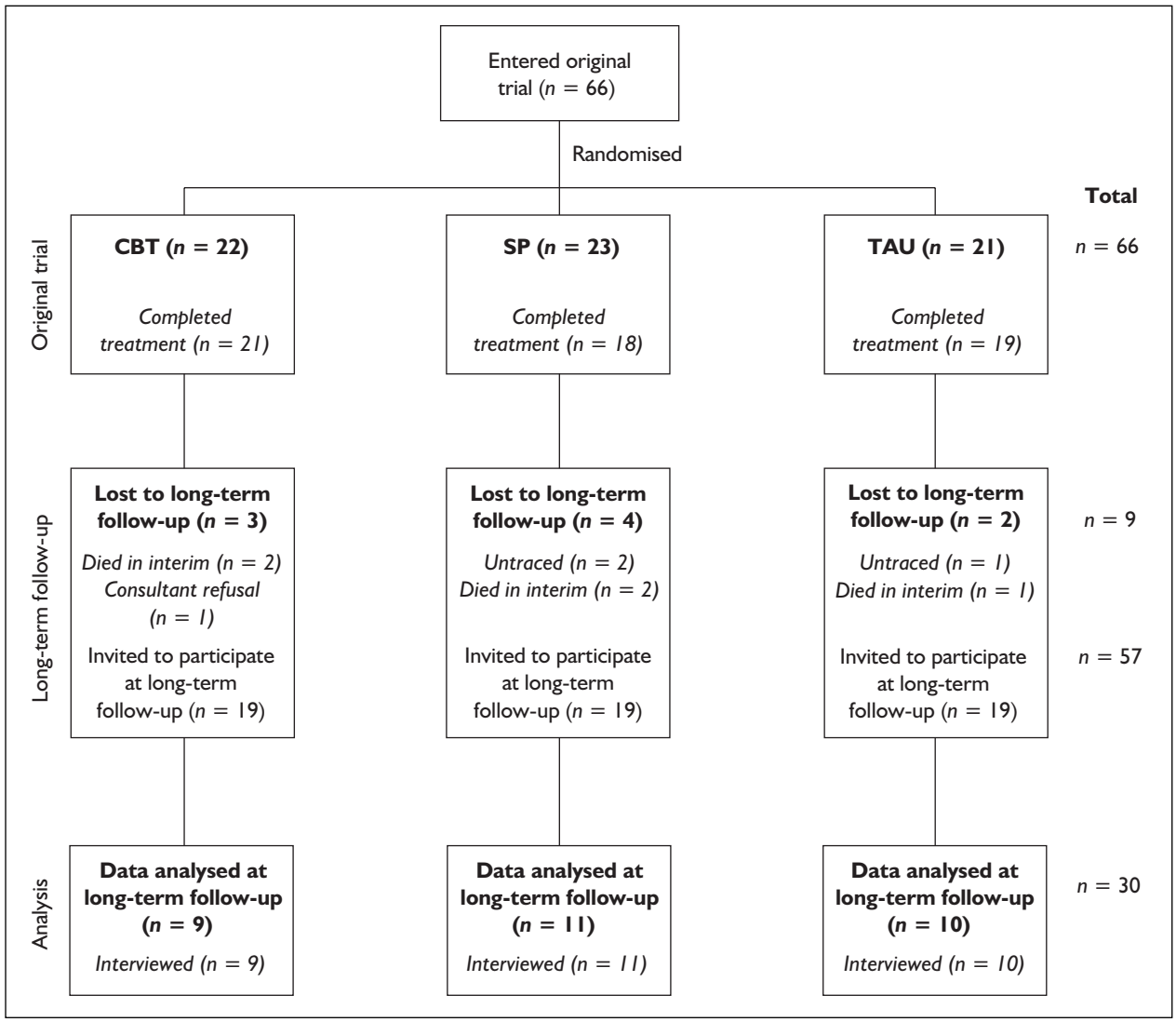


FIGURE 37 *Study 9: flow diagram of participants from original trial to long-term follow-up*

