Efficacy and cost-effectiveness of a specialist depression service versus usual specialist mental health care to manage persistent depression: a randomised controlled trial

*Richard Morriss, Anne Garland, Neil Nixon, Boliang Guo, Marilyn James, Catherine Kaylor-Hughes, Richard Moore, Rajini Ramana, Christopher Sampson, Timothy Sweeney, Tim Dalgleish, and the NIHR CLAHRC Specialist Mood Disorder Study Group**

Published Online

August 3, 2016 http://dx.doi.org/10.1016/ S2215-0366(16)30143-2

*Members of CLAHRC Specialist Mood Disorder Study Group are listed in the Acknowledgments

Department of Psychiatry and Applied Psychology, Institute of Mental Health(Prof R Morriss MD, B Guo PhD, C Kaylor-Hughes DPhil) and Division of Rehabilitation and Ageing, School of Medicine (Prof M James PhD, C Sampson MSc), University of Nottingham, Nottingham, UK;

Adult Mental Health Directorate, Nottinghamshire Healthcare Trust, Nottingham, UK (A Garland RMN, N Nixon MD, T Sweeney RMN);

Independent Practitioner, Cambridge, UK (R Moore PhD);

Cambridge and Peterborough Partnership NHS Foundation Trust, Cambridge, UK (R Ramana MBBS);

Medical Research Council Cognition and Brain Sciences Unit, Cambridge, UK (Prof T Dalgleish PhD)

Correspondence to: Prof Richard Morriss,

Department of Psychiatry and Applied Psychology, Institute of Mental Health, University of Nottingham, Nottingham NG7 2TU, UK richard.morriss@nottingham.ac.uk

Summary

Background Persistent moderate or severe unipolar depression is common and expensive to treat. Clinical guidelines recommend combined pharmacotherapy and psychotherapy. Such treatments can take up to 1 year to show an effect, but no trials of suitable duration have been done. We investigated the efficacy and cost-effectiveness of outpatient- based, specialist depression services (SDS) versus treatment as usual (TAU) on depression symptoms and function.

Methods We did a multicentre, single-blind, patient-level, parallel, randomised controlled trial (RCT), as part of the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care (CLAHRC) study, in three mental health outpatient settings in England. Eligible participants were in secondary care, were older than 18 years, had unipolar depression (with a current major depressive episode, a 17-item Hamilton Depression Rating Scale [HDRS17] score of ≥16, and a Global Assessment of Function [GAF] score of ≤60), and had not responded to 6 months or more of treatment for depression. Randomisation was stratified by site with allocation conveyed to a trial administrator, with research assessors masked to outcome. Patients were randomised (1:1) using a computer-generated pseudo-random code with random permuted blocks of varying sizes of two, four, or six to either SDS (collaborative care approach between psychiatrists and cognitive behavioural therapists for 12 months, followed by graduated transfer of care up to 15 months) or to the TAU group. Intention-to-treat primary outcome measures were changes in HDRS17 and GAF scores between baseline and 6, 12, and 18 months' follow-up. We will separately publish follow-up outcomes for months 24 and 36. Clinical efficacy and cost-effectiveness were examined from health and social care perspectives at 18 months, as recommended by the National Institute for Health and Care Excellence. This trial is registered at ClinicalTrials.gov (NCT01047124) and the ISRCTN registry (ISRCTN10963342); the trial has ended.

