THE LANCET Psychiatry

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Goodyer IM, Reynolds S, Barrett B, et al. Cognitive behavioural therapy and short-term psychoanalytical psychotherapy versus a brief psychosocial intervention in adolescents with unipolar major depressive disorder (IMPACT): a multicentre, pragmatic, observer-blind, randomised controlled superiority trial. *Lancet Psychiatry* 2016; published online Nov 30. http://dx.doi.org/10.1016/S2215-0366(16)30378-9.

Appendix i-iv

Effectiveness And Cost-Effectiveness Of Cognitive Behaviour Therapy And Short-Term Psychoanalytic Psychotherapy Compared With Brief Psychosocial Intervention In Maintaining Reduced Depressive Symptoms 12 months after end of treatment in Adolescents with Unipolar Major Depression (IMPACT): A Pragmatic Superiority Randomised Controlled Trial

Ian M Goodyer, Shirley Reynolds, Barbara Barrett, Sarah Byford, Bernadka Dubicka, Jonathan Hill, Fiona Holland, Raphael Kelvin, Nick Midgley, Chris Roberts, Rob Senior, Mary Target, Barry Widmer, Paul Wilkinson, Peter Fonagy

Appendix i: Planned interventions and summary of clinical protocols

Planned interventions

IMPACT was a pragmatic superiority randomised controlled trial comparing the relative clinical effectiveness of three psychological treatments each with evidence of clinical efficacy being associated with clinical remission in the short term (i.e. 3-6 months) . These treatments are available in CAMHS NHS practice although distribution around the UK is not standardised. The three treatment approaches tested in this study were all manualised.

A duty of care by clinical staff to patients was observed in all clinical arms. This included parent support and engagement, explanation of treatment principles, maintenance and support of family during individual treatment, individual risk management strategies and contact with other agencies where appropriate.

Comprehensive treatment protocols were developed for the trial and designed for delivery by practitioners working in routine NHS CAMHS settings. The rational for using treatment manuals as guides to therapy is that:

- Manuals aid dissemination of treatment methods into clinical practice.
- They help to standardize the intervention between therapists and across site.
- They form the basis for audiotape ratings of treatment adherence and differentiation and thus ensure that the interventions have been given properly in keeping with the trial protocol.

The three treatments differed in the total number of sessions they offered over the study period. The number of sessions offered for each treatment were as follows:

- BPI – Up to 12 sessions, consisting of up to 8 individual andup to 4 family/parent sessions, to be delivered over 20 weeks.
- STPP – Up to 28 individual sessions plus up to 7 parent sessions to be delivered over 30 weeks.
- CBT - Up to 20 individual sessions plus up to 4 family/parent sessions to be delivered over 30 weeks.

The treatments are described below.

1. Brief Psychosocial Intervention (BPI)

Brief Psychosocial Intervention (BPI) is a brief structured intervention for the treatment of moderate to severe unipolar major depression in adolescents. [1,2] The clinical care approach originally used in the ADAPT RCT was the basis for BPI used in this trial [3,4,5] In the ADAPT study, the forerunner of BPI was described as Specialist Clinical Care (SCC: referred to as a non-manualised treatment as usual [TAU] in CAMHS) together with Fluoxetine 20mg-60mg daily, was as effective as TAU+Fluoxetine+CBT for moderate to severely depressed adolescents in routine NHS practice. [5] SCC was reformulated for the current study and described into a treatment manual. [23] Prescribing an SSRI is not a part of BPI per se but can be added and fully integrated if improvement is not judged to be occurring after 2-4 weeks as per the NICE guidelines of 2005.

Meta analytic studies of adolescent psychotherapies highlight the central therapeutic importance of care that is structured, evidence driven and founded on interpersonal effectiveness, warmth and trust. [6,7] The incorporation of collaborative care for depression in adults has been shown to provide added value for the treatment of depression in adults over above psychological and or medication treatments. [8,9]

BPI intervention is based on re-structuring and codification of the principles and practises found in the domains of skilled assessment, listening, information giving, advising, problem solving, safety, caring and explaining about adolescent depression. The duration and number of treatment sessions in the BPI manual is based on clinical experience gained with using specialist clinical care in the ADAPT study.

BPI was delivered in this study as the standard reference psychosocial intervention. Emphasis was placed on the importance of psychoeducation about depression, and action oriented, goal-focused and interpersonal activities as therapeutic strategies. Specific advice was given on improving and maintaining mental and physical hygiene, engaging in pleasurable activities, engaging and maintaining schoolwork and peer relations and diminishing solitariness. BPI did not use cognitive or reflective analytic techniques. There was therefore no discussion of unconscious conflict and no deliberate effort to modify maladaptive models of attachment relationships. Neither was there any focus on changing cognitions and negative cognition-driven behaviours were not deconstructed. BPI consisted of up to 12 sessions, consisting of up to 8 individual and up to 4 family/parent sessions, delivered over 20 weeks. Liaison with external agencies and personnel e.g. teachers, social care and peer group were commonly undertaken.

Case management in BPI

Since BPI case management has a rational and relational framework case management is founded on the three principles of:

- Interpersonal effectiveness.
- Understanding of mental states.
- Activation and problem solving.

The case management process is integrated through the development of a formulation which is a general construct summarising the probable relationship between the above 3 constructs. The formulation is developed as a series of prospective working hypotheses that can be tested and evaluated against subsequent progress within the therapy. BPI is delivered within this framework in up to 12 sessions, consisting of up to 8 individual and up to 4 family/parent sessions, over 20 weeks.

Therapy was delivered with the following strategies and principles being utilised throughout.

- Effective engagement, activation and problem solving.
- Diagnostic accuracy, and mental state evaluation.
- Sharing understanding and knowledge of the impairments and consequences of symptoms; the "lived experience" including effects in other settings such as school or peer relationships.
- Attention to accuracy in conducting a risk assessment and its management.
- Sharing aetiological description: defining risk and protective factors.
- A psycho-educative approach that at all points aims to help "activate" and empower, including parents and family as necessary.
- An approach that includes understanding of the role of medication, its appropriate use and how it sits within the care package
- A jointly agreed, collaboratively developed, and shared, management plan
- All delivered in a fashion that can help the child, young person and parents to manage and cope with their emotional expression.

Therapists, training and supervision

BPI therapists in this study were drawn from a range of professional backgrounds including mental health nursing, clinical psychology, psychiatry and mental health social work. The majority (>80%) of therapists were however psychiatrists in specialist CAMHS training as well as consultants. In the IMPACT trial to be eligible for training as a BPI therapist clinicians had to:

- Have had a minimum of 6 months supervised or independent work in a multidisciplinary child and adolescent mental health setting.
- Have already established sufficient competence and skills to be allowed to undertake independent mental health assessment and treatment of adolescents with moderate to severe depression.

BPI practitioners had basic training in BPI: reading of the manual; confirmation by the supervising clinician that they met the criteria to become a BPI therapist; attendance at a BPI training day; continued access to the BPI manual and ongoing supervision fitting in with usual local CAMHS NHS supervisory structures. The regional leads for BPI met and problem solved supervisory issues in relation to BPI on a regular basis across the IMPACT study period.

2. Short Term Psychoanalytic Psychotherapy (STPP)

Psychoanalytic psychotherapy with children and young people is a well-established specialist treatment for emotional and developmental difficulties in childhood and adolescence, with an emerging evidence-base. [70-71] It is one of several psychological therapies recommended by NICE as equally effective in the acute treatment of child and adolescent depression. [1] Its intellectual roots are drawn from psychoanalysis, child development, attachment theory and developmental psychopathology.

In this trial all therapists were approved as psychoanalytically trained by the Association of Child Psychotherapists UK. The short-term model of psychoanalytic psychotherapy (STPP) used shares therapeutic principles with time-limited psychodynamic work for adults with depression for which there is now a substantial evidence-base. [2] It is a 28-session model, with parents or carers being offered up to 7 additional sessions by a separate parent worker. STPP aims to elaborate and increase the coherence of the young person's mental models of attachment relationships and thereby improve their capacity for affect regulation as well as the capacity for making and maintaining positive relationships with other. [3]
The STPP method [2,3,4] draws on a long history in the UK of psychoanalytic work with depressed children and young people [74] including an unpublished manual used in an earlier clinical trial, in which short-term psychoanalytic psychotherapy for children with depression demonstrated good outcomes. [5] As with the other manuals used in the IMPACT study, the STPP manual [71] provided a guide to practice but not a recipe or a step-by-step guide, and drew on the existing skills and training of child and adolescent psychotherapists already working in the NHS.

STPP aims to elaborate and increase the coherence of the young person's mental models of attachment relationships and thereby improve their capacity for affect regulation as well as the capacity for making and maintaining positive relationships with others. When treatment is successful, it should free the young person to engage in normal adolescent development including educational attainment and independent peer group development involving a degree of separation from their primary carers. [6,7, 8]

The techniques of child and adolescent psychotherapy are primarily based on close and detailed observation of the relationship the child or young person makes with their therapist. The therapist introduces the therapeutic task to the young person as one of understanding feelings and difficulties in their life. The therapist's stance is non-judgemental and enquiring and conveys the value of understanding: the aim is to put into words conscious and unconscious thoughts and feelings. Through actions and words, the therapist attempts to convey an openness to all forms of psychic experience – current preoccupations, memories, day-dreams, nocturnal dreams and phantasies – but will be attuned specifically to evidence of unconscious phantasies which underlie the young person's relationship to self and others. This attentiveness to unconscious phenomena is specific to psychoanalytic psychotherapy, and is related to the theoretical importance attributed to these deeper less accessible layers of the mind.

With all adolescents, most particularly those with difficult early years' experiences, there is a need for the therapist to be in a state of mind characterised by availability to the reception of projected contents (anxieties,

affects, uncertainties) of the adolescent's mind. The patient's experience of the therapist receiving, holding in mind, and thinking about this projected material is a central feature of the therapy. The adolescents are helped to gain ownership of previously disowned part of themselves, and are strengthened by identification with another person (i.e. the therapist) experienced as capable of making meaning in this way and thus enabling more mature thinking to take place.

The STPP therapist and/or parent worker requires an alertness to the need, at times, for active communication and liaison with other significant individuals and agencies in the adolescent's life. This may include external agencies such as school/college, youth and social services, and also mental health colleagues, including Child and Adolescent Psychiatrists, where there are issues about risk and a possible need for medication or hospitalization. Prescribing an SSRI is not a part of STPP per se but can be added and fully integrated if improvement is not judged to be occurring after 2-4 weeks as per the NICE guidelines of 2005. [12]

Support for parents or carers, offered concurrently and in parallel with individual therapy for children and adolescents, is a well-established practice in the UK. There is some evidence that psychoanalytic therapy is more effective when undertaken with concurrent parent support work. [7] Parent support aims to help with parental anxieties and develop greater understanding about their relationship to their son or daughter. The duration of treatment and number of sessions prescribed is based on prior studies and clinical experience with adolescent patients.

Therapists, training and supervision

To be eligible to practice as an STPP therapist in the IMPACT study the clinician had to:

- Have undertaken a four-year postgraduate professional training, leading to membership of the Association of Child Psychotherapists (ACP) or be fourth-year trainee members of the ACP.
- Those doing parent work were individuals with at least 6 months CAMHS experience following professional training in child psychotherapy, clinical or counselling psychology, child mental health nursing, family therapy or psychiatry.

STPP training was designed and delivered on the basis that prospective STPP practitioners already have all the fundamental competencies and skills required to deliver all the components of STPP. Building upon these existing skills STPP training for IMPACT comprised: reading of the STPP manual; confirmation by the clinician that they met the criteria to become an STPP therapist; and attendance at an STPP training day.

STPP supervision by a consultant Child and Adolescent Psychotherapist was provided as part of routine practice within the CAMHS team.

3. Cognitive Behaviour Therapy (CBT)

Cognitive Behaviour Therapy (CBT) therapy in this trial is based on the classical form originally developed for adults with depression. This posits that emotional disorders are characterised by pervasive information processing biases which increase vulnerability to depression in the context of environmental stress, and which maintain and amplify core symptoms of depression including hopelessness, low mood, and irritability. The

focus of CBT is to identify the information processing biases that maintain depression and low mood and to amend these through a process of *collaborative empiricism* between the therapist and client.

It was adapted for this study to include parental involvement, a large focus on engagement and an emphasis on the use of behavioural techniques. [1,2] CBT included up to 20 sessions plus up to 4 family/marital sessions over 30 weeks. CBT therapists were routine CAMHS clinicians and were either clinical psychologists, or other clinicians who had received post qualification training in CBT. CBT emphasizes 'collaborative empiricism': i.e. explicit, tangible and shared goals between therapist and young person, and clear structured sessions. CBT links thoughts, feelings and behaviours and techniques includes behavioural activation; identifying and challenging negative automatic thoughts; developing adaptive thoughts and relapse prevention. Topics introduced within a therapy session are extended and supported outside the session by tasks completed by the client between sessions and reviewed at each subsequent session. CBT was delivered to the adolescent alone or to the young person and parent(s) flexibly. A formulation was developed at the start of therapy, which included consideration of parental, and family factors in the development and maintenance of depression. Where it was considered relevant parent(s) were involved in therapy session, by negotiation, to support the young person in treatment.

In this study CBT was manualised and incorporated adaptations for working with adolescents (as opposed to adults) including inclusion of simplified and aged appropriate cognitive techniques as well as the flexibility to take a behavioural focus if cognitive work was considered too demanding for a young person. A number of additional amendments were made including a greater focus on engagement in therapy, on building the therapeutic alliance, and on working collaboratively with parents and schools. Parents were involved in treatment sessions as indicated by the formulation and the preferences of the family. There were no separate sessions for parents.

Treatment length for CBT was a maximum of 20 sessions, delivered weekly, tapering to every 2 weeks as needed for relapse prevention, plus up to 4 family/parent sessions. Sessions were structured with an agenda set by the therapist and young person at the start of every session and out of session assignments agreed between the therapist and young person. Typically, early sessions (1-4) focused on relationship building, understanding the young person's current presentation and experience, and psycho-education, including the CBT model. A provisional formulation of the young person's difficulties, incorporating family history, key life events and transitions, recent stressors, and coping strategies was developed with the young person (and parent where relevant). Subsequently the formulation guided treatment. This included using CBT techniques to treat non depressive comorbid symptoms of anxiety, obsessions and compulsions and oppositional behaviours.

Mid-treatment focused on identifying and modifying the behavioural and cognitive processes that maintained depression and low mood for that young person. Behavioural work included activity scheduling, ratings of mastery and pleasure and reinforcement of engagement in activities. Cognitive work included identifying dysfunctional and unhelpful automatic thoughts and thought challenging using a range of techniques including behavioural experiments. Modifications to the core CBT model, such as the use of mindfulness were permitted depending on the individual formulation. The end of treatment was marked by a focus on relapse prevention.

Typically this included a revisit to the formulation, identifying potential risk and vulnerability factors, problem solving, and building resilience. Prescribing an SSRI is not a part of CBT per se but can be added and fully integrated if improvement is not judged to be occurring after 2-4 weeks as per the NICE guidelines of 2005. [12]

Therapists, training and supervision

CBT therapists were NHS staff from a range of professional backgrounds including clinical and counselling psychology, nursing, and occupational therapy. They delivered CBT for depression as part of their routine clinical practice in multi-disciplinary Child and Adolescent Mental Health services.

CBT therapists had to have received specialist training in CBT, either as part of their core professional training (i.e. as a clinical psychologist) or as post qualification training (i.e. as a nurse or occupational therapist). They were eligible to be IMPACT CBT therapists if they routinely used CBT in their NHS clinical work and if they could demonstrate some pre or post qualification training in CBT.

CBT training was delivered as a one day workshop within services. It was designed as a top-up training for individuals who already had core CBT skills. The core features of the treatment manual were described and the practicalities and constraints of delivering CBT within the context of a research trial were discussed. All clinicians had copies of the CBT manual and familiarised themselves with it.

CBT supervision was provided as part of routine practice within the CAMHS team.

Prescribing of Fluoxetine during the trial

For all three arms Fluoxetine or another SSRI could be added where clinicians judged that combination therapy may accelerate the time to remission following NICE guidelines for a major depression episode in adolescents. [12] A test dose of 10 mg was given for 48 hours followed by 20 mg as a single dose. If there was no improvement within 2-4 weeks the dose can be adjusted upwards to 60 mg maximum.