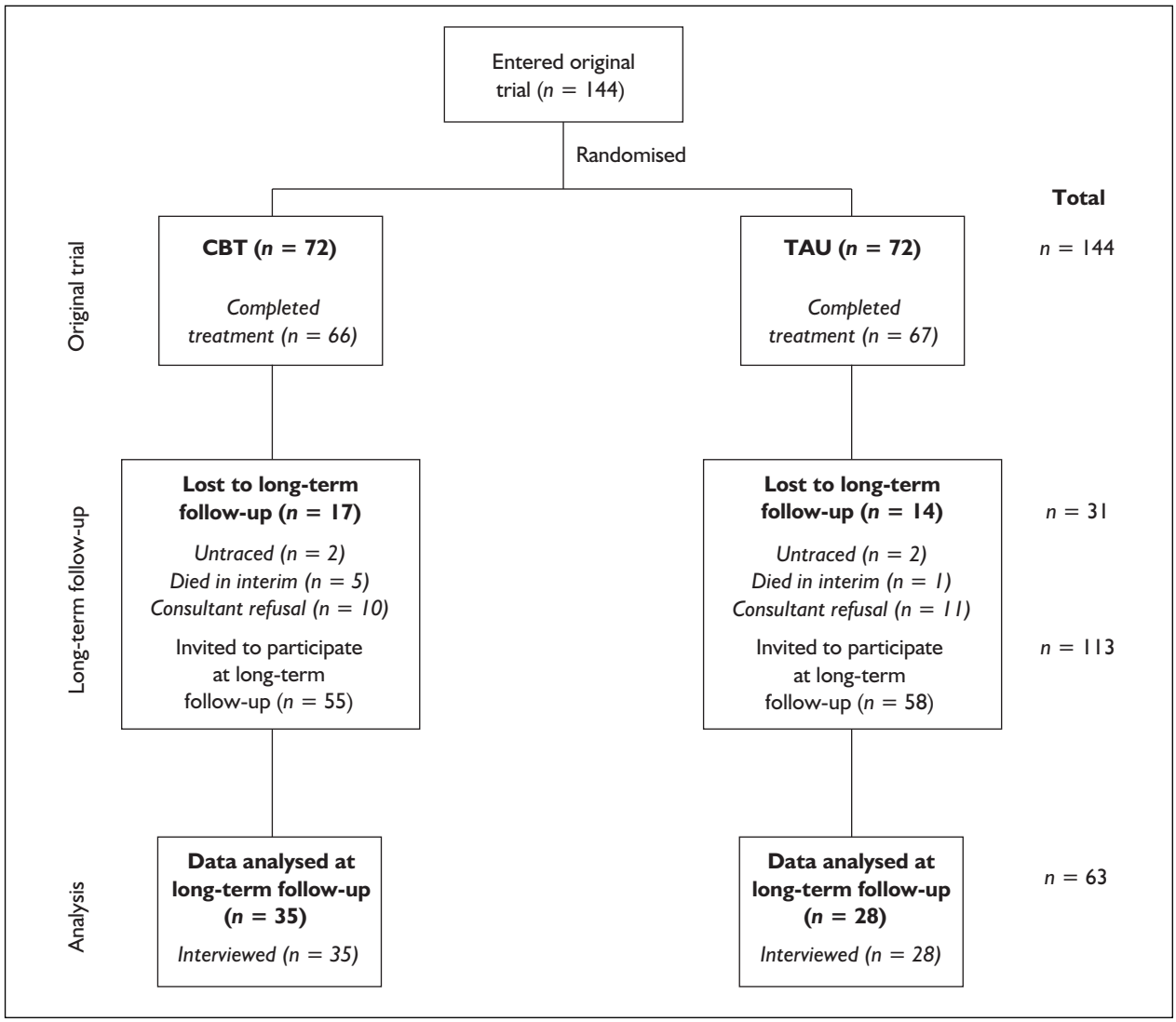


FIGURE 38 *Study 10: flow diagram of participants from original trial to long-term follow-up*

Health Technology Assessment Programme

Prioritisation Strategy Group

Members

Chair,
Professor Tom Walley,
Director, NHS HTA Programme, Department of Pharmacology & Therapeutics, University of Liverpool