Findings 307 patients were assessed for eligibility between Dec 21, 2009, and Oct 31, 2012. 94 patients were assigned to TAU and 93 patients to SDS, and were included in intention-to-treat analyses. The changes from baseline to 6 months in HDRS17 and GAF scores did not significantly differ between treatment groups (mean change difference in HDRS17 score –1·01 [95% CI –3·30 to 1·28], p=0·385; and in GAF score 1·33 [–2·92 to 5·57], p=0·538). Primary outcome data were available for 134 (72%) patients at 12 months. We noted no differences at 12 months' follow-up between SDS and TAU for mean HDRS17 score (14·8 [SD 7·9] in the SDS group vs 17·2 [7·3] in the TAU group; p=0·056) or GAF score (60·4 [11·7] vs 55·8 [12·7]; p=0·064), and the changes from baseline to 12 months in HDRS17 and GAF scores did not significantly differ between treatment groups (mean change difference in HDRS17 score –2·45 [95% CI –5·04 to 0·14], p=0·064; and in GAF score 4·12 [–0·11 to 8·35], p=0·056). The mean change in HDRS17 score from baseline to 18 months was significantly improved in the SDS group compared with the TAU group (13·6 [SD 8·8] in the SDS group vs 16·1 [6·6] in the TAU group; mean change difference –2·96 [95% CI –5·33 to –0·59], p=0·015), but the GAF scores showed no significant differences between the groups (61·2 [SD 13·0] vs 57·7 [11·9]; mean change difference 3·82 [–9·3 to 8·57], p=0·113). We reported no deaths, but one (1%) patient was admitted to hospital for myocardial infarction, and three episodes of self- harm were reported in three (2%) patients (two receiving TAU, one receiving SDS care). The incremental cost-effectiveness ratio of SDS versus TAU was £43 603 per quality-adjusted life-year.

Interpretation Compared with usual specialist mental health secondary care, SDS might improve depression symptoms for patients with persistent moderate to severe depression, but functional outcomes and economic benefits are equivocal.

Funding National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care, UK Medical Research Council, Nottinghamshire Healthcare NHS Foundation Trust, Derbyshire Healthcare NHS Foundation Trust, Cambridgeshire and Peterborough NHS Foundation Trust, University of Nottingham.

Research in context

Evidence before the study

The Clinical Guideline for Depression (2004) from the National Institute for Clinical Excellence (NICE) had not found any randomised controlled trial (RCT) of collaborative care or other service interventions for chronic or treatment-resistant depression in secondary mental health care. Again, no RCT was identified by NICE in their clinical update in 2009. We searched PubMed and the Cochrane Library for RCTs published between Jan 1, 2003, and May 4, 2016, in English. We used the same search criteria for RCTs and systematic reviews, as outlined by NICE in their 2004 and 2009 guidelines, using the terms "treatment resistant depression", "recurrent depression", "atypical depression", "psychotic depression", "chronic depression", "persistent depression", or "specialist mental health", "collaborative care", "stepped care", "combined modality care", or "psychotherapy". One inpatient RCT reported clinically important benefits in depression symptoms and function from an intensive programme of interpersonal therapy plus medication compared with intensive clinical management with medication. One single-centre feasibility study of a brief intervention of collaborative care for chronic or persisting moderate or severe depression compared with usual outpatient care in a specialist mental health setting reported equivocal findings. A meta-analysis with meta-regression of RCTs of combined antidepressant and psychotherapy compared with other clinical management for chronic depression reported benefits of combined treatment in chronic major depression, but not dysthymia, although substantial heterogeneity existed in depression symptom outcomes. A subsequently completed RCT of psychodynamic therapy added to usual care in specialist mental health outpatient settings also suggested efficacy for chronic persistent major depression, but only after 24 months.

Added value of this study

Our study was the first multicentre, outpatient RCT in patients with persistent moderate or severe unipolar depression, comparing collaborative care involving combined psychological and pharmacological treatment delivered by a specialist mental health depression service (SDS) with usual specialist mental health care under the direction of a consultant psychiatrist. We showed some evidence of improvement in depression symptoms, but not function, with SDS by 18 months' follow-up compared with usual care, but SDS cost more than usual care.

Implications of all the available evidence

Overall, in people with persisting depression in specialist mental health settings, collaborative care involving integrated psychological and pharmacological treatment lasting at least 12 months shows evidence of efficacy against depression symptoms compared with usual care after 18–24 months. However, SDS is more expensive in the short term than usual care, and therefore its implementation is dependent on treatment centres' willingness to pay the upfront costs of treatment. Brief similar interventions are unlikely to be effective.