References

BPI

1. Kelvin RG, Dubicka B, Wilkinson PO, Goodyer IM. Brief Psychosocial Intervention (BPI) : A Specialist Clinical Care Treatment Manua For CAMHS Use. 2010, University of Cambridge , online treatment manual, dev.psychiatry.cam.ac.uk.

2. Kelvin RG, Dubicka B, Wilkinson PO, Goodyer IM. A Brief Psychosocial Intervention for adolescent depression: unpublished manuscript. 2015. Website address.

3. Goodyer I, Dubicka B, Wilkinson P, Kelvin R, Roberts C, Byford S, et al. Selective serotonin reuptake inhibitors (SSRIs) and routine specialist care with and without cognitive behaviour therapy in adolescents with major depression: randomised controlled trial. *BMJ* 2007;**335**:142.

4. Kelvin R, Wilkinson P, Goodyer I. *Managing Acute Depressive Episodes: Putting it Together in Parctice*. In: Birmaher B, Rey J, editors. Treating Child and Adolescent Depression. Philadelphia, USA: Lippincott, Williams and Wilkins 2009 162-73.

5. Goodyer IM, Dubicka B, Wilkinson P, Kelvin R, Roberts C, Byford S, et al. A randomised controlled trial of cognitive behaviour therapy in adolescents with major depression treated by selective serotonin reuptake inhibitors.The ADAPT trial. *Health Technol Assess*. 2008;**12**:60.

6. Weisz JR, Jensen-Doss AJ, Hawley KM. Evidence-based youth psychotherapies versus usual clini- cal care: A meta-analysis of direct comparisons. *Am Psychologist* 2006;**61** 671–89.

7. Weisz JR, McCarty CA, Valeri SM. Effects of psychotherapy for depression in children and adoles- cents: A meta-analysis. . *Psychol Bulletin*. 2008;**132:** 132– 49.

8. Archer J, Bower P, Gilbody S, Lovell K, Richards D, Gask L, et al. *Collaborative care for depression and anxiety problems.* The Cochrane database of systematic reviews. 2012;**10**:CD006525.

9. Coventry PA, Hudson JL, Kontopantelis E, Archer J, Richards DA, Gilbody S, et al. Characteristics of effective collaborative care for treatment of depression: a systematic review and meta-regression of 74 randomised controlled trials. *PloS One* 2014;**9**:e108114.

STPP

1. Abbass AA, Rabung S, Leichsenring F, Refseth JS, Midgley N. Psychodynamic psychotherapy for children and adolescents: a meta-analysis of short-term psychodynamic models. *J Am Acad Child Adolesc Psychiatry* 2013;**52**: 863-75.

2. Midgley N, Kennedy E. Psychodynamic psychotherapy for children and adolescents: a critical review of the evidence base. *J Child Psychol Psychiatry* 2011;**37**:1-29.

3. Hopkins K, Crosland P, Elliott N, Bewley S, Clinical Guidelines Update Committee B. Diagnosis and management of depression in children and young people: summary of updated NICE guidance. *BMJ* 2015;**350**:h824.

4. Target M, Fonagy P, (2nd ed) - *The long-term followup of child analytic treatments (AFC3).* Fonagy P, editor. London: Institute of Psychanalysis; 2002:141-146.

5. Busch F, M. R, Shapiro T. *The treatment of depression.* American Psychiatric Association; Washington: 2004.

6. Trowell J, Joffe I, Campbell J, Clemente C, Almqvist F, Soininen M, et al. Childhood depression: a place for psychotherapy. An outcome study comparing individual psychodynamic psychotherapy and family therapy. Eu Child Adolesc Psychiatry. 2007;**16**:157-67.

7. Midgley N, Cregeen S, Hughes C, Rustin M. Psychoanalytic psychotherapy as a treatment for depression in adolescents. *Child Adolesc Psychiatr Clin N Am* 2012;**22**: 67-82.

8. Fonagy P, Target M. *The history and current status of outcome research at the Anna Freud Centre.* The Psychoanalytic study of the child. 2002;**57**:27-60.

CBT

1. Spirito A, Esposito-Smythers C, Wolff J, Uhl K. Cognitive-behavioral therapy for adolescent depression and suicidality. *Child Adolesc Psychiatr Clin N Am* 2011;**20**:191-204.

2. Graham P, Reynolds S (editors). *Cognitive Behaviour Therapy for Children and Families.* 3 ed. Cambridge: Cambridge University Press; 2015.

Appendix i (continued)

Psychological Adherence and Differentiation Instruments

Comparative Psychotherapy Process Scale – External Rater form (CPPS-ER)
The CPPS is a measure that assesses the degree to which a therapist uses techniques of psychodynamic-interpersonal (PI) and/or cognitive behavioural psychotherapy (CB) in an entire psychotherapy session.[1] Developed from an extensive empirical review of the comparative psychotherapy process literature, all items are rated on a seven-point Likert Scale, ranging from 0 ("not at all characteristic"), 2 ("somewhat characteristic"), 4 ("characteristic"), to 6 ("extremely characteristic").[1] The 20-item measure includes ten PI items and ten CB items, forming two distinct sub-scales. The psychometric properties of the CPPS have been well established in psychotherapy with adults.[1] Internal consistency of both scales has been good to excellent: Cronbach's α of ·82 to ·92 for the PI scale and ·75 to ·94 for the CB scale.[2,3] Inter-rater reliability is reported as rating from good through to excellent (ICC 0·6 to 0·75). [2,4]

Brief Psychological Intervention Scale (BPI-S)
The BPI-S is a new scale, developed specifically for use in this study to assess treatment adherence to BPI. The 18 key components of the BPI manualised treatment were identified using expert consensus in the IMPACT team. A pilot investigation conducted by the BPI experts used a sample of five tapes to develop the adherence scale. Following this phase the measure was operationalised as an eight-item measure with three 'core' and five 'general' items, rated as a Likert Scale (0 – no evidence, 1 – passing evidence, 2 – some evidence, to 3 –clear evidence).

The three core items are: (i) Activation and problem solving; (ii) Interpersonal effectiveness; and (iii) Attention to mental state-current presentation or diagnosis. The five general items are: (i) Attention to vulnerability and protective factors, (ii) Psycho-education; (iii) Setting case management within a BPI framework; (iv) Attending to the social context of the patient; and (v) Making effort to help the patient manage their emotional expression. These eight items were chosen to (a) capture important treatment principles (relevance), based on the BPI manual; and (b) cover all relevant treatment principles (comprehensiveness), as outlined in the BPI manual.

For each item, a score of two or more was considered an adequate level of adherence. Overall, a BPI therapy session was judged to be 'adherent' if:

i. At least two out of three 'core' items were rated as 2 or above

ii. And a total of at least four out of the eight items were rated as 2 or above.

When this revised standard was applied to the five taped sessions previously rated, 100% agreement was obtained between the experts who rated four sessions as adherent and one session as not adherent.

Training for five independent raters was completed over two days. The raters were all trained in BPI and experienced senior clinicians with medical and psychiatric qualifications, and achieved high levels of inter-rater reliability (>80%) by the end of the training. Feedback from the raters during the training process indicated high levels of face validity indicated by good comprehension of the BPI adherence scale and an understanding of the rating measure and procedure. Each session was watched in its entirety, and then rated by the two judges independently; but raters were not blind to the treatment arm, as only BPI sessions were rated using the BPI-S.

Appendix ii

Statistical Analysis Plan (SAP) and Supplementary Statistical Tables

IMPACT TRIAL Statistical and Health Economics Data Analysis Plan
Version 1.0
Chris Roberts, Sarah Byford, Fiona Holland, Barbara Barrett
18th November 2014
Approved By the IMPACT Investigator Committee
Chief Investigator Ian M Goodyer

This document considers the final statistical and health economic analyses for the IMPACT trial. The purpose of the SAP is to document the confirmatory statistical analyses of the trial thereby controlling for statistical analyses bias. The statistical analyses follow the principles of ICH E9.

1. Aims

Improving Mood with Psychoanalytic and Cognitive Therapies, the IMPACT Study, will determine whether both medium intensity Cognitive Behavioural Therapy (CBT) [up to 20 sessions] and high intensity short term psychoanalytic therapy (STPP) [up to 28 sessions] are superior in reducing relapse compared to low intensity specialist clinical care (SCC) that is primarily advice and support [up to 12 sessions] in adolescents with moderate to severe depression attending routine child or adolescent mental health clinics. An additional aim is to establish the cost-effectiveness of CBT and STPP compared to SCC.

2. Study Design

Originally it was proposed that the trial will run in six CAMH clinics in each of three centres, giving 18 clinics with a minimum of one therapist for each treatment modality in each clinic and ten patients per treatment modality recruited in each clinic. This gives a total sample size of 540.

The ADAPT trial gave an SD of 14·6 at 28 weeks follow-up and correlation between baseline and follow-up of 0·41 for MFQ, proposed primary outcome of this study. We have assumed five points on the MFQ to be the minimum clinically important difference. This is approximately 25% of the change in the MFQ scale from baseline to 28 weeks. It is equivalent to a one point improvement on five of the 34 items of the scale. It is a standardize effect size of 0·34 (small to medium) and corresponds to non-overlap between treatments of approximately 25% (Cohen, 1988). Table 1 below gives estimates of power for Superiority, Non-Inferiority, and Equivalence designs for an intra-therapist correlation coefficient of 0·0, 0·025, or 0·05. Provided that the intra-cluster correlation is less than 0·025 a superiority analysis comparing CBT with STPP will have a power of over 80%. By virtue of the increased sample size specialist comparisons of the specialist treatments (CBT and STPP) with treatment as usual (SCC) will have substantial power. These power calculations assume a cross-sectional analysis, but statistical analysis will be based on longitudinal data using a linear mixed effects model (LME, see Section 8·2). Use of such a model will increase the power of the statistical analysis as data is in effect shared across follow-up time-points. This power calculation assumed a 90% follow-up as 92% follow-up at 28 weeks was achieved in ADAPT.

Allocation to treatment group was by minimisation controlling for severity (defined by MFQ score), sex, age, and recruiting region.

2·1 Statistical Hypotheses

The study was designed with a two level hypothesis: i) Both CBT and STPP will show superiority effects compared to BPI in the primary outcomes at 52 and 86 weeks; ii) CBT will show non-inferiority effects to STPP at 52 weeks; iii) STPP will show superiority effects compared to CBT at 86 weeks.

Table 1. Power assuming 18 therapists for each treatment modality, and ten patients per therapist

Intra-therapist correlation	Superiority	Inferiority	Equivalence
CBT vs STPP			
0	88%	93%	87%
0·025	80%	88%	75%
0·05	73%	82%	64%
(CBT+STPP) vs BPI			
0	96%	98%	96%
0·025	91%	95%	90%
0·05	85%	91%	82%

3. Outcome Measures

3·1 Primary outcome measure
Depressive symptoms over 36 to 86 weeks measured by the adolescent self report (Mood and Feelings Questionnaire, MFQ).

3·2 Secondary outcome measures

Along with the primary outcome, the secondary outcome measures are shown in the last column of Table 2 by frequency of collection and type of report. All are assumed to be continuous variables.

3·3 Hierarchy of young person versus parent reporting on various scales

When a young person and the parent both complete a particular questionnaire then the young person's data will form the basis of the main inference. The results from the parent will be supplementary.

4. Data analyses

Data analyses will be carried out by a statistician based in Biostatistics, Institute of Population Health, Manchester University, under the supervision of the trial statistician (CR) in conjunction with the IMPACT trial coordinator and trial centre in Cambridge. Economic data analyses will be carried out by a health economist based in the Centre for the Economics of Mental and Physical Health, Institute of Psychiatry, Kings College London, under the supervision of the trial economist (SB).

5. Data Sources

5·1 Pre randomisation data
Data are required for completion of the CONSORT diagram pre randomisation. These include:

- Numbers of potential participants assessed
- Numbers excluded after initial assessment by reason
- Numbers invited to baseline research interview
- Numbers excluded after baseline research interview by reason
- Numbers consenting and randomised by treatment arm.

5·2 Demographic data and patient characteristics prior to randomisation of randomised patients

The information collected at baseline consisted of basic demographic data (gender, age, twin, adopted/fostered, complications with pregnancy/labour/delivery, ethnicity, living arrangements, education, family employment) and clinical data (current medical problems, current medication of subject and other family member currently or in the past suffering any medical, emotional, or behavioural problems), plus the standard schedules as shown in Table 2.

5·3 Therapist/Care Provider and Treatment data

As suggested in the CONSORT guidance extension for trials of non-pharmacological interventions, information will be gathered regarding the characteristics of all therapists and care providers for each intervention. A statistical summary of this data will be prepared.

For each patient the intended therapist assigned at randomisation will be recorded along with the number of sessions attended.

The following minimum data will be collected for each trial therapy session (SCC, CBT, STPP):

- Therapist id
- Type of therapy delivered
- Individuals present in session

This will be aggregated to determine the number and type of therapy sessions received from each therapist.

A Kaplan-Meier plot and the associated log-rank test will be presented for time from randomisation to start of trial therapy and time to completion of trial therapy by a) treatment arm and b) by regional research centre (EA, NL, NW).

Summary statistics will be provided for duration of therapy by treatment arm.

5·4 Follow-up Assessments

The follow-up schedule is shown in Table 2.

6 Handling Missing Data and Slotting of Assessments

6·1 Item Non-response in Scale Measures

For questionnaire instruments, item non-response will be dealt with using a pro-rating strategy. Provided that at least 50% of items are available the observed total (for the completed items) and the number of items completed will be used to calculate an adjusted total as follows:

Adjusted total = Observed total * Total number of items in scale/Number of items completed

Note, this is equivalent to replacing the missing item by the average of available data for that dimension. The extent of pro-rata estimation will be reported for each scale for each treatment arm.

The NEO-FFI with five subscales and DSC with two subscales (global and affective) will have each subscale pro-rated and analysed separately.

6·2 Missing baseline covariate data

Subjects will not be excluded from outcome analyses due to missing baseline data. Where baseline covariate data (current not lifetime) cannot be obtained across different questionnaires, simple imputation (White & Thompson, 2005) which is based on multiple regression will be used. The following covariates will be used (see Section 8·2): region, comorbid behaviour disorders (CD+ODD), all anxiety disorders combined, SSRI use at baseline, age at randomisation, and sex. In addition baseline severity (MFQ score) will also be used. Substitution or imputation will not be used for post-baseline outcomes (see Section 8·3 for reasons).

6·3 Slotting of assessment measures

In RCT studies there is often a delay in starting the therapy post randomisation and the visits (research assessments) are scheduled relative to the start of therapy rather than from time since randomisation. To provide summary statistics we need to assign each actual assessment to a target assessment week based on pre-defined intervals from time since randomisation. The assessment will be assigned to one of the following scheduled visit weeks based on the interval it falls into for the time (in weeks) from randomisation:

Scheduled Visit	Interval in Weeks since randomisation
0	<= rand date
6	>0 - 11
12	12 - 25
36	26 - 46
52	47 - 64
86	65+

Note, these bands may have to be modified when the data is inspected. To avoid bias this will be carried out blind to outcome scores by calculating summary statistics on the completeness of the primary outcome and will take account of the results of analyses on time to start and completion of trial therapies.

If, for a given assessment window, there is more than one measurement in the band then the measurement nearest to the week from randomisation will be used for descriptive statistics.

7 Descriptive Analyses of randomised patients

7·1 Baseline Characteristics

Patients in the three treatment groups (SCC, CBT, STPP) will be described separately with respect to the characteristics given in Section 5·2.

Numbers (with percentages) for binary and categorical variables, and ordered categories plus means, standard deviation, median plus minimum and maximum values for continuous variables will be presented. Consistent with CONSORT guidance there will be no tests of statistical significance or confidence intervals for differences between randomised groups for any baseline variable.

All baseline measurement scales will be summarised separately for adolescent and parent responses, by treatment arm.

7·2 Follow-up

All measurement scales in Table 2 taken during follow-up will be summarised separately for adolescent and parents, by visit and each treatment arm. Note, the assignment of data to a specific assessment visit will use slotting as described in Section 6·3.

7·3 Missing follow-up data

For the primary outcome measure (MFQ) the frequencies (with percentages) of patient losses to follow-up at 6, 12, 36, 52, and 86 weeks after randomisation will be reported and compared between arms. For each subject, the provision of a measurement at each time point will be based on the slotting procedure given in Section 6·3.