Professor Bruce Campbell, Consultant Vascular & General Surgeon, Royal Devon & Exeter Hospital

Dr Edmund Jessop, Medical Advisor, National Specialist, Commissioning Advisory Group (NSCAG), Department of Health, London

Professor Jon Nicholl, Director, Medical Care Research Unit, University of Sheffield, School of Health and Related Research

Dr John Reynolds, Clinical Director, Acute General Medicine SDU, Radcliffe Hospital, Oxford

Dr Ron Zimmern, Director, Public Health Genetics Unit, Strangeways Research Laboratories, Cambridge

HTA Commissioning Board

Members

Programme Director,
Professor Tom Walley,
Director, NHS HTA Programme, Department of Pharmacology & Therapeutics, University of Liverpool

Chair,
Professor Jon Nicholl,
Director, Medical Care Research Unit, University of Sheffield, School of Health and Related Research

Deputy Chair,
Professor Jenny Hewison,
Professor of Health Care Psychology, Academic Unit of Psychiatry and Behavioural Sciences, University of Leeds School of Medicine

Dr Jeffrey Aronson
Reader in Clinical Pharmacology, Department of Clinical Pharmacology, Radcliffe Infirmary, Oxford

Professor Deborah Ashby, Professor of Medical Statistics, Department of Environmental and Preventative Medicine, Queen Mary University of London

Professor Ann Bowling, Professor of Health Services Research, Primary Care and Population Studies, University College London

Dr Andrew Briggs, Public Health Career Scientist, Health Economics Research Centre, University of Oxford

Professor John Cairns, Professor of Health Economics, Public Health Policy, London School of Hygiene and Tropical Medicine, London

Professor Nicky Cullum, Director of Centre for Evidence Based Nursing, Department of Health Sciences, University of York

Mr Jonathan Deeks, Senior Medical Statistician, Centre for Statistics in Medicine, University of Oxford

Dr Andrew Farmer, Senior Lecturer in General Practice, Department of Primary Health Care, University of Oxford

Professor Fiona J Gilbert, Professor of Radiology, Department of Radiology, University of Aberdeen

Professor Adrian Grant, Director, Health Services Research Unit, University of Aberdeen

Professor F D Richard Hobbs, Professor of Primary Care & General Practice, Department of Primary Care & General Practice, University of Birmingham

Professor Peter Jones, Head of Department, University Department of Psychiatry, University of Cambridge

Professor Sallie Lamb, Professor of Rehabilitation, Centre for Primary Health Care, University of Warwick

Professor Stuart Logan, Director of Health & Social Care Research, The Peninsula Medical School, Universities of Exeter & Plymouth

Dr Linda Patterson, Consultant Physician, Department of Medicine, Burnley General Hospital

Professor Ian Roberts, Professor of Epidemiology & Public Health, Intervention Research Unit, London School of Hygiene and Tropical Medicine

Professor Mark Sculpher, Professor of Health Economics, Centre for Health Economics, Institute for Research in the Social Services, University of York

Dr Jonathan Shapiro, Senior Fellow, Health Services Management Centre, Birmingham

Ms Kate Thomas, Deputy Director, Medical Care Research Unit, University of Sheffield

Ms Sue Ziebland, Research Director, DIPEx, Department of Primary Health Care, University of Oxford, Institute of Health Sciences

Current and past membership details of all HTA 'committees' are available from the HTA website (www.ncchta.org)

Diagnostic Technologies & Screening Panel

Members

Chair,
Dr Ron Zimmern, Director of the Public Health Genetics Unit, Strangeways Research Laboratories, Cambridge

Ms Norma Armston, Lay Member, Bolton

Professor Max Bachmann Professor of Health Care Interfaces, Department of Health Policy and Practice, University of East Anglia

Professor Rudy Bilous Professor of Clinical Medicine & Consultant Physician, The Academic Centre, South Tees Hospitals NHS Trust

Dr Paul Cockcroft, Consultant Medical Microbiologist and Clinical Director of Pathology, Department of Clinical Microbiology, St Mary's Hospital, Portsmouth