Introduction

The Global Burden of Disease Study (GBD) 2013[1] identified major depressive disorder as the second leading cause of years lived with disability in the world. Major depressive disorder is an expensive condition with direct costs of £1·7 billion per year (through use of services, mainly non-psychiatric National Health Service [NHS] use and social service delivery) and indirect costs of £7·5 billion per year (mainly through lost employment) in England alone in 2008.[2] These costs escalate substantially as 7% of patients with depression in secondary mental health care develop a chronic unremitting course,[3] for whom evidence indicates will have particularly high rates of admissions to hospital, functional impairment, and suicide.[4]

Despite these high human and economic costs, research on treatment for chronic depression remains largely limited to single interventions, such as pharmacotherapy or psychotherapy. However, even intensive use of these individual approaches leaves many cases unremitted. The STAR*D randomised controlled trial[5] (RCT) in 3671 outpatients with major depressive disorder showed that after 12 months of systematic and tailored pharmacological or psychological treatment, only 67% of patients would achieve remission over 1 year. Clinical guidelines, such as those of the National Institute for Health and Care Excellence (NICE),[6] therefore recommended specialist multi-professional care and combined psychotherapy and pharmacotherapy for chronic moderate-to-severe depression in patients who had not responded to generic specialist mental health care treatment. However, only one small pilot trial[7] (n=64) has investigated the effect of combining psychotherapy and optimised pharmacotherapy in an integrated specialist depression service (SDS) for patients with persistent moderate-to-severe unipolar depression. The trial reported mixed results, showing increased remission rates at the end of 4 months of treatment in the absence of significant differences in response or depressive symptom reduction, compared with community treatment as usual (TAU).[7] A meta-analysis[8] of RCTs of psychological treatment showed that treatment response for chronic major depressive disorder was proportional to the number of treatment sessions offered.

Therefore, we investigated the clinical efficacy and cost-effectiveness of an SDS on depression symptoms and function compared with usual specialist mental health treatment as directed by a consultant psychiatrist.

Methods

Study design and participants

We did a multicentre, single-blind, patient-level, parallel-group, randomised controlled study in secondary care psychiatry settings within three NHS trusts in England (Derbyshire Healthcare NHS Foundation Trust, Nottinghamshire Healthcare NHS Foundation Trust, and Cambridgeshire and Peterborough NHS Foundation Trust). We recruited patients with persistent moderate or severe depression. The protocol is available online.[9]

The definition of persistent depression was deliberately pragmatic, encouraging the referral of patients who had not responded to secondary mental health care for at least 6 months, and in whom their main current problem was depression. Inclusion criteria were that the patient was thought by the referrer to have primary unipolar depression (i.e., not caused by another psychiatric disorder); aged older than 18 years; able and willing to give oral and written informed consent to participate in the study; must have been offered or received direct and continuous care from one or more health professionals in the preceding 6 months and currently be under the care of a secondary care mental health team; had a structured clinical interview for DSM-IV (SCID) diagnosis of major depressive disorder with a current major depressive episode;[10] met five of nine NICE criteria for symptoms of moderate depression; had a 17-item Hamilton Depression Rating Scale[11] (HDRS) of at least 16; and had a Global Assessment of

Functioning[12] (GAF) score of 60 or lower, to ensure that the patient had a current moderate or severe major depressive episode. Patients were excluded if they were in receipt of emergency care for suicide risk, at risk of severe neglect, or a homicide risk, but patients were not excluded because of such risk provided the risk was adequately contained in their current care setting and the primary medical responsibility for care was with the referral team. They were also excluded if they did not speak fluent English; were pregnant; or had unipolar depression secondary to a primary psychiatric or medical disorder, except when bipolar disorder was identified by the research team after referral as unipolar depression because an SDS would be expected to manage bipolar depression in clinical practice.

All participants were referred to the study by a mental health professional working in one of the three participating specialist mental health organisations. All participating patients provided both oral and written informed consent to participate in the study. Ethics approval was obtained from the National Research Ethics Service in Derby, UK. The statistical and economic analysis plan is available online.[13]

Randomisation and