Treatment arm and selected baseline characteristics (see Section 5·2) of subjects providing an adolescent outcome measure at the week 6 visit and those with missing data will be compared using a logistic regression model. Similarly, separate logistic regression models will be used to investigate patterns of failure to provide outcome measures at the later follow-up weeks.
These analyses will be used to develop an understanding of the missing data mechanism. These models will be repeated for the primary parent MFQ measure. The reasons for end of treatment and study discontinuation will be tabulated by treatment arm.

Table 2 Summary of Baseline and Follow-up Assessments
(^Mandatory measures reported : others are for additional studies)

	Weeks						Type
	0	6	12	36	52	86	
Adolescent Self report							
^MFQ: Mood and Feelings Questionnaire	✓	✓	✓	✓	✓	✓	P
^RCMAS: Revised Children's Manifest Anxiety Scale	✓	✓	✓	✓	✓	✓	S
^LOI: Leyton Obsessional Inventory	✓	✓	✓	✓	✓	✓	S
^Behaviours Checklist	✓	✓	✓	✓	✓	✓	S
RSES Rosenberg's Self Esteem Scale	✓	✓	✓	✓	✓	✓	
DEQ: Depressive Experiences Questionnaire ***(two subscales)***	✓				✓	✓	
DES-IV: Differential Emotion Scale-IV	✓		✓		✓		
DSC: Depressed States Checklist ***(two subscales)***	✓		✓		✓		
NEO-FFI: NEO-Five Factor Inventory ***(five subscales)***	✓				✓	✓	
DEEP	✓						
^RTSHIA: Risk-Taking & Self-Harming Inventory for Adolescents ***(two subscales)***	✓	✓	✓	✓	✓	✓	
RRS: Ruminative Responses Scale	✓	✓	✓	✓	✓	✓	
^EQ-5D: EuroQol measure of Health Related Quality of Life	✓	✓	✓	✓	✓	✓	
Interviewer Completed Measures							
^K-SADS-PL: Kiddie Schedule for Affective Disorders and Schizophrenia Present and Lifetime	✓	✓	✓	✓	✓	✓	
ZAN:BPD: Zanerini Rating Scale for Borderline Personality Disorder	✓				✓		
C-SSRS: Classification Suicide Severity Rating Scale	✓	✓	✓	✓	✓	✓	
CA-SUS: Child and Adolescent Service Use Schedule	✓	✓	✓	✓	✓	✓	
^HoNOSCA: Health of the Nation Outcome Scale for Children and Adolescents	✓	✓	✓	✓	✓	✓	S
EECI*: Expectations and Experience of Therapy Interview	✓			✓		✓	
Clinician Completed							
CGI: Clinical Global Impressions Scale	✓	✓	✓	✓	✓	✓	
WAI-S[×]: Working Alliance Inventory-Short		✓	✓	✓			
Family							
FAD: Family Assessment Device	✓		✓		✓		
APQ: Alabama Parenting Questionnaire	✓		✓		✓		
Life Events Questionnaire	✓		✓		✓		
Friendships Questionnaire	✓		✓		✓		
Parent self report							
SCL:90: Symptom Checklist 90 (global severity index)	✓	✓	✓	✓	✓		

Week 0 refers to baseline i.e. prior to randomisation

Type: P=Primary; S=Secondary outcome measure

Due to rationalisation of some of the scales there may be insufficient data for formal analyses. However, summary statistics will be provided for each mandatory scale.
**Analysed as part of IMPACT-ME substudy (at weeks 36 and 86 follow-up only for London participants)*
[×]Completed by adolescent and parent at same time points.

7·4 Quality Control of Measures

Observer Reliability between and within research sites

Intra and inter observer reliability will be considered using graphical methods and relevant summary statistics including intra-class correlation coefficients and kappa coefficients.

8 Statistical analysis of outcome comparing treatments

Extensive data cleaning of outcome and baseline data will be conducted without the treatment group allocations attached to the dataset. Results of these preliminary analyses will be reviewed by the trial research team to identify data errors and carry out preliminary checks regarding distributional assumptions prior to linking the treatment allocation to the follow-up data.

The analyses comparing treatments will be conducted applying the principle of intention to treat (ITT). No interim analyses of outcome data will be carried out unless specifically requested by the trial data monitoring and ethics committee.

8·1 Statistical inference between treatments

Within the protocol we considered both superiority and non-inferiority as potentially relevant hypotheses. The following hypotheses are stated in the protocol:

i) CBT will show superiority effects compared to SCC in the primary outcomes at 52 and 86 weeks
ii) STPP will show superiority effects compared to SCC in the primary outcomes at 52 and 86 weeks
iii) CBT will show non inferiority effects to STPP at 52 weeks
iv) STPP will show superiority effects compared to CBT at 86 weeks.

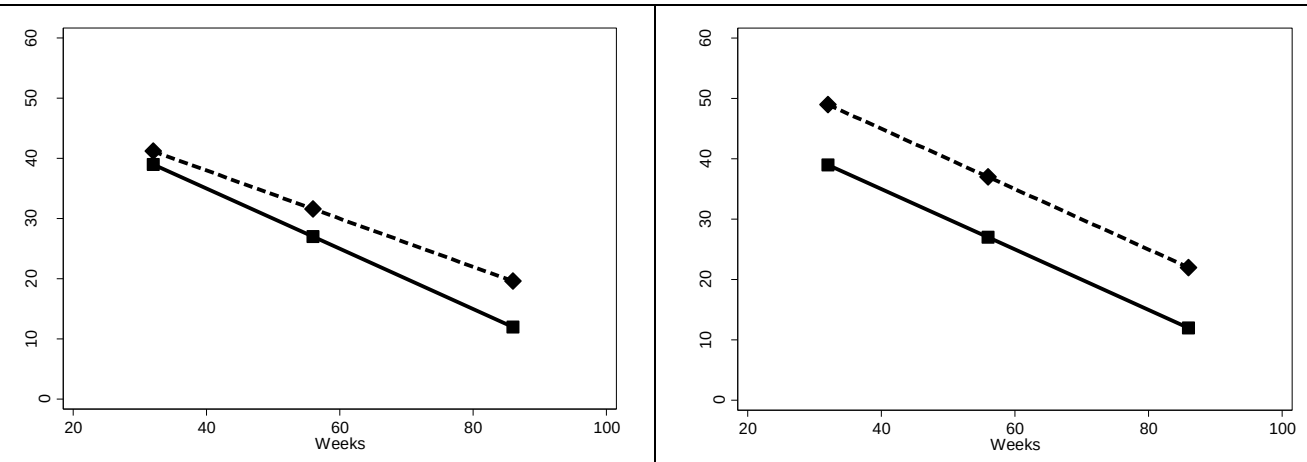
The hypotheses will be addressed using a linear mixed models analysis.

8·2 Treatment Effect Estimation for the Primary Outcome and other Continuous Outcome Measures

The intervention may influence outcome in two ways. Firstly, there may be a faster rate of recovery by 36 weeks and/or reduced clinically meaningful symptom recurrence between weeks 36 to 86 in one group than the other. Differential changes in symptoms over time can be estimated using a time with intervention group interaction. Secondly, there may be a systematic difference between intervention groups during follow-up, which is measured by a main effect. The statistical analysis of the primary outcome measure (MFQ) and the secondary continuous measures (see Table 2) will estimate the treatment effect using linear mixed effects models (LME, also know as random effects or random coefficient models). For all models, time (from randomisation considered as a continuous variable) will be centred based on the available data for the particular analysis being undertaken.

Because the aim of this study is to establish the longer term benefits of therapy we will consider only data over the post-treatment period for the primary analyses. All measures from week 36 onwards will be used for the statistical analyses for this purpose. By using data from week 36 onwards this should yield up to three measures per subject. This and the fact that time is continuous rather than discrete will reduce the potential for model identifiability problems given the number of random effects, time points and interaction terms.

Models with and without a time with treatment interaction



Each model will adjust for baseline values of the outcome under consideration and the pre-specified prognostic variables as shown in Table 3:

Table 3 Fixed Baseline Covariates for the Each Outcome Measure

Primary: MFQ	SR baseline MFQ, plus RCMAS, LOI, and BC scores at baseline Other baseline*: treatment arm, region, sex, age at randomisation in years, and use of SSRI at baseline
Secondary: RCMAS	SR baseline RCMAS, plus MFQ, LOI, and BC scores at baseline Other baseline: see MFQ outcome
Leyton Obsessional Inventory	SR baseline LOI, plus MFQ, RCMAS, and BC scores at baseline Other baseline: see MFQ outcome
Behaviours Checklist	SR baseline BC, plus MFQ, RCMAS, and LOI scores at baseline Other baseline: see MFQ outcome
HoNOSCA	Co-morbid behaviour disorders (i.e., a diagnosis of oppositional defiant disorder or conduct disorder) and all anxiety disorders combined+ Other baseline: see MFQ outcome

SR= Self-report

*****These covariates will be used in all lme models

+ For these two disorders a binary variable will be created for absent (coded as 0) versus a diagnosis of "Yes" or a "high clinical index" (coded as 1).

Models will also include a subject level random intercept and correlated random coefficient for time. In addition, therapist will be included as a random effect subject to model fitting constraints.

First, a model with a time with intervention group interaction will be fitted. If there is a significant treatment by time interaction, inference for the interaction will be reported and separate adjusted treatment effects for the three pairs of treatments will be estimated for 52 and 86 weeks from the model. The hypothesis of non-inferiority of CBT relative to STPP at 52 weeks will be addressed by considering the 95% confidence interval of the treatment effect.

If the interaction between time and treatments is not significant, this term will be omitted from the model. Adjusted treatment effects will be estimated and tested using this simplified model. Non-inferiority will be considered using the 95% confidence interval of the treatment effect.

To assess the treatment effect while receiving therapy, random intercept LME models will be fitted to data prior to week 36 post randomisation, made up mostly of notional week 6 and 12 assessment data.

8·3 LME Inference and missing data

Of note, by using maximum likelihood for these models, "Missing At Random" is assumed for drop-out i.e., missing outcome data is conditional on observed data. Under this assumption it is assumed that future behaviour, given the past, is the same for all, whether a subject drops out or not. This allows distributional information to be "borrowed" from those who remain on the trial and applied to those who drop-out given they have the same covariate set up until the time of dropout. Therefore, the estimand of the treatment effect is what would be seen if all subjects had remained on the study until the end.

8·4 LME Model Diagnostics

Normal probability plots will be used to check distributional assumptions of the model for residuals of within and between subject variance terms. Where there is evidence of non-normality outcome data may be transformed.

8·5 Longitudinal Models for the parent MFQ and other Continuous Outcome Measures

The analysis of the parent data and the secondary outcome measures will be essentially the same as the primary analyses of the adolescent data.

8·6 Models for binary and ordinal outcome data

Binary data will be analysed using longitudinal logistic regression and ordered categorical secondary outcome measures such as the CGI scale will be analysed using an ordinal logistic regression model with random intercept and gradient terms on the log-odds scale.

9 Further Analyses

9·1 Adherence to therapy

Summary statistics on the number of trial therapy sessions attended by each subject will be tabulated by arm. A frequency distribution of number of sessions will also be presented by arm. In addition, the percentage of target total sessions will be summarised by arm.

Based on input from specialists, a binary variable for adherence is defined as follows for the three modalities:

STPP: Eight sessions is considered as the minimum therapeutic dose: thus adherence = 0 when seven or fewer sessions in total are completed, otherwise adherence = 1.

CBT: Six sessions is considered as the minimum therapeutic dose: thus adherence = 0 when five or fewer sessions are completed, otherwise adherence = 1.

BPI: Three sessions is considered as the minimum therapeutic dose: thus adherence = 0 when two or fewer sessions are completed, otherwise adherence = 1.

Adherence will be summarised by arm and this may be used in a secondary causal analysis of treatment effects which will be investigated separately from the main statistical analysis following a proposed discussion of causal pathways. Analyses will estimate the propensity to receive/adhere to treatment, accounting for SSRI usage at baseline as a dichotomous variable.

9·2 Moderator Analyses

The following are the pre-specified moderators of treatment which will be investigated, one at a time, to determine whether they interact statistically with therapy group based on MFQ outcome data over the short term (i.e. >0 and < 36 weeks post randomisation) analysed using linear random intercept models:

- MFQ score at baseline
- Age at randomisation
- Sex
- Region
- SSRI prescribing at baseline

Based on the MFQ outcome data from week 36 onwards the same treatment interactions will be tested using random effects (intercept and slope) models.

In addition, an interaction between MFQ score and SSRI usage at baseline will be examined.

10 Economic evaluation

10·1 Perspective

In the first instance the economic evaluation will take a service perspective, which will include the use of all hospital, community health, costs in addition to mainstream education and social services. Secondly, we will undertake analyses from a societal perspective, which in addition to the service costs will include the out of pocket costs of travel to treatment that fall to carers and any productivity losses for the study participant or their carer as a result of illness.

10·2 Calculation of total costs

For each piece of service use information collected in the CA-SUS, a unit cost (for example a cost per hour with a professional, a cost per inpatient night, a cost per unit of a drug) will be applied and the total costs calculated. The total cost per participant is calculated by summing all costs. All unit costs will be for the financial year 2012-2013. Costs between 52 and 86 weeks will be discounted at a rate of 3·5% because cost-effectiveness results should reflect the present value of costs and benefits, as recommended by the National Institute for Health and Clinical Excellence (NICE) 2013. Sensitivity analysis using rates of 1·5% will also be presented in additional analyses.

All NHS hospital contacts will be costed using NHS reference costs (Department of Health 2011). Unit costs of community health and social services will be taken from national publications (Curtis 2011) and education costs from government published statistics (https://www.gov.uk/government/statistics/la-and-school-expenditure-financial-year-2012-to-2013). Medications will be costed using information in the British National Formulary (British Medical Association and Royal Pharmaceutical Society of Great Britain 2010). Contacts with criminal justice sector services using available data from published sources (e.g. HM Prison Service 2009). Where necessary, costs will be inflated to 2012-2013 rates using the Hospital and Community Health Services inflation indices or the Retail Price inflation indices, as appropriate (Curtis 2011).

The cost of the CBT, STPP, and SCC interventions will be calculated on the basis of the salary of the therapist plus overhead expenses (administrative, managerial, and capital). Calculation of the indirect time, including preparation and supervision of therapists, will be based on information provided by the trial therapists on the ratio of direct face-to-face contact compared with other intervention-related activities using the bottom-up approach (Drummond et al. 2005) used in similar research (Byford et al. 2007) to generate a cost per hour with each study therapist and clinician. Sensitivity analyses will vary the assumptions used in generating the intervention unit costs to investigate the impact of low and high cost estimates on the results of the study.

Productivity losses will be calculated for the adolescent (if they are in employment) and the parent or carer using the human capital approach, which involves multiplying the individual's salary by reported days off work due to illness.

10·3 Calculation of QALYs

QALYs will be calculated on the basis of the EuroQol EQ-5D health state classification instrument which has five domains: mobility; self-care; usual activities; pain/discomfort; and anxiety/depression. For each domain the respondent chooses one of three levels of functioning, good to poor. The three levels for each of the five domains are used to define 243 health states (Glick et al. 2007). The health states will then be given a utility score using responses from a representative sample of adults in the UK (Dolan et al. 1995). QALYs in the second year will be discounted at a rate of 3·5%, as recommended by NICE (National Institute for Health and Clinical Excellence 2013). QALYs will be calculated as the area under the curve as defined by the utility values at baseline, six, 12, 36, 52, and 86 weeks follow-up and it will be assumed that changes in utility score over time will follow a linear path (Richardson and Manca 2004).

10·4 Service use

Differences in the use of services between randomised groups will be compared descriptively. No statistical comparisons will be made.