Professor Adrian K Dixon, Professor of Radiology, University Department of Radiology, University of Cambridge Clinical School

Dr David Elliman, Consultant Paediatrician/ Hon. Senior Lecturer, Population Health Unit, Great Ormond St. Hospital, London

Professor Glyn Elwyn, Primary Medical Care Research Group, Swansea Clinical School, University of Wales Swansea

Mr Tam Fry, Honorary Chairman, Child Growth Foundation, London

Dr Jennifer J Kurinczuk, Consultant Clinical Epidemiologist, National Perinatal Epidemiology Unit, Oxford

Dr Susanne M Ludgate, Medical Director, Medicines & Healthcare Products Regulatory Agency, London

Professor William Rosenberg, Professor of Hepatology, Liver Research Group, University of Southampton

Dr Susan Schonfield, Consultant in Public Health, Specialised Services Commissioning North West London, Hillingdon Primary Care Trust

Dr Phil Shackley, Senior Lecturer in Health Economics, School of Population and Health Sciences, University of Newcastle upon Tyne

Dr Margaret Somerville, PMS Public Health Lead, Peninsula Medical School, University of Plymouth

Dr Graham Taylor, Scientific Director & Senior Lecturer, Regional DNA Laboratory, The Leeds Teaching Hospitals

Professor Lindsay Wilson Turnbull, Scientific Director, Centre for MR Investigations & YCR Professor of Radiology, University of Hull

Professor Martin J Whittle, Associate Dean for Education, Head of Department of Obstetrics and Gynaecology, University of Birmingham

Dr Dennis Wright, Consultant Biochemist & Clinical Director, Pathology & The Kennedy Galton Centre, Northwick Park & St Mark's Hospitals, Harrow

Pharmaceuticals Panel

Members

Chair,
Dr John Reynolds, Chair Division A, The John Radcliffe Hospital, Oxford Radcliffe Hospitals NHS Trust

Professor Tony Avery, Head of Division of Primary Care, School of Community Health Services, Division of General Practice, University of Nottingham

Ms Anne Baileff, Consultant Nurse in First Contact Care, Southampton City Primary Care Trust, University of Southampton

Professor Stirling Bryan, Professor of Health Economics, Health Services Management Centre, University of Birmingham

Mr Peter Cardy, Chief Executive, Macmillan Cancer Relief, London

Professor Imti Choonara, Professor in Child Health, Academic Division of Child Health, University of Nottingham

Dr Robin Ferner, Consultant Physician and Director, West Midlands Centre for Adverse Drug Reactions, City Hospital NHS Trust, Birmingham

Dr Karen A Fitzgerald, Consultant in Pharmaceutical Public Health, National Public Health Service for Wales, Cardiff

Mrs Sharon Hart, Head of DTB Publications, *Drug & Therapeutics Bulletin*, London

Dr Christine Hine, Consultant in Public Health Medicine, South Gloucestershire Primary Care Trust

Professor Stan Kaye, Cancer Research UK Professor of Medical Oncology, Section of Medicine, The Royal Marsden Hospital, Sutton

Ms Barbara Meredith, Lay Member, Epsom

Dr Andrew Prentice, Senior Lecturer and Consultant Obstetrician & Gynaecologist, Department of Obstetrics & Gynaecology, University of Cambridge

Dr Frances Rotblat, CPMP Delegate, Medicines & Healthcare Products Regulatory Agency, London

Professor Jan Scott, Professor of Psychological Treatments, Institute of Psychiatry, University of London

Mrs Katrina Simister, Assistant Director New Medicines, National Prescribing Centre, Liverpool

Dr Richard Tiner, Medical Director, Medical Department, Association of the British Pharmaceutical Industry, London

Dr Helen Williams, Consultant Microbiologist, Norfolk & Norwich University Hospital NHS Trust

Therapeutic Procedures Panel

Members

Chair,
Professor Bruce Campbell,
Consultant Vascular and General Surgeon, Department of Surgery, Royal Devon & Exeter Hospital