10·5 Costs

Total cost per participant over follow-up will be calculated and analysed for both a service and a societal perspective. Although costs are not expected to be normally distributed, analysis will compare mean costs using standard t-tests/analysis of co-variance with covariates as described in Section 8·2. The robustness of the parametric tests will be confirmed using non-parametric, bias-corrected bootstrapping (Barber & Thompson, 2000). The following comparisons will be made:

1. CBT v BPI at 52 and 86 weeks
2. STPP v BPI at 52 and 86 weeks
3. CBT v STPP v BPI at 52 and 86 weeks

10·6 Cost-utility analysis

Cost-utility analysis will be undertaken using quality adjusted life years (QALYs) calculated from the EQ-5D as the measure of effect. Cost-utility will be assessed through the calculation of incremental cost-effectiveness ratios (ICERs) –the ratio of the additional cost of one intervention compared with another over the additional effects of one intervention over another. The following primary cost-utility analyses will be carried out using a service perspective:

1. CBT vs BPI at 86 weeks
2. STPP vs BPI at 86 weeks
3. A three-way analysis which will involve pair-wise comparisons between CBT, STPP, and BPI and a three-way comparison at the 86-week follow-up. When more than two strategies are compared, ICERs are calculated using rules of dominance and extended dominance (Johannesson & Weinstein, 1993). Strategies will be ranked by cost, from the least expensive to the most expensive, and if a strategy is more expensive and less effective than the previous strategy, it is said to be dominated and is excluded from the calculation of ICERs. This process compares strategies in terms of observed differences in costs and effects, regardless of the statistical significance of the difference.

In addition, a secondary analysis will make the same comparisons using a societal perspective and also using data from the 52 weeks follow-up.

Uncertainty around the costs and effectiveness estimates will be represented by cost-effectiveness acceptability curves, which will be calculated using the net benefit approach (Briggs, 2001). Net benefits for the sample using values for λ (willingness to pay for an additional QALY) ranging from £0 to a maximum value of £50,000 will be calculated. A bootstrap replication of 5000 means for each net benefit estimate will be created, adjusted for baseline covariates outlined in Section 8·2.

The proportion of these replications that are greater than zero will indicate the probability that the intervention is cost-effective for each value of λ. Plotting these probabilities on a graph creates a cost-effectiveness

acceptability curve, which depicts graphically the probability that the estimated cost-effectiveness ratio falls below the specified willingness to pay values (Van Hout et al., 1994).

10·7 Sensitivity analyses

A number of one-way sensitivity analyses will be undertaken to test the robustness of the results to the assumptions made in the economic evaluation. These will include, but will not be limited to:

- Variation of the cost of the interventions, dependent on seniority of therapists, time in direct contact with patients and other assumptions.
- Variation in the rate used for discounting of costs and outcomes in the second year to 1·5% as recommended by NICE (2013).

Potential Additional Analyses

Mediator Analyses

Analysis of treatment mediators will depend on there being evidence of a treatment effect and will therefore be part of a later exploratory analysis. The effect of treatment on the mediators will be investigated separately from the main statistical analysis following the proposed discussion of causal pathways. The proposed mediators are:

STPP:
STPP involves reflective and dynamic processes directly with the patient focussing on potential underlying unconscious abnormalities stemming from experience dependent learning. Parent support is a key element in this therapy and we hypothesise that improvements in parent well-being will mediate the efficacy of STPP by 86 weeks. This will be expressed as:
Lower Global Severity Index scores of the SCL-90 over the course of treatment will be associated with better response to STPP revealed as lower self reported depression scores by 86 weeks.

CBT:
CBT involves a very clear focus on current abnormalities and distortions of thinking processes and their ruminative style that serves to maintain and potentially amplify the pathological cognitive reasoning about the self, the future, and the world. We hypothesise that self-reported ruminations about negative cognitions will mediate the efficacy of CBT by 86 weeks. This will be expressed as:
Lower self reported total rumination score over the course of treatment will be associated with a better response to CBT revealed as lower self reported depression scores by 86 weeks.

BPI:
BPI is a pragmatic treatment involving here and now advice and support to aid understanding of illness and remedy clear-cut maladaptive behaviours in the environment such as social withdrawal and solitariness. We hypothesise that reducing solitariness and increasing behavioural sociability through a focus on well-being will mediate the efficacy of SCC by 86 weeks. This will be expressed as:
High friendship scores over the course of the treatment will be associated with a better response to SCC revealed as lower self-report depression scores by 86 weeks.

References

Barber, J. A. & Thompson, S. G. (2000). Analysis of cost data in randomised trials: an application of the non-parametric bootstrap. Statistics in Medicine, vol. 19, pp. 3219-3236.

Briggs, A. (2001). A Bayesian approach to stochastic cost-effectiveness analysis. International Journal of Technology Assessment in Health Care, vol. 17, pp. 69-82.

British Medical Association & Royal Pharmaceutical Society of Great Britain (2010). BNF 59 BMJ Books/Pharmaceutical Press, London.

Byford, S., Barrett, B., Roberts, C., Wilkinson, P., Dubicka, B., Kelvin, R. G., White, L., Ford, C., Breen, S., & Goodyer, I. (2007). Cost-effectiveness of selective serotonin reuptake inhibitors and routine specialist care with and without cognitive behavioural therapy in adolescents with major depression. Br J Psychiatry, vol. 191, no. 6, pp. 521-527.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences*. Hillsdale, New Jersey.: Lawrence Erlbaum Associates, Inc.,.

Curtis, L. (2011). Unit costs of health and social care 2010 PSSRU, University of Kent, Canterbury.

Department for Education and Skills (2002). LEA and School Information Service, The Stationery Office, London, LEA Level information - Standard Report 06.

Department of Health (2011). Reference costs 2010, Department of Health, London.

Dolan, P., Gudex, C., Kind, P., & Williams, A. (1995). A social tariff for EuroQoL: results from a UK general population survey, University of York: Centre for Health Economics, York.

Drummond, M., Sculpher, M., Torrance, G. L., O'Brien, B. J., & Stoddart, G. L. (2005). Methods for the economic evaluation of health care programmes, Third Edition edn, Oxford University Press, Oxford.

Glick, H. A., Briggs, A., & Polsky, D. (2007). Economic evaluation in clinical trials Oxford University Press, Oxford.

HM Prison Service (2009). Prison Service Annual Report and Accounts 2008, TSO, London.

Johannesson, M. & Weinstein, M. C. (1993). On the decision rules of cost-effectiveness analysis. Journal of Health Economics, vol. 12, no. 4, pp. 459-467.

National Institute for Health and Clinical Excellence (2013). Guide to the methods of technology appraisal, NICE, London.

Richardson, G. & Manca, A. (2004). Calculation of quality adjusted life years in the published literature: a review of methodology and transparency. Health Economics, vol. 13, pp. 1203-1210.

Van Hout, B. A., Al, M. J., Gordon, G. S., & Rutten, F. H. (1994). Costs, effects and cost-effectiveness ratios alongside a clinical trial. Health Economics, vol. 3, pp. 309-319.

White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. Statistics in Medicine 2005; 24: 993-1007

Addendum to IMPACT SAP (31st March 2016)

Section 1 and throughout document Replace specialist clinical care (SCC) with the term brief psychological intervention (BPI).
Section 6.2 The SSRI covariate used for imputation and analyses was "SSRI prescribed before trial entry" (where if missing information then it was assumed not to be prescribed) and not "SSRI use at baseline".
Section 6.3 The time of assessment caused some duplication in the slotting procedure where if more than one measurement was assigned to a band then only the nearest to the week from randomisation was to be used for the summary statistics. As the summary statistics from this approach were difficult to interpret we adopted the standard approach of reporting by researcher assessment instead.
Section 7.2 Only the primary and secondary outcome measurement scales were summarised.
Behaviours checklist (BC) is also known as the Antisocial Behaviours questionnaire (ABQ). Since the distribution of ABQ total score was highly skewed with the standard deviation larger than the mean at many time-points and medians of zero at weeks 52 and 86 for each group we considered ABQ as a binary outcome coded as one if the ABQ score was ≥1.
Section 8.2 Because MFQ and RCMAS are correlated the latter was omitted where both were originally listed as baseline covariates in an outcome model (except for the RCMAS outcome where MFQ was dropped). The covariates in Table 3 of the SAP were replaced by the ones shown in Table 1 of this Addendum with ABQ entered on the 3 point scale.
The SAP states:
"First, a model with a time with intervention group interaction will be fitted. If there is a significant treatment by time interaction, inference for the interaction will be reported and separate adjusted treatment effects for the three pairs of treatments will be estimated for 52 and 86 weeks from the model. The hypothesis of non-inferiority of CBT relative to STPP at 52 weeks will be addressed by considering the 95% confidence interval of the treatment effect."

Based on these interaction models the marginal effect of treatment was estimated at 36, 52 weeks and 86 weeks post randomisation for the following two comparisons rather than three to match the protocol hypotheses:

(i) STPP against CBT and
(ii) (CBT and STPP) against BPI

Note, the sample size calculation used a significance level of 2.5% to allow for this multiplicity. A Bonferroni correction was not applied to the p-values, but it is suggested that readers use a 2.5% significance level to maintain the family-wise 5% significance level at a particular point of assessment.
Section 9.1 Adherence is now redefined as therapeutic dose. For STPP this was changed from ≥8 to ≥6 sessions based on consultation with experts in this therapy field.

Section 9.2 Following detailed discussion between the PI's the original list of moderators were replaced by the following as they were deemed relevant:
Hypotheses for the DEQ at baseline:
1) Elevated relatedness/dependent scores will be associated with a relatively better response in the STPP group compared to BPI or CBT groups.
2) Elevated self-critical/identity scores will be associated with a relatively better response in the CBT group compared to BPI or STPP groups.

Hypotheses for the RRS at baseline

1) Higher scores will show a better treatment response in the CBT compared to the BPI and STPP arms.

Additional Analyses not specified in the SAP
In order to gain a better understanding of patterns over time in diagnosis, medication prescription and adverse events the following were undertaken. The results are presented in the HTA report.

1. We investigated change over time using GEE longitudinal analyses for Unipolar major depressive disorder (MDD) and MFQ total score >25 outcomes. The analysis of ABQ was changed from an mixed model to GEE since the data was not normally distributed.

2. Summaries on SSRI prescription prior to trial entry and also during follow-up overall and also split by <36 weeks and ≥36 weeks post randomisation to match the two analyses time periods were provided.
3. Adverse event reporting.

Addendum to IMPACT SAP (8th April 2016)

Section 1 and throughout document Replace specialist clinical care (SCC) with brief psychological intervention (BPI).
Section 4 Replace 'Centre for the Economics of Mental and Physical Health, Institute of Psychiatry' with 'King's Health Economics, Institute of Psychiatry, Psychology and Neuroscience'.
Section 6.2 The SSRI covariate used for imputation and analyses was "SSRI prescribed before trial entry" (where if missing information then it was assumed not to be prescribed) and not "SSRI use at baseline".
Section 6.3 The time of assessment caused some duplication in the slotting procedure where if more than one measurement was assigned to a band then only the nearest to the week from randomisation was to be used for the summary statistics. As the summary statistics from this approach were difficult to interpret we adopted the standard approach of reporting by researcher assessment instead.
Section 7.2 Only the primary and secondary outcome measurement scales were summarised.
Behaviours checklist (BC) is also known as the Antisocial Behaviours questionnaire (ABQ). Since the distribution of ABQ total score was highly skewed with the standard deviation larger than the mean at many time-points and medians of zero at weeks 52 and 86 for each group we considered ABQ as a binary outcome coded as one if the ABQ score was ≥1.
Table 2 Replace 'EQ-5D: EuroQol measure of health-related quality of life' with 'EQ-5D-3L: EuroQol measure of health-related quality of life (three level version)'. This is for clarity given the development of the new EQ-5D-5L (five level version).
Section 8.2 Because MFQ and RCMAS are correlated the latter was omitted where both were originally listed as baseline covariates in an outcome model (except for the RCMAS outcome where MFQ was dropped). The covariates in Table 3 of the SAP were replaced by the ones shown in Table 1 of this Addendum with ABQ entered on the 3 point scale.
The SAP states:
"First, a model with a time with intervention group interaction will be fitted. If there is a significant treatment by time interaction, inference for the interaction will be reported and separate adjusted treatment effects for the three pairs of treatments will be estimated for 52 and 86 weeks from the model. The hypothesis of non-inferiority of CBT relative to STPP at 52 weeks will be addressed by considering the 95% confidence interval of the treatment effect."

Based on these interaction models the marginal effect of treatment was estimated at 36, 52 weeks and 86 weeks post randomisation for the following two comparisons rather than three to match the protocol hypotheses:

(iii) STPP against CBT and
(iv) (CBT and STPP) against BPI

Note, the sample size calculation used a significance level of 2.5% to allow for this multiplicity. A Bonferroni correction was not applied to the p-values, but it is suggested that readers use a 2.5% significance level to maintain the family-wise 5% significance level at a particular point of assessment.
Section 9.1 Adherence is now redefined as therapeutic dose. For STPP this was changed from ≥8 to ≥6 sessions based on consultation with experts in this therapy field.
Section 9.2 Following detailed discussion between the PI's the original list of moderators were replaced by the following as they were deemed relevant:
Hypotheses for the DEQ at baseline:
1) Elevated relatedness/dependent scores will be associated with a relatively better response in the STPP group compared to BPI or CBT groups.
2) Elevated self-critical/identity scores will be associated with a relatively better response in the CBT group compared to BPI or STPP groups.

Hypotheses for the RRS at baseline
1) Higher scores will show a better treatment response in the CBT compared to the BPI and STPP arms.
Section 10.1 Replace section with the following to match with the perspective as originally planned and described in the published protocol: 'The economic evaluation will take a societal perspective, including the use

of all health, social care, education and criminal justice sector resources plus family costs in the form of travel to trial intervention sessions and productivity losses of the primary carer resulting from their child's illness.'
Section 10.2 Remove 'Sensitivity analysis using rates of 1.5% will also be presented in additional analyses' and 'Sensitivity analyses will vary the assumptions used in generating the intervention unit costs to investigate the impact of low and high cost estimates on the results of the study'. Sensitivity analyses focus on key areas of uncertainty in an economic evaluation, which cannot always be predicted in advance. Actual sensitivity analyses undertaken are described in full in the 'Additional analyses not specified in the SAP' below.
Section 10.5 Remove '52 weeks' from all comparisons in line with what was originally planned and described in the published protocol. Analyses at 52 weeks were discussed during the course of the trial, but final agreement was to stick to the original plan.
Section 10.6 A further comparison was added between CBT and STPP, in line with the clinical comparisons undertaken.
Section 10.6 Remove sentence 'In addition, a secondary analysis will make the same comparisons using a societal perspective and also using data from the 52 weeks follow-up.' This is no longer relevant given the removal of analyses at 52 weeks noted above.
Section 10.7 This section is removed in its entirety. As noted above, sensitivity analyses focus on key areas of uncertainty in an economic evaluation, which cannot always be predicted in advance. Actual sensitivity analyses undertaken are described in full in the 'Additional analyses not specified in the SAP' below.

Additional Analyses not specified in the SAP
In order to gain a better understanding of patterns over time in diagnosis, medication prescription and adverse events the following were undertaken. The results are presented in the HTA report.

4. We investigated change over time using GEE longitudinal analyses for Unipolar major depressive disorder (MDD) and MFQ total score >25 outcomes. The analysis of ABQ was changed from an mixed model to GEE since the data was not normally distributed.
5. Summaries on SSRI prescription prior to trial entry and also during follow-up overall and also split by <36 weeks and ≥36 weeks post randomisation to match the two analyses time periods were provided.
6. Adverse event reporting.

As noted above, sensitivity analyses undertaken on the economic data are hard to predict in advance as they are dependent on the assumptions made in the costing and analysis of the economic data. The following sensitivity analyses were carried out to test the robustness of the assumptions made:

1. The cost of sessions offered but not attended was explored by increasing the cost from the assumption of zero applied in the main analysis (which assumes professionals are able to make use of the time available to undertake alternative tasks) to 50% of the cost of a session (which assumes professionals make some use of the time available, but not all). Data were calculated as the number of sessions offered minus the number of sessions attended, which may not be exactly equivalent to the number of DNAs (did not attend) as sessions may have been offered but cancelled or rearranged. This analysis should therefore be interpreted with caution.
2. The impact of missing data was considered using multiple imputation of missing values.
3. Due to the variation in the timing of follow-up, cost per week was calculated and analysed, in addition to total cost over 86-weeks.

Appendix ii (continued)

Additional Results From the Analyses

Symptom Characteristics of young people entering the trial

Table A1 gives the prevalence of concurrent depressive symptoms from the K-SADS-PL. The most prevalent symptom was sleep disturbance (92%) followed by depressed mood (84%). The mean number of symptoms was 8·4 for the BPI group, 8·7 for CBT, and 8·3 for STPP. Recent suicide attempts refer current major depression episode. Lifetime suicide attempts refer to all lifetime except current episode.