Dr Aileen Clarke,
Reader in Health Services Research, Public Health & Policy Research Unit, Barts & the London School of Medicine & Dentistry, London

Dr Matthew Cooke, Reader in A&E/Department of Health Advisor in A&E, Warwick Emergency Care and Rehabilitation, University of Warwick

Dr Carl E Counsell, Clinical Senior Lecturer in Neurology, Department of Medicine and Therapeutics, University of Aberdeen

Ms Amelia Curwen, Executive Director of Policy, Services and Research, Asthma UK, London

Professor Gene Feder, Professor of Primary Care R&D, Department of General Practice and Primary Care, Barts & the London, Queen Mary's School of Medicine and Dentistry, London

Professor Paul Gregg, Professor of Orthopaedic Surgical Science, Department of General Practice and Primary Care, South Tees Hospital NHS Trust, Middlesbrough

Ms Bec Hanley, Co-Director, TwoCan Associates, Hurstpierpoint

Ms Maryann L Hardy, Lecturer, Division of Radiography, University of Bradford

Professor Alan Horwich, Director of Clinical R&D, Academic Department of Radiology, The Institute of Cancer Research, London

Dr Simon de Lusignan, Senior Lecturer, Primary Care Informatics, Department of Community Health Sciences, St George's Hospital Medical School, London

Professor Neil McIntosh, Edward Clark Professor of Child Life & Health, Department of Child Life & Health, University of Edinburgh

Professor James Neilson, Professor of Obstetrics and Gynaecology, Department of Obstetrics and Gynaecology, University of Liverpool

Dr John C Pounsford, Consultant Physician, Directorate of Medical Services, North Bristol NHS Trust

Karen Roberts, Nurse Consultant, Queen Elizabeth Hospital, Gateshead

Dr Vimal Sharma, Consultant Psychiatrist/Hon. Senior Lecturer, Mental Health Resource Centre, Cheshire and Wirral Partnership NHS Trust, Wallasey

Dr L David Smith, Consultant Cardiologist, Royal Devon & Exeter Hospital

Professor Norman Waugh, Professor of Public Health, Department of Public Health, University of Aberdeen

Expert Advisory Network

Members

Professor Douglas Altman, Director of CSM & Cancer Research UK Med Stat Gp, Centre for Statistics in Medicine, University of Oxford, Institute of Health Sciences, Headington, Oxford

Professor John Bond, Director, Centre for Health Services Research, University of Newcastle upon Tyne, School of Population & Health Sciences, Newcastle upon Tyne

Mr Shaun Brogan, Chief Executive, Ridgeway Primary Care Group, Aylesbury

Mrs Stella Burnside OBE, Chief Executive, Office of the Chief Executive. Trust Headquarters, Altnagelvin Hospitals Health & Social Services Trust, Altnagelvin Area Hospital, Londonderry

Ms Tracy Bury, Project Manager, World Confederation for Physical Therapy, London

Professor Iain T Cameron, Professor of Obstetrics and Gynaecology and Head of the School of Medicine, University of Southampton

Dr Christine Clark, Medical Writer & Consultant Pharmacist, Rossendale

Professor Collette Clifford, Professor of Nursing & Head of Research, School of Health Sciences, University of Birmingham, Edgbaston, Birmingham

Professor Barry Cookson, Director, Laboratory of Healthcare Associated Infection, Health Protection Agency, London

Professor Howard Cuckle, Professor of Reproductive Epidemiology, Department of Paediatrics, Obstetrics & Gynaecology, University of Leeds

Dr Katherine Darton, Information Unit, MIND – The Mental Health Charity, London

Professor Carol Dezateux, Professor of Paediatric Epidemiology, London

Mr John Dunning, Consultant Cardiothoracic Surgeon, Cardiothoracic Surgical Unit, Papworth Hospital NHS Trust, Cambridge

Mr Jonothan Earnshaw, Consultant Vascular Surgeon, Gloucestershire Royal Hospital, Gloucester

Professor Martin Eccles, Professor of Clinical Effectiveness, Centre for Health Services Research, University of Newcastle upon Tyne