Table A1: Depressive symptoms recorded at baseline research assessment

Depressive Symptom	BPI (n=155)		CBT (n=154)		STPP (n=156)		Total (N=465)	
	Freq.	(%)	Freq.	(%)	Freq.	(%)	Freq.	(%)
Two Weeks Prior to Baseline Assessment								
Sleep disturbance	141	(91·.0)	141	(91·.6)	145	(92·9)	427	(91·8)
Depressed Mood	131	(84·5)	134	(87·0)	125	(80·1)	390	(83·9)
Disturbed Concentration, inattention	112	(72·3)	119	(77·3)	118	(75·6)	349	(75·1)
Fatigue, lack energy	117	(75·5)	113	(73·4)	111	(71·2)	341	(73·3)
Worthlessness	108	(69·7)	101	(65·6)	105	(67·3)	314	(67·5)
Anhedonia, apathy	96	(61·9)	104	(67·5)	103	(66·0)	303	(65·2)
Irritable, anger	97	(62·6)	104	(67·5)	91	(58·3)	292	(62·8)
Suicidal Ideation	95	(61·3)	91	(59·1)	97	(62·2)	283	(60·9)
Decreased Appetite	71	(45·8)	78	(50·6)	71	(45·5)	220	(47·3)
Hopelessness	74	(47·7)	66	(42·9)	71	(45·5)	211	(45·4)
Indecision	47	(30·3)	62	(40·3)	51	(32·7)	160	(34·4)
Guilt	53	(34·2)	51	(33·1)	45	(28·8)	149	(32·0)
Agitation	43	(27·7)	53	(34·4)	50	(32·1)	146	(31·4)
Psychomotor retardation	37	(23·9)	38	(24·7)	36	(23·1)	111	(23·9)
Weight loss	29	(18·7)	25	(16·2)	23	(14·7)	77	(16·6)
Increased appetite	21	(13·5)	23	(14·9)	23	(14·7)	67	(14·4)
Weight gain	15	(9·7)	12	(7·8)	15	(9·6)	42	(9·0)
Hallucinations	12	(7·7)	16	(10·4)	6	(3·8)	34	(7·3)
Delusions	4	(2·6)	5	(3·2)	5	(3·2)	14	(3·0)
Recent Suicidal Attempt	3	(1·9)	2	(1·3)	7	(4·5)	12	(2·6)
Lifetime Suicidal Attempt	57	(36·8)	48	(31·2)	55	(35·3)	160	(34·4)

Table A2 gives a detailed breakdown of co-morbid psychiatric diagnoses recorded in the baseline K-SADS-PL by treatment group. A total of 225 (48%) were concurrently comorbid for at least one other psychiatric disorder. Of these 134 (29%) and 60 (13%) had one and two comorbidities, respectively. The maximum number of comorbidities was five in the BPI group and four in the other two groups. Overall, the most frequent comorbid diagnoses were generalised anxiety disorder and social phobia. There were no marked differences between the three treatment groups in these characteristics.

Table A2: Co-morbidity at baseline research assessment recorded by K-SADS-PL

Comorbid diagnosis	BPI (n=155)		CBT (n=154)		STPP (n=156)		Total (n=465)	
	Freq.	(%)	Freq.	(%)	Freq.	(%)	Freq.	(%)
Generalised anxiety disorder	34	(21·9)	34	(22·1)	31	(19·9)	99	(21·3)
Social Phobia	19	(12·3)	20	(13·0)	22	(14·1)	61	(13·1)
Oppositional defiant disorder	14	(9·0)	18	(11·7)	12	(7·7)	44	(9·5)
Specific phobia	16	(10·3)	13	(8·4)	6	(3·8)	35	(7·5)
Post-traumatic stress disorder	6	(3·9)	12	(7·8)	14	(9·0)	32	(6·9)
Separation anxiety disorder	6	(3·9)	9	(5·8)	5	(3·2)	20	(4·3)
Conduct disorder	7	(4·5)	2	(1·3)	5	(3·2)	14	(3·0)
Obsessive compulsive disorder	2	(1·3)	5	(3·2)	3	(1·9)	10	(2·2)
Panic without Agoraphobia	2	(1·3)	3	(1·9)	2	(1·3)	7	(1·5)
Agoraphobia	3	(1·9)	1	(0·6)	3	(1·9)	7	(1·5)
Alcohol abuse	2	(1·3)	2	(1·3)	2	(1·3)	6	(1·3)
Panic with Agoraphobia	2	(1·3)	1	(0·6)	2	(1·3)	5	(1·1)
Attention deficit hyperactivity disorder	2	(1·3)	1	(0·6)	1	(0·6)	4	(0·9)
Bulimia nervosa	0	(0)	1	(0·6)	2	(1·3)	3	(0·6)
Substance abuse	3	(1·9)	0	(0)	0	(0)	3	(0·6)
Anorexia nervosa	0	(0)	0	(0)	2	(1·3)	2	(0·4)
Substance dependence	1	(0·6)	0	(0)	1	(0·6)	2	(0·4)
Enuresis	1	(0·6)	0	(0)	1	(0·6)	2	(0·4)
Alcohol dependence	0	(0)	0	(0)	1	(0·6)	1	(0·2)
Encopresis	0	(0)	0	(0)	0	(0)	0	(0)

Table A3: Number (%) of clinical sessions attended by treatment allocation (The recommended number of treatment session were 12 for BPI, 20 for CBT, and 28 for STPP.)

Number of sessions	BPI Freq.	BPI (%)	BPI (%≥)	CBT Freq.	CBT (%)	CBT (%≥)	STPP Freq.	STPP %	STPP (%≥)
Missing	6	(3·9)	-	6	(3·9)	-	2	(1·3)	-
0	11	(7·1)	(100.0)	15	(9·7)	(100.0)	21	(13·5)	(100.0)
1	12	(7·7)	(92·6)	11	(7·1)	(89·9)	8	(5·1)	(86·4)
2	13	(8·4)	(84·6)	8	(5·2)	(82·4)	9	(5·8)	(81·2)
3	9	(5·8)	(75·8)	4	(2·6)	(77·0)	4	(2·6)	(75·3)
4	15	(9·7)	(69·8)	6	(3·9)	(74·3)	6	(3·8)	(72·7)
5	9	(5·8)	(59·7)	10	(6·5)	(70·3)	8	(5·1)	(68·8)
6	12	(7·7)	(53·7)	11	(7·1)	(63·5)	5	(3·2)	(63·6)
7	10	(6·5)	(45·6)	6	(3·9)	(56·1)	4	(2·6)	(60·4)
8	4	(2·6)	(38·9)	8	(5·2)	(52·0)	13	(8·3)	(57·8)
9	8	(5·2)	(36·2)	5	(3·2)	(46·6)	1	(0·6)	(49·4)
10	6	(3·9)	(30·9)	6	(3·9)	(43·2)	3	(1·9)	(48·7)
11	7	(4·5)	(26·8)	7	(4·5)	(39·2)	6	(3·8)	(46·8)
12	9	(5·8)	(22·1)	5	(3·2)	(34·5)	2	(1·3)	(42·9)
13	3	(1·9)	(16·1)	7	(4·5)	(31·1)	6	(3·8)	(41·6)
14	5	(3·2)	(14·1)	8	(5·2)	(26·4)	2	(1·3)	(37·7)
15	3	(1·9)	(10·7)	2	(1·3)	(20·9)	3	(1·9)	(36·4)
16	2	(1·3)	(8·7)	4	(2·6)	(19·6)	2	(1·3)	(34·4)
17	1	(0·6)	(7·4)	7	(4·5)	(16·9)	1	(0·6)	(33·1)
18	4	(2·6)	(6·7)	2	(1·3)	(12·2)	-	-	(32·5)
19	1	(0·6)	(4·0)	3	(1·9)	(10·8)	5	(3·2)	(32·5)
20	1	(0·6)	(3·4)	8	(5·2)	(8·8)	3	(1·9)	(29·2)
21	1	(0·6)	(2·7)	3	(1·9)	(3·4)	4	(2·6)	(27·3)
22	-	-	(2·0)	1	(0·6)	(1·4)	4	(2·6)	(24·7)
23	1	(0·6)	(2·0)	-	-	(0·7)	5	(3·2)	(22·1)
24	-	-	(1·3)	1	(0·6)	(0·7)	5	(3·2)	(18·8)
25	-	-	(1·3)	-	-	(0·0)	7	(4·5)	(15·6)
26	-	-	(1·3)	-	-	(0·0)	4	(2·6)	(11·0)
27	-	-	(1·3)	-	-	(0·0)	4	(2·6)	(8·4)
28	-	-	(1·3)	-	-	(0·0)	6	(3·8)	(5·8)
29	-	-	(1·3)	-	-	(0·0)	1	(0·6)	(1·9)
33	1	(0·6)	(1·3)	-	-	(0·0)	-	-	(1·3)
39	-	-	(0·7)	-	-	(0·0)	1	(0·6)	(1·3)
42	-	-	(0·7)	-	-	(0·0)	1	(0·6)	(0·6)
43	1	(0·6)	(0·7)	-	-	(0·0)	-	-	(0·0)
Total	155	100		154	100		156	(100·0)	
Median (IQR) [a]	6	(4,11)		9	(5,14)		11	(5,23)	

[a] Calculated for young people receiving one or more session.

Table A4: Antidepressant (AD) prescribing during treatment and follow-up

<36 weeks	BPI (n=122)		CBT (n=120)		STTP (n=122)	
Medication	Freq.	%	Freq.	%	Freq.	%
Citalopram	3	2·5	5	4·2	3	2·5
Fluoxetine	29	23·8	27	22·5	23	18·9
Sertraline	3	2·5	3	2·5	9	7·4
Any AD	34	27·9	33	27·5	32	26·2
Not receiving medication	88	72·1	87	72·5	90	73·8
≥36 weeks	(n=125)		(n=125)		(n=124)	
Medication	Freq.	%	Freq.	%	Freq.	%
Citalopram	9	7·2	9	7·2	6	4·8
Fluoxetine	36	28·8	30	24·0	24	19·4
Sertraline	12	9·6	5	4·0	13	10·5
Any AD	50	40·0	43	34·4	43	34·7
Not receiving medication	75	60·0	82	65·6	81	65·3
All Follow-up	(n=137)		(n=137)		(n=137)	
Any AD	56	40·9	55	40·1	50	36·5
Not receiving medication	81	59·1	82	59·9	87	63·5

Table A5: Response rates and time from randomisation for the primary outcome (MFQ) by assessments

Assessment	BPI				CBT				STPP			
	response rate		Time since Randomisation		response rate		Time since randomisation		response rate		Time since randomisation	
	Freq.	(%)	mean	(Min ,Max)	Freq.	(%)	mean	(Min ,Max)	Freq.	(%)	mean	(Min ,Max)
Baseline	155	(100)			154	(100)			156	(100)		
6 week	99	(64)	11·0	(6,25)	104	(68)	12·3	(7,41)	107	(69)	11·1	(6,21)
12 week	112	(72)	17·6	(12,33)	106	(69)	19·0	(11,38)	108	(69)	17·6	(12,28)
36 week	105	(68)	42·3	(36,54)	104	(68)	42·9	(35,63)	109	(70)	41·5	(31,52)
52 week	105	(68)	59·2	(51,76)	111	(72)	60·3	(48,92)	110	(71)	59·3	(50,85)
86 week	116	(75)	95·4	(73,132)	123	(80)	94·9	(82,147)	114	(73)	95·1	(69,149)

Table A6: Linear Mixed Effects (LME) models estimates of main effects of treatment and time with treatment interactions with therapist, participant, and slope random effects for data from 36 weeks onwards post randomisation

Outcome measure	Treatment effect	(95% c.i.)	p-value[a]
Primary			
MFQ			
Time-treatment interaction			
STPP vs CBT	0·008	(-0·058 to 0·074)	0·812
CBT vs BPI	0·023	(-0·043 to 0·089)	
STPP vs BPI	0·031	(-0·035 to 0·097)	
(CBT+STPP) vs BPI	0·027	(-0·030 to 0·084)	0·361
Treatment main effect[c]			
STPP vs CBT	0·411	(-2·901 to 3·723)	0·808
CBT vs BPI	-2·591	(-5·860 to 0·678)	
STPP vs BPI	-2·179	(-5·487 to 1·128)	
(CBT+STPP) vs BPI	-2·385	(-5·226 to 0·456)	0·100
Secondary			
RCMAS			
Time-treat interaction			
STPP vs CBT	-0·012	(-0·732 to 0·049)	0·701
CBT vs BPI	0·069	(0·007 to 0·131)	
STPP vs BPI	0·057	(-0·005 to 0·120)	
(CBT+STPP) vs BPI	0·063	(0·009 to 0·117)	0·022
Treatment main effect [b]			
STPP vs CBT	0·488	(-2·450 to 3·425)	0·751
CBT vs BPI	-2·140	(-5·052 to 0·772)	
STPP vs BPI	-1·652	(-4·601 to 1·297)	
(CBT+STPP) vs BPI	-1·896	(-4·432 to 0·640)	0·116
HoNOSCA			
Time-treat interaction			
STPP vs CBT	0·0002	(-0·039 to 0·039)	0·993
CBT vs BPI	0·016	(-0·022 to 0·054)	
STPP vs BPI	0·016	(-0·023 to 0·055)	
(CBT+STPP) vs BPI	0·016	(-0·017 to 0·049)	0·348
Treatment main effect [b]			
STPP vs CBT	0·612	(-0·785 to 2·008)	0·391
CBT vs BPI	-1·055	(-2·414 to 0·303)	
STPP vs BPI	-0·444	(-1·820 to 0·932)	
(CBT+STPP) vs BPI	-0·749	(-1·925 to 0·426)	0·207

[a] p-value based on a likelihood ratio test where a significance level is 0·025 should be used to control for two comparisons.

[b] Treatment main effects are based on the time-treatment interaction model.

[c] Treatment main effect is averaged across centred time since randomisation because there is no interaction between time and treatment.

A total of 392 (392/465, 84%) participants were retained over the follow up period and use in the primary analyses (BPI,132 (85%) of 155, CBT,133 (86%) of 154 , STPP, 127 (81%) of 156) .

Appendix iii

Economic evaluation methods and results

Aim
The aim of the economic evaluation was to investigate the cost-effectiveness of psychological treatments for adolescent depression and in particular to determine whether the additional cost of the two specialist treatments, CBT and STPP, can be justified by improvements in effectiveness and/or decreased use of health and social care services compared to BPI by 86 weeks follow up.

Perspective
The a priori perspective of the economic evaluation was societal, including the use of all health, social care, education and criminal justice sector resources plus family costs in the form of travel to trial intervention sessions and productivity losses of the primary carer resulting from their child's illness. However, criminal justice, travel costs and productivity losses were not found to be relevant to this population, being very low, and were excluded from the analysis.

Method of economic evaluation
The primary economic analysis was a cost-effectiveness analysis with outcomes expressed as quality adjusted life years (QALYs), as recommended by NICE.[6]

Calculation of costs
The process of calculating costs was separated into the identification, measurement and valuation of relevant resources.

Identification of resources
Relevant resources were identified based on the results of previous studies in adolescent depression and in discussion with study clinicians and patient representatives.[7] Resource use was collected in the following domains:

Delivery of the BPI, CBT and STPP interventions
Use of NHS secondary care services

- Inpatient stays (mental health and all medical specialties)
- Outpatient appointments (mental health and all medical specialties)
- Accident and emergency attendances

Use of NHS primary care services

- General practitioner (in surgery, at home, and by telephone)
- Community nurse (e.g. practice nurse, district nurse, health visitor, midwife)
- Community paediatrician
- Community mental health service
- Community medical professional e.g. physiotherapist
- School based mental health and medical professionals

Use of medication in the following areas

- Antidepressants
- Sleeping tablets
- Mood stabilisers/antipsychotics

Use of social care and education sector services

- Foster care and residential care
- Staffed accommodation e.g. hostel
- Social worker
- Specialist education facilities
- Education psychologist
- Family support worker
- Youth worker
- Youth offending team worker

Measurement of resources

Trial interventions
The trial therapists recorded details of attendance and non-attendance at treatment sessions, and duration of treatment sessions for each study participant throughout the trial.