Professor Pam Enderby, Professor of Community Rehabilitation, Institute of General Practice and Primary Care, University of Sheffield

Mr Leonard R Fenwick, Chief Executive, Newcastle upon Tyne Hospitals NHS Trust

Professor David Field, Professor of Neonatal Medicine, Child Health, The Leicester Royal Infirmary NHS Trust

Mrs Gillian Fletcher, Antenatal Teacher & Tutor and President, National Childbirth Trust, Henfield

Professor Jayne Franklyn, Professor of Medicine, Department of Medicine, University of Birmingham, Queen Elizabeth Hospital, Edgbaston, Birmingham

Ms Grace Gibbs, Deputy Chief Executive, Director for Nursing, Midwifery & Clinical Support Services, West Middlesex University Hospital, Isleworth

Dr Neville Goodman, Consultant Anaesthetist, Southmead Hospital, Bristol

Professor Alastair Gray, Professor of Health Economics, Department of Public Health, University of Oxford

Professor Robert E Hawkins, CRC Professor and Director of Medical Oncology, Christie CRC Research Centre, Christie Hospital NHS Trust, Manchester

Professor Allen Hutchinson, Director of Public Health & Deputy Dean of ScHARR, Department of Public Health, University of Sheffield

Dr Duncan Keeley, General Practitioner (Dr Burch & Ptnrs), The Health Centre, Thame

Dr Donna Lamping, Research Degrees Programme Director & Reader in Psychology, Health Services Research Unit, London School of Hygiene and Tropical Medicine, London

Mr George Levvy, Chief Executive, Motor Neurone Disease Association, Northampton

Professor James Lindesay, Professor of Psychiatry for the Elderly, University of Leicester, Leicester General Hospital

Professor Julian Little, Professor of Human Genome Epidemiology, Department of Epidemiology & Community Medicine, University of Ottawa

Professor Rajan Madhok, Medical Director & Director of Public Health, Directorate of Clinical Strategy & Public Health, North & East Yorkshire & Northern Lincolnshire Health Authority, York

Professor David Mant, Professor of General Practice, Department of Primary Care, University of Oxford

Professor Alexander Markham, Director, Molecular Medicine Unit, St James's University Hospital, Leeds

Dr Chris McCall, General Practitioner, The Hadleigh Practice, Castle Mullen

Professor Alistair McGuire, Professor of Health Economics, London School of Economics

Dr Peter Moore, Freelance Science Writer, Ashtead

Dr Sue Moss, Associate Director, Cancer Screening Evaluation Unit, Institute of Cancer Research, Sutton

Mrs Julietta Patnick, Director, NHS Cancer Screening Programmes, Sheffield

Professor Tim Peters, Professor of Primary Care Health Services Research, Academic Unit of Primary Health Care, University of Bristol

Professor Chris Price, Visiting Chair – Oxford, Clinical Research, Bayer Diagnostics Europe, Cirencester

Professor Peter Sandercock, Professor of Medical Neurology, Department of Clinical Neurosciences, University of Edinburgh

Dr Eamonn Sheridan, Consultant in Clinical Genetics, Genetics Department, St James's University Hospital, Leeds

Dr Ken Stein, Senior Clinical Lecturer in Public Health, Director, Peninsula Technology Assessment Group, University of Exeter

Professor Sarah Stewart-Brown, Professor of Public Health, University of Warwick, Division of Health in the Community Warwick Medical School, LWMS, Coventry

Professor Ala Szczepura, Professor of Health Service Research, Centre for Health Services Studies, University of Warwick

Dr Ross Taylor, Senior Lecturer, Department of General Practice and Primary Care, University of Aberdeen

Mrs Joan Webster, Consumer member, HTA – Expert Advisory Network

Feedback

The HTA Programme and the authors would like to know your views about this report.

The Correspondence Page on the HTA website (http://www.ncchta.org) is a convenient way to publish your comments. If you prefer, you can send your comments to the address below, telling us whether you would like us to transfer them to the website.

We look forward to hearing from you.

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 23 8059 5639 Email: hta@soton.ac.uk
http://www.ncchta.org

ISSN 1366-5278