Other health, social care and education sector services
Data on use of all other services included in the study perspective were collected using the Child and Adolescent Service Use Schedule. The CA-SUS was developed using data from several child and adolescent mental health trials and was further modified and successfully employed in a previous trial in adolescent depression.[7] The CA-SUS was completed with participants and family members in interview with a researcher at baseline and at the 6, 12, 36, 52 and 86 week follow-up interviews. At baseline, information covered the previous three months. At each of the follow-up interviews, service use since the previous interview was recorded; in this way, the entire period from baseline to final follow-up was covered. The CA-SUS asks participants for the number and duration of contacts with various services and professionals.

Valuation of resources
To calculate the total cost of the resources used by each study participant, a unit cost was applied to each resource use item. All unit costs were for the financial year 2011/12, uprated, where necessary, using the Hospital and Community Health Services Index.[8] Costs in the second year were discounted at a rate of 3·5% as recommended by NICE.[6] All unit costs are summarised in Table A10.

Table A7: Unit costs applied to economic data

Service	Unit	Cost (£)
CBT	Per session	71-111
STPP	Per session	64-190
BPI	Per session	58-171
Medication	Per daily dose	various
Inpatient	Per night	495-632
Outpatient	Per appointment	30-624
Accident and Emergency	Per attendance	131-155
Ambulance	Per trip	230
GP surgery	Per minute of patient contact	3·40
GP home	Per home visit minute	4·30
GP telephone	Per minute of patient contact	3·38
Practice nurse	Per minute of face-to-face contact	0·88
District nurse, health visitor, midwife	Per home visit minute	1·03
CAMHS team	Per contact	225
Counsellor/therapist	Per minute of client contact	1·08
Social worker	Per minute	3·43
Support worker/ youth worker	Per minute	0·61

Education psychologist	Per minute	2·27
Physiotherapist	Per contact	80
Speech and language therapist	Per contact	88
Dietician	Per contact	71
Youth offending team worker	Per minute	3·43

Trial treatments

Treatment sessions were costed on the basis of the profession and grade of the therapist that delivered each session for each trial participant. The length of the treatment sessions was extracted from the average duration of treatment recorded in the session record forms. Average duration of sessions was 45 minutes for BPI, 50 minutes for STPP, and 55 minutes for CBT. For the base case analysis, only the costs of the sessions that the young person attended were included. This assumption was employed because of an understanding that clinicians are usually able to do something else during the time freed up by missed appointments. In a sensitivity analysis, an estimate of the cost of the sessions that were offered but not attended was included. The data for this analysis came from the records held by the trial therapists and are the closest data to non-attendance available. The rate of non-attended sessions was included at 50% of the cost of a full session, which assumes professionals make some use of the time available, but not all.

Antidepressants and other medication
The total cost of antidepressants prescribed and other included medication costs were calculated using daily dose information and costs of the generic drug as listed in the British National Formulary.[9]

Secondary care service
Unit costs for all hospital services were taken from the National Schedule of NHS Reference costs for 2011/12.[10]

Primary care services and social care and voluntary services
For NHS primary care services, social workers, and support workers costs contained in the Unit Costs of Health and Social Care and NHS Reference costs were used.[8,10]

Calculation of Quality Adjusted Life Years (QALYs)
QALYs were calculated using the area under the curve approach after the health states from the EQ-5D were converted into utility scores using responses from a representative sample of adults in the UK.[11] It was assumed that changes in utility score over time followed a linear path.[12] QALYs in the second year were discounted at a rate of 3·5% as recommended by NICE and all analyses were adjusted for baseline utility scores to take into consideration the impact any baseline differences will have on the area under the curve.[6,13]

Data analysis
For base case calculations, complete case analysis was used, with the impact of missing data explored in sensitivity analyses. All analyses were carried out on an intention to treat basis using STATA (www.stata.com).

Resource use
Resource use by the study participants is reported descriptively for each group at 86 weeks as mean use and percentage of the group who had at least one contact. No statistical comparisons between use of services are made to avoid problems associated with multiple testing, and because the focus of the economic evaluation is on cost and cost-effectiveness.

Difference in costs and QALYs
A number of tests for differences in costs at 86 weeks between randomised groups were completed:

1) CBT v BPI
2) STPP v BPI
3) CBT v STPP

These were analysed using linear regression models with the following pre-specified covariates: baseline costs (total cost over the previous three months), region (East Anglia, North London, North West), behavioural disorder at baseline (measured using the K-SADS-PL), and antidepressant use at baseline. The validity of the results were confirmed using bias-corrected, non-parametric bootstrapping (repeat re-sampling).[14] Despite the skewed nature of cost data, this approach is recommended to enable inferences to be made about the arithmetic mean.[15]

Cost-effectiveness analyses

For cost-effectiveness, analysis moves from considering differences in costs and outcomes in terms of statistical significance to analysing costs and outcomes together in a decision-making context[16]. The cost-effectiveness analysis, undertaken using QALYs calculated from the EQ-5D measure of health-related quality of life, was completed for the following comparisons:

1) CBT v BPI
2) STPP v BPI
3) CBT v STPP
4) CBT v STPP v BPI

Initially, incremental cost-effectiveness ratios (ICERs) were calculated, which are the difference in mean cost divided by the difference in mean effect.[17] Because ICERs are calculated from four sample means and are therefore subject to statistical uncertainty, 5000 re-samples (bootstrapping) from the cost and outcomes data were used to generate a distribution of mean costs and effects.[15] These distributions were plotted onto the cost-effectiveness plane for interpretation. Replications that fall in the South-West quadrant of the plane suggest that the intervention is less costly and less effective than the comparator, and those that fall in the South-East quadrant suggest that the intervention is less costly and more effective than the comparator. Replications in the North-West quadrant suggest the intervention is more costly and less effective than the comparator, while those in the North-East quadrant suggest the intervention is more costly and more effective than the comparator.

The bootstrapped distributions were also used to calculate the probability that each of the treatments is the optimal choice, subject to a range of possible maximum values (the ceiling ratio, λ) that a decision-maker might be willing to pay for a unit improvement in outcome. To explore the uncertainty that exists around estimates of mean costs and effects as a result of sampling variation and uncertainty regarding the maximum value of λ, cost-effectiveness acceptability curves (CEACs) are presented by plotting these probabilities for a range of possible values of the ceiling ratio (λ).[18] All analyses used baseline costs, region, and behavioural disorder at baseline as covariates.

Sensitivity analyses

A number of sensitivity analyses were carried out to test the robustness of the assumptions made:

1. The cost of sessions offered but not attended was explored by increasing the cost from the assumption of zero applied in the main analysis (which assumes professionals are able to make use of the time available to undertake alternative tasks) to 50% of the cost of a session (which assumes professionals make some use of the time available, but not all). Data were calculated as the number of sessions offered minus the number of sessions attended, which may not be exactly equivalent to the number of DNAs (did not attend) as sessions may have been offered but cancelled or rearranged. This analysis should therefore be interpreted with caution.
2. The impact of missing data was considered using multiple imputation of missing values.
3. Due to the variation in the timing of follow-up, cost per week was calculated and analysed.

Economic evaluation results

Data completeness

At 86 weeks, full CA-SUS service use data was available for 94 participants (61%) in the CBT group, 91 (58%) in the STPP group, and 92 (59%) in the BPI group, which was 60% of the total number randomised.

Outliers
The cost data were examined to consider the impact of highly influential observations, defined by Weichle et al [19] as those whose exclusion result in major changes in the results. Two observations were identified as above the 99th percentile for total costs, but only one of these would have increased parameter estimates by a factor of 1.4. Therefore this one observation was removed from the main analysis as recommended.[19]

Resource use
All resources used over the 86-week follow-up period are summarised by group in Table A11.

Trial treatment
For the sample of participants with full service use information, the average number of treatment sessions attended by the young people was 7·97 in the BPI group, 9·73 in the CBT group and 13·85 in the STPP group. The numbers differ slightly from those reported in the main paper because they are the results for the sub-sample of participants for whom we had full service use data. On average, the number of sessions attended was lower than the number of sessions planned (BPI 12 sessions, CBT 20 sessions, STPP 28 sessions).

Other health and social services
Overall there was little difference between randomised groups in levels of service use over the 86 week follow-up (see Table A11). Levels of mental health admissions were low (less than 2%) across all randomised groups. There were slight variations in non-mental health admissions, with 13% of the STPP group being admitted compared to 8% in the BPI group and 5% in the CBT group. Overall up to a fifth of participants had a non-mental health admission. Accident and emergency attendances were not uncommon (BPI 23·40%, CBT 12·63%, STPP 19·57%), but average levels of attendance were less than one contact in each group.

GPs were the most widely used service, accessed by 66%, 72%, and 64% of participants in the BPI, CBT, and STPP groups, respectively. Use of community mental health services, excluding the trial interventions, was highest in the BPI group (46% of BPI participants) compared to 38% and 29% of the CBT and STPP groups, respectively. Rates of social services contacts were also highest in the BPI group.

Antidepressant medication
Over the course of the study, patients were allowed to receive an SSRI in addition to psychological treatment if they met National Institute of Health and Care Excellence guidelines for combined treatment to aid clinical remission by end of treatment. The proportion of participants prescribed antidepressant medication at any point over the 86-week follow-up was around 30% in each group.

Table A8: Service use (unit), Mean, SD, over 86-week follow-up

	BPI (n=96)			CBT (n=95)			STPP (n=92)		
	Mean	SD	%	Mean	SD	%	Mean	SD	%
Treatment (sessions)	7·97	5·19	96.67	9·73	6·54	93.62	13·85	10·41	91.21
Mental health inpatient (night)	0·02	0·20	1·04	0·08	0·72	2·11	0·00	0·00	0.00
Non-Mental health inpatient (night)	0·26	1·06	8·34	0·11	0·57	5·26	0·42	1·54	13·04
Mental health outpatient (attendance)	0·01	0·10	1·04	0·05	0·51	1·05	0·00	0·00	0.00
Non-mental health outpatient (attendance)	0·65	1·83	18·75	0·35	1·19	13·68	0·75	1·90	23·91
Accident and emergency (attendance)	0·45	1·61	22·91	0·14	0·38	12·63	0·35	0·80	19·57
General practitioner (contact)	2·79	5·00	66·67	2·40	4·07	71·58	2·60	3·79	64·13
Community medical services (contact)	0·12	0·43	8·33	0·09	0·49	5·26	0·37	2·26	4·35
Community mental health services (contact)	4·93	11·12	45·83	5·64	14·08	37·89	3·80	10·85	29·35
Community social services (contact)	1·33	3·74	20·83	0·95	4·02	11·58	6·88	62·32	14·13
Education support services (contact)	1·32	5·18	25·00	1·61	6·90	15·79	3·11	11·00	27·17
Antidepressant medication (any prescribed)			30·77			28·57			31·73
Other medication (any prescribed)			2·13			4·21			4·34

Total cost

Treatment costs
On average the cost of the trial interventions was lowest for CBT (£904·57) and highest for STPP (£1396·72), with BPI costing £1292·91. These differences reflect variation in the number and duration of treatment sessions and the cost of the professionals providing the therapy.

Total costs over follow-up
The broadly similar levels of service use reported in Table A11 translated into similar total health, social care and education costs per participant over the 86 week follow-up across the three groups: £1368·04 in the BPI group, £1459·26 in the CBT group, and £1668·51 in the STPP group. Including the cost of the trial interventions generated total costs per participant over the 86 week follow-up of £2678·39 for BPI, £2379·01 for CBT, and £3081·70 for STPP (see Table A12).

The results of the between group comparisons, detailed in Table A13, show that there were no significant differences in costs between groups. Bootstrapped confidence intervals were similar to those calculated from the linear regression models so are not presented here.

Table A9: Total cost per participant (£), Mean, SD, over 86-week follow-up

	BPI (n=90)		CBT (n=92)		STPP (n=91)	
	Mean	SD	Mean	SD	Mean	SD
Health, social care and education costs	1385.48	2807.69	1474.43	3496.52	1684.98	3441.00
Treatment costs	1292·91	851.29	904·57	607·25	1396·72	1133·41
Total costs	2678·39	2678·39	2379·01	3643·85	3081·70	3573·17

Table A10: Between group differences in total costs over 86-week follow-up

	Coefficient	95% confidence interval	p-value
CBT versus BPI (n=180)	-338·54	(-1333·17 to 656·09)	0·503
STPP versus BPI (n=174)	609·55	(-406·73 to 1625·83)	0·238
CBT versus STPP (n=178)	-709·23	(-1836·04 to 417·58)	0·216

* Adjusted for region and baseline cost, behavioural disorder and antidepressant use

Outcomes

Health-related quality of life

EQ-5D scores at baseline and all follow-up points are detailed in Table A14. Utility scores were generally higher in the CBT group compared to BPI and STPP, where a higher score denotes higher levels of health-related quality of life. However, differences were small and at the 86 week follow-up, scores were marginally higher in the BPI group followed by the STPP group. The QALYs show very little between group differences: BPI group 1.241 QALYs, CBT group 1·228 QALYs, and STPP 1·246 QALYs. There were no significant between group differences in QALYs as shown in Table A15.

Table A11: EQ-5D scores and QALYs over 86-week follow-up*

Assessment point		BPI		CBT		STPP	
	n	Mean	SD	Mean	SD	Mean	SD
Baseline	447	0·596	0·275	0·578	0·281	0·569	0·258
t1 (week 6)	303	0·622	0·278	0·685	0·236	0·674	0·275
t2 (week 12)	310	0·713	0·236	0·714	0·267	0·680	0·259
t3 (week 36)	290	0·730	0·262	0·797	0·227	0·765	0·233
t4 (week 52)	295	0·771	0·227	0·803	0·232	0·792	0·257
t5 (week 86)	307	0·817	0·228	0·780	0·256	0·808	0·240
QALYs	294	1·241	0·270	1·228	0·304	1·246	0·293

*Higher EQ-5D scores and higher QALYs denote better quality of life

Table A12: Between group differences in QALYs over 86-week follow-up

	Coefficient	95% confidence interval	p-value
CBT versus BPI (n=195)	-0·009	(-0·091 to 0·074)	0·839
STPP versus BPI (n=193)	0·000	(-0·081 to 0·082)	0·992
CBT versus STPP (n=200)	-0·019	(-0·103 to 0·064)	0·648

*Adjusted for region and baseline EQ-5D score, behavioural disorder and antidepressant use

Cost-effectiveness analysis

CBT v BPI

For the CBT versus BPI comparison, CBT is less costly but slightly less effective in terms of QALYs than BPI. As a result, the replications produced in the scatterplot in Figure A1 are mainly in the South-West and South-East quadrants reflecting lower costs in the CBT group (points below the x-axis) and the very small difference in outcomes between the two groups (points evenly spread across the y-axis). The cost-effectiveness acceptability curve (CEAC) in Figure A2 shows that for all levels of willingness to pay per QALY there is a higher probability that CBT is more cost-effective than BPI.

Figure A1: Scatter plot of differences in costs versus differences in QALYs for CBT versus BPI

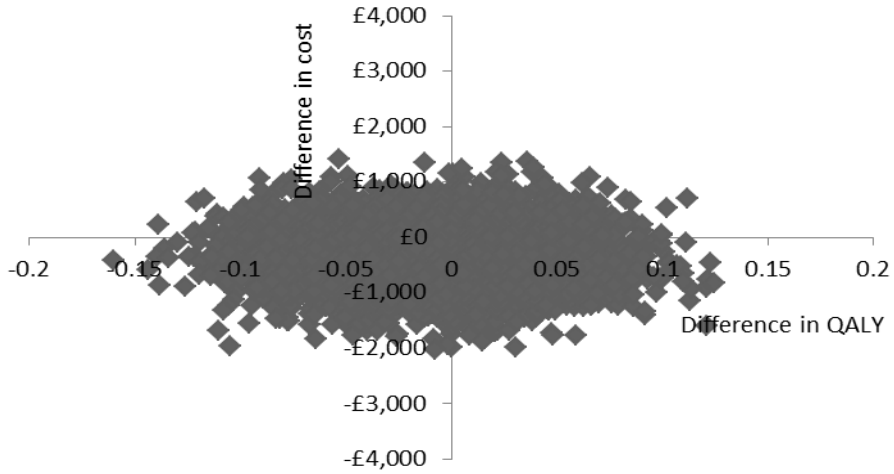
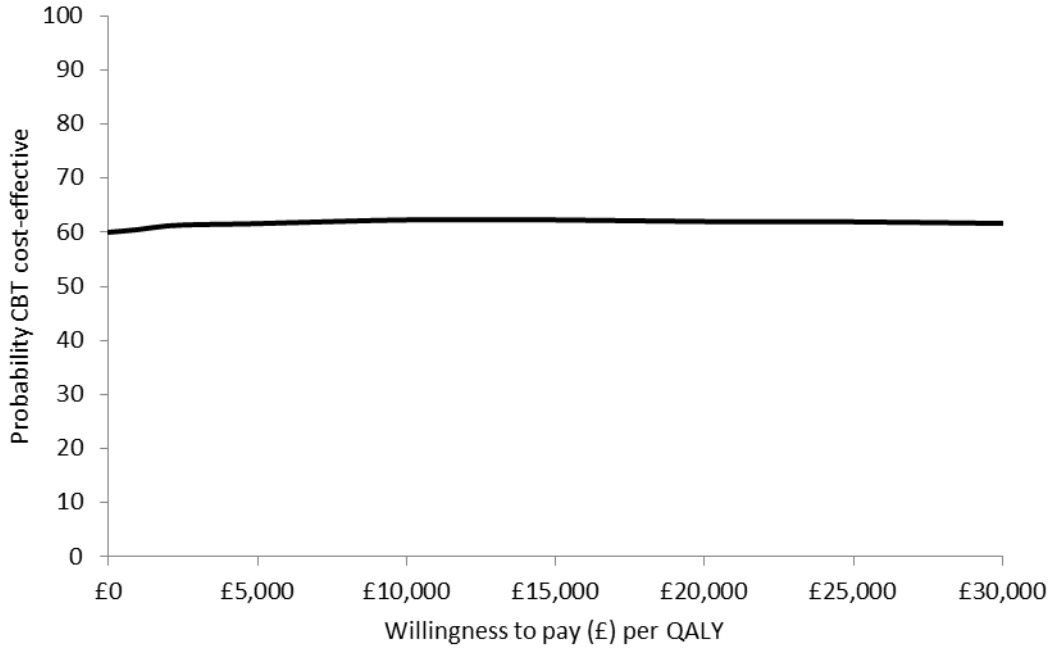


Figure A2: Cost-effectiveness acceptability curve showing the probability that CBT is cost-effective compared to BPI for different values a decision maker might be willing to pay for improvements in QALYs



STPP v BPI

For the STPP versus BPI comparison, costs were on average £403 more in the STPP group than the BPI group and QALYs were similar. The bootstrapped replications for STPP v BPI are shown in Figure A3. The majority are in the North-East and North-West quadrants, reflecting the higher costs in the STPP group (points above the x-axis). The CEAC in Figure A4 shows that there are no willingness to pay values where the probability of STPP being cost-effective compared to BPI is greater than 23%, within the £20,000-£30,000 ceiling level of willingness to pay considered acceptable by NICE [6].

Figure A3: Scatter plot of differences in costs versus differences in QALYs for STPP versus BPI

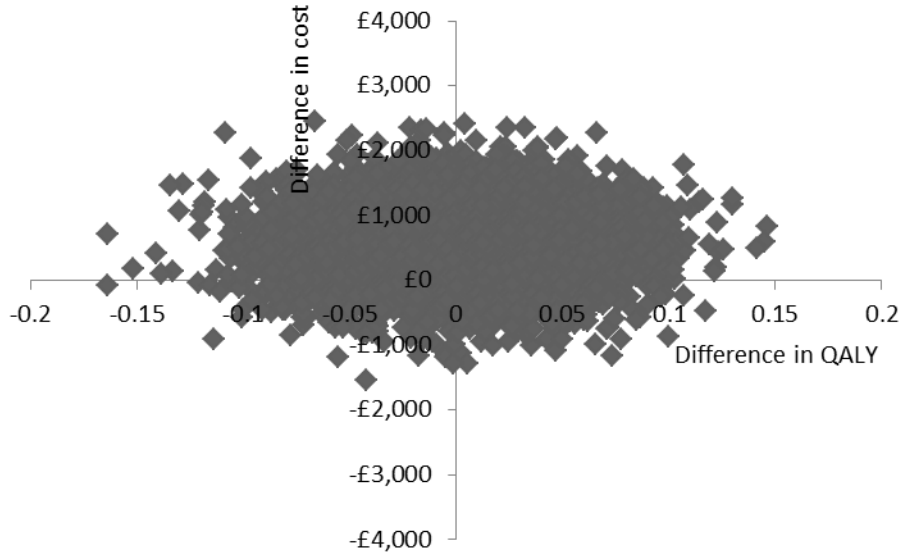
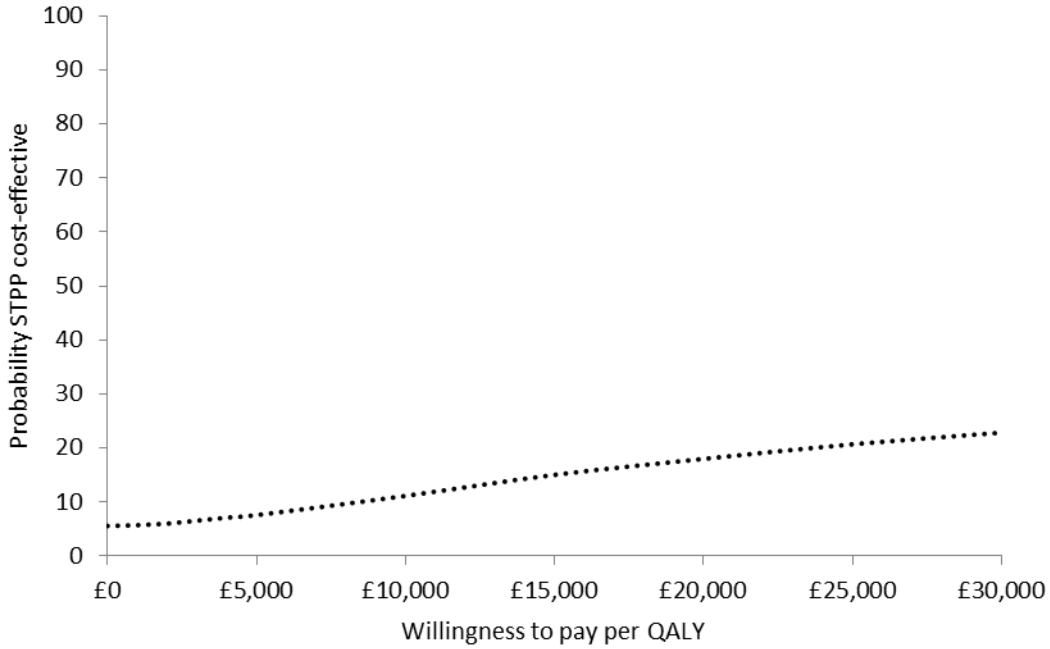


Figure A4: Cost-effectiveness acceptability curve showing the probability that STPP is cost-effective compared to BPI for different values a decision maker might be willing to pay for improvements in QALYs



CBT v STPP

Comparing the two intensive psychological treatments, CBT and STPP, total costs per participant over the 86 week follow-up were on average £703 lower in the CBT group and outcomes 0·02 QALYs worse. As a result, the replications in the scatterplot in Figure A5 are mostly in the South-West quadrant. The CEAC shown in Figure A6 suggests that the probability that CBT is cost-effective compared to STPP for all willingness to pay values is greater than 50%.

Figure A5: Scatter plot of differences in costs versus differences in QALYs for CBT versus STPP

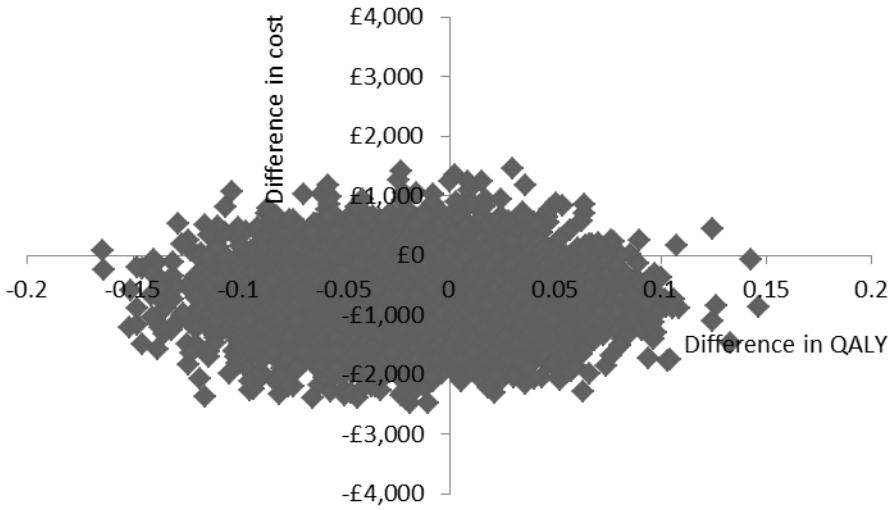
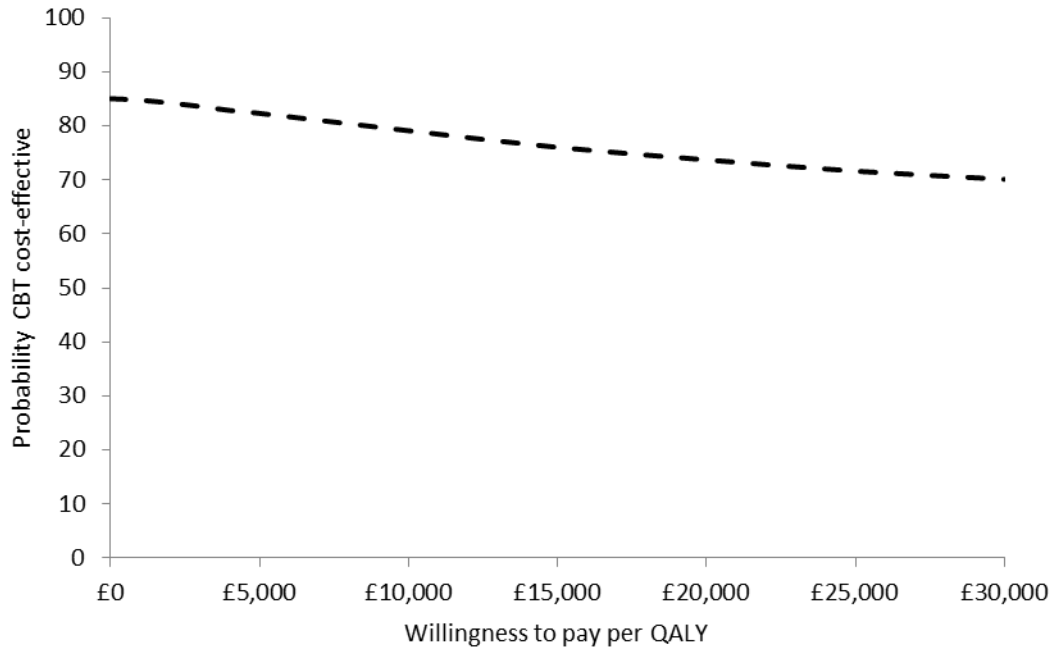


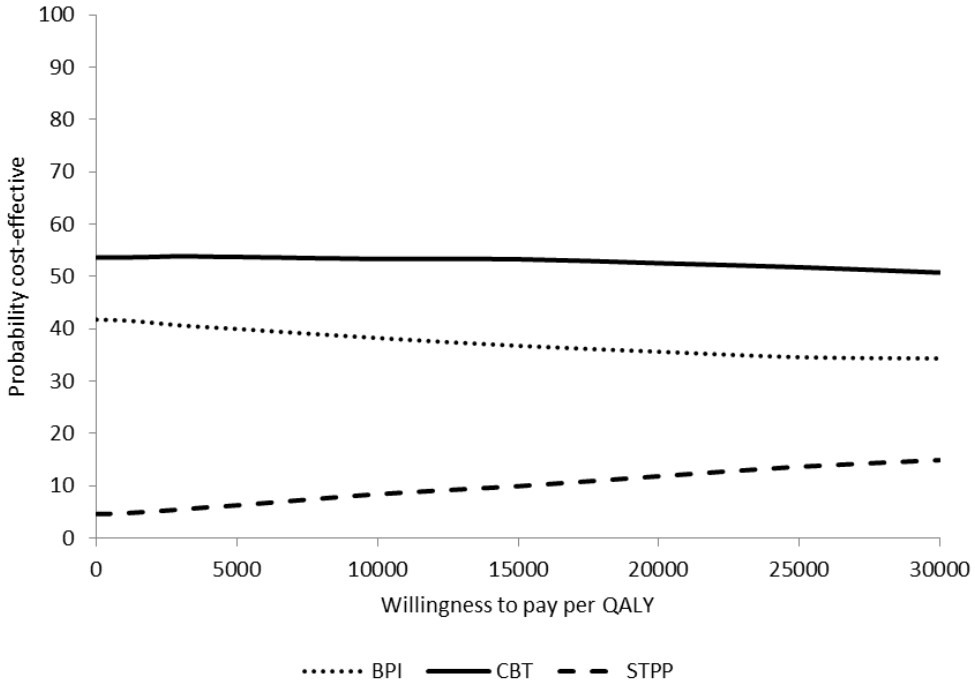
Figure A6: Cost-effectiveness acceptability curve showing the probability that CBT is cost-effective compared to STPP for different values a decision maker might be willing to pay for improvements in QALYs



CBT v STPP v BPI

The three interventions were compared head to head in a three-way comparison. The CEACs in Figure A7 show that for all values that a decision maker might be willing to pay for a QALY, CBT has the highest probability of being cost-effective.

Figure A7: Cost-effectiveness acceptability curve showing the probability that BPI, CBT and STPP are cost-effective for different values a decision-maker might be willing to pay for improvements in QALYs



Sensitivity analysis

The results of the sensitivity analyses are detailed in Tables A16 and A17. Multiple imputation did not alter the direction of the differences in cost, nor did re-analysis using cost per week rather than cost over the entire follow-up period. Including an estimate of the cost of sessions that were scheduled but which the young person did not attend, however, altered the order between the three interventions.

For the sample with full economic data, the average number of sessions that were offered but were not attended were three in the BPI group, 14 in the CBT group, and six in the STPP group. The inclusion of the cost of these sessions (at 50% of the cost of a full session) resulted in the average cost of CBT (£3,050) becoming more expensive than the BPI mean cost (£2,939), with STPP remaining the most costly group (mean cost £3,364).

Whilst there remain no statistically significant differences in cost between the groups, this change in direction impacts upon the cost-effectiveness analyses for the comparison of CBT and BPI. Figure A8 shows the scatter plot for this comparison; the majority of the replications are in the North-East and North-West quadrants denoting higher costs in the CBT group (points above the x-axis). The very similar outcomes mean that the CEAC in Figure A9 suggests that the probability that CBT is cost-effective compared to BPI is less than 50% for all values a decision maker might be willing to pay for a QALY. Figure A10 shows a head to head comparison of all three groups in terms of cost-effectiveness and including a cost for sessions missed. It demonstrates that there is a higher probability of BPI being cost-effective compared to CBT and STPP, for all values of willingness to pay.

Table A13: Sensitivity analyses for costs (£) over 86-week follow-up

	BPI		CBT		STPP	
	Mean	SD	Mean	SD	Mean	SD
Base case analysis	2678·39	2881·89	2379·01	3643·85	3081·70	3573·17
Non-attendance at 50% cost	2907·30	2939·08	3050·05	5891·69	3364·14	3563·08
Multiple imputation	-	-	-	-	-	-
Total cost per week	28·76	31·63	25·25	38·35	32·42	35·84

Table A14: Between group differences for sensitivity analysis at 86-week follow-up

Comparison	Sensitivity analysis	Coefficient	95% confidence interval	p-value
CBT v BPI	Base case	-338·54	(-1333·17 to 656·09)	0·503
	Non-attendance at 50% cost	185·15	(-392·71 to 1657·16)	0·225
	Multiple imputation	-425·07	(-1384·58 to 534·43)	0·381
	Total cost per week	-3·95	(-14·58 to 6·68)	0·464
STPP v BPI	Base case	609·55	(-406·73 to 1625·83)	0·238
	Non-attendance at 50% cost	632·21	(-392·71 to 1657·16)	0·225
	Multiple imputation	448·95	(-609·77 to 1507·66)	0·399
	Total cost per week	6·12	(-4·47 to 16·72)	0·256
CBT v STPP	Base case	-709·23	(-1836·04 to 417·58)	0·216
	Non-attendance at 50% cost	-429·79	(-1955·24 to 1095·65)	0·579
	Multiple imputation	-891·47	(-1951·81 to 168·86)	0·098
	Total cost per week	-7·46	(-19·10 to 4·17)	0·207

* All adjusted for region and baseline cost, behavioural disorder and antidepressant use

Figure A8: Sensitivity analysis – Scatter plot of differences in costs versus differences in QALYs for CBT versus BPI with non-attendance at 50% session cost

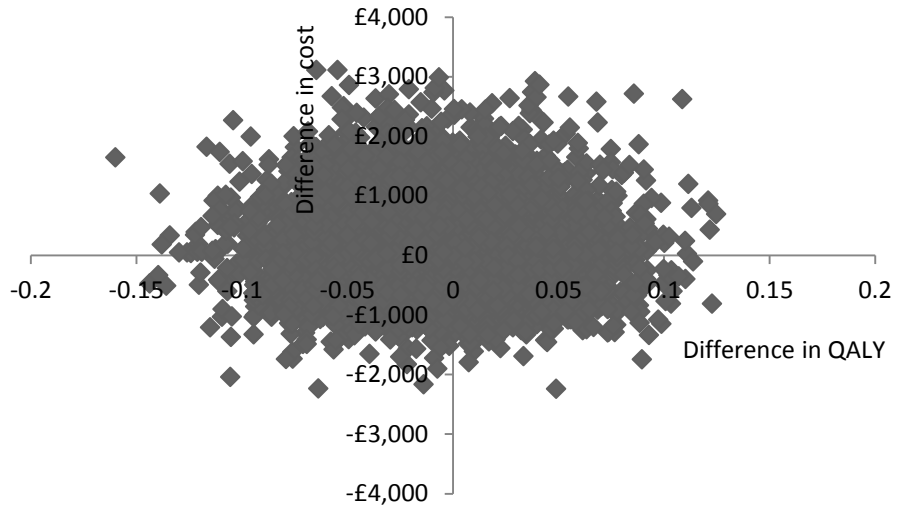


Figure A9: Sensitivity analysis – Cost-effectiveness acceptability curve showing the probability that CBT is cost-effective compared to BPI for different values a decision maker might be willing to pay for improvements in QALYs with non-attendance at 50% session cost

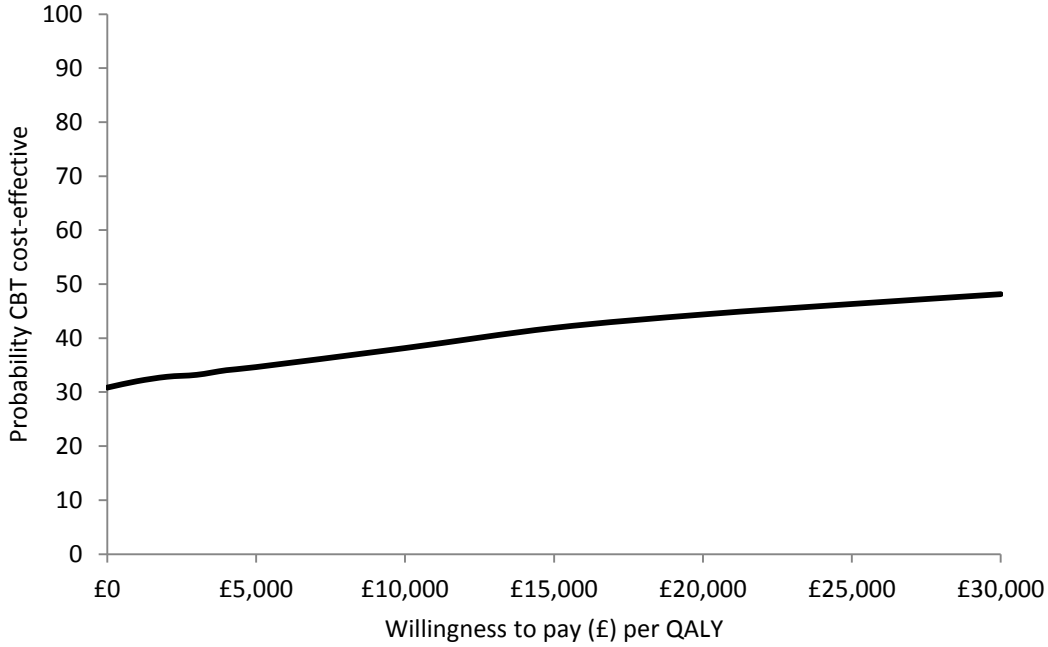
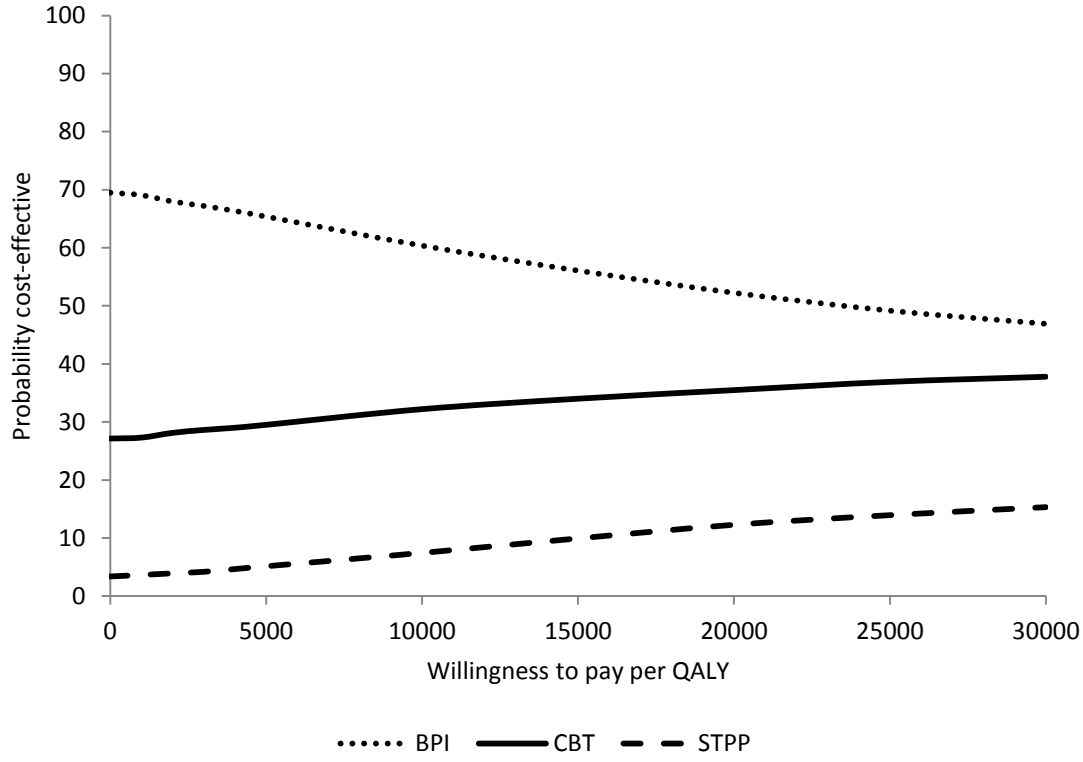


Figure A10: Sensitivity analysis – Cost-effectiveness acceptability curve showing the probability that CBT, STPP and BPI are cost-effective for different values a decision-maker might be willing to pay for improvements in QALYs with non-attendance at 50% session cost



Append iv:

Moderation of treatment effects

Little is understood regarding factors that may influence treatment response in depressed adolescents. This study included 2 putative cognitive processes that the literature suggests may moderate therapeutic response to different psychological treatments. These are:

i) Individual differences in self-reported ruminative thinking whilst depressed. A ruminative response style is defined as persistently brooding or dwelling on current depressive thoughts and feelings, often to the exclusion of other themes in the patient's life. [20]

ii) The quality of predominant depressive experiences, which is, defined as possessing a thinking style (dependent or self-critical) likely to predispose or be associated with depressive illness but not synonymous with a pattern of symptoms. [21]

Ruminative response style

Rumination is the compulsively focused attention on the symptoms of one's distress, and on its possible causes and consequences, as opposed to its solutions. [22] Rumination is similar to worry except rumination focuses on bad feelings and experiences from the past, whereas worry is concerned with potential bad events in the future. [23] Both rumination and worry are associated with clinical anxiety and depression. [23] Rumination has been widely studied as a cognitive vulnerability factor for depression, however its measures have not been unified. In the Response Styles Theory proposed by Nolen-Hoeksema [23] rumination is defined as "compulsively focused attention on the symptoms of one's distress, and on its possible causes and consequences, as opposed to its solutions". Because the

Response Styles Theory has been empirically supported, this conceptual model of rumination is the most widely used.

Depressive experiences style

Both theoretical assumptions and empirical findings suggest that adult patients with clinical depression may be characterized by a cognitive styles of excessive preoccupation with relatedness (principally focused on disappointment with relationships) and self-definition or identity (principally focused on self-criticism). [24] As such individuals with depression may be predominantly troubled by one of the following issues which have been shown as 2 independent factors in the depression experiences style self report scale:

i) High concerns about the quality of interpersonal relatedness with feelings of emptiness and loneliness, and intense fears of being abandoned and left unprotected. Termed ***dependent/relatedness***

ii) Possessing an extremely self-critical attitude together with feelings of worthlessness, guilt, failure, and self-blame. Termed **self-critical/identity**

Moderator effects on the primary outcome were investigated by adding an interaction between the moderator variable and treatment allocation to the primary analysis model. Table 16 gives the estimates of the treatment by moderator effect for each of the moderator hypotheses proposed in the methods section (see Chapter 7 for details of measures and chapter 8 for analytic strategy and hypotheses). A negative estimate in this table indicates that a higher score of the moderator lowered the MFQ for the treatment relative to the comparator, that is an increase in the beneficial treatment effect.

First, we hypothesized that young people with elevated dependency sub scale sum scores on the Depressive Experiences Questionnaire (DEQ) would have greater reduction in MFQ if they received STPP rather than BPI or CBT treatment than those with lower scores. Before 36 weeks the direction of the effect was consistent with our hypothesis but this was not statistically significant (p=0.168). After 36 weeks there was clearly no evidence of an effect (p=0.918).

Secondly, we hypothesized that young people with elevated self-critical sun scale sum scores on the Depressive Experiences Questionnaire would have a better response if they received CBT rather than either BPI or STPP treatment. The direction of the effect was consistent with our hypothesis both before and after 36 weeks but with a significant trend by 36 weeks (0.053) but no evidence subsequently (p=0.384).

Finally, we hypothesized that higher total scale scores for rumination response style of thinking when depressed (RSS) would show a better treatment response for CBT than BPI or STPP treatment. There was no evidence for such an effect either before (p=0.671) or after (p=0.976) thirty-six weeks

Table A15: Treatment moderator analyses for the primary outcome (MFQ) based on the LME model with main effects for treatment with a moderator by treatment interaction

	<36 weeks			≥ 36 weeks		
	Mod. Effect	(95% c.i.)	p-value	Mod. Effect	(95% c.i.)	p-value
DEQ Dependency STPP vs (BPI+CBT)	-0.21	(-0.51 to 0.09)	0.168	0.02	(-0.35 to 0.39)	0.918
STPP vs BPI	-0.29	(-0.64 to 0.06)		-0.01	(-0.44 to 0.43)	
STPP vs CBT	-0.13	(-0.48 to 0.23)		0.05	(-0.40 to 0.49)	
DEQ self-criticism CBT vs (BPI+ STPP)	-0.36	(-0.72 to 0.05)	0.053	-0.20	(-0.66 to 0.25)	0.383
CBT vs BPI	-0.42	(-0.85 to 0.02)		-0.21	(-0.74 to 0.32)	
CBT vs STPP	-0.31	(-0.72 to 0.10)		-0.20	(-0.73 to 0.33)	
Ruminative response scale CBT vs (BPI + STPP)	0.04	(-0.14 to 0.22)	0.671	0.004	(-0.23 to 0.23)	0.975
CBT vs BPI	0.02	(-0.18 to 0.22)		-0.06	(-0.31 to 0.19)	
CBT vs STPP	0.07	(-0.14 to 0.28)		0.08	(-0.18 to 0.35)	

Note, negative effects indicate benefit for high scores of moderators

Appendix ii and iii references

1. Hilsenroth MJ, Blagys MD, Ackerman SJ, Bonge DR, Blais MA. Measuring Psychodynamic-Interpersonal and Cognitive-Behavioral Techniques: Development of the Comparative Psychotherapy Process Scale. Psychotherapy: Theory, Research, Practice, Training, 2005; 42(3): 340-56.
2. Goldman R, Hilsenroth MJ, Gold J, Owen J. Psychotherapy integration and alliance: Use of cognitive-behavioral techniques within a short-term psychodynamic treatment model. . Journal of Psychotherapy Integration, 2013; 23: 373–85.
3. Hilsenroth MJ, Blagys MD, Ackerman SJ, Bonge DR, Blais MA. Measuring Psychodynamic-Interpersonal and Cognitive-Behavioral Techniques: Development of the Comparative Psychotherapy Process Scale. Psychotherapy: Theory, Research, Practice, Training 2005; 42 340-56.
4. Stein M, F. P, J. S, Hilsenroth M. A training outline for conducting psychotherapy process ratings: An example using Therapist Technique. Counsellig Psychotheapy Research 2010; 10: :50 –9.
5. Linden M. How to define, find and classify side effects in psychotherapy: from unwanted events to adverse treatment reactions. Clinical psychology & psychotherapy. 2013;20(4):286-96.
6. National Institute for health and Care Excellence. Guide to the methods of technology appraisal. London: NICE, 2013.
7. Byford S, Barrett B, Roberts C, et al. Cost-effectiveness of selective serotonin reuptake inhibitors and routine specialist care with and without cognitive behavioural therapy in adolescents with major depression. British Journal of Psychiatry 2007; 191: 521-7.
8. Curtis, L. (2012) Unit costs of health and social care 2012. Canterbury: PSSRU.
9. British Medical Association and the Royal Pharmaceutical Society. British National Formulary. London: BMJ Books & Pharmaceutical Press.; 2011.
10. Department of Health (2013) NHS Reference Costs 2011-2. London: Department of Health.
11. Dolan P, Gudex C, Kind P, Williams A. A social tariff for EuroQoL: results from a UK general population survey. Centre for Health Economics, York: University of York, 1995.

12. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. Health Economics 2005; 14(5): 487-96.
13. Richardson G, Manca A. Calculation of quality adjusted life years in the published literature: a review of methodology and transparency. Health Economics 2004; 13(12): 1203-10.
14. Efron B, Tibshirani R. An introduction to the bootstrap. New York: Chapman and Hall; 1993.
15. Thompson SG, Barber JA. How should cost data in pragmatic randomised trials be analysed? BMJ 2000; 320(7243): 1197-200.
16. Claxton, K., Sculpher, M., McCabe, C., Briggs, A., Akehurst, R., Buxton, M., Brazier, J., & O'Hagan, T. 2005. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. Health Economics, 14, 339-347
17. Drummond M, Sculpher M, Torrance GL, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press; 2005.
18. Fenwick E, Byford S. A guide to cost-effectiveness acceptability curves. Br J Psychiatry 2005; 187: 106-8.
19. Weichle T, Hynes DM, Durazo-Arvizu R, Tarlov E, Zhang QI. Impact of alternative approaches to assess outlying and influential observations on health care costs. . Springerplus 2013; 18: 614.
20. Nolen-Hoeksema S. The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. . *J Abn Psychology* 2000;**109**:504–11.
21. Blatt SJ, Schaffer CE, Bers SA, Quinlan DM. Psychometric properties of the Depressive Experiences Questionnaire for adolescents. *J Pers Assessment* 1992;**59**:82-98.
22. Smith JM, Alloy LB. A roadmap to rumination: a review of the definition, assessment, and conceptualization of this multifaceted construct. *Clin Psychoi Review* 2009;**29**:116-28.
23. Nolen-Hoeksema S, Wisco BE, Lyubomirsky S. Rethinking Rumination. Perspectives on psychological science : J Assoc Psychol mScience. 2008;**3**:400-24.
24. Blatt SJ, Luyten P. A structural-developmental psychodynamic approach to psychopathology: Two polarities of experience across the life span. . *Dev Psychopathology* 2009;**21**:793–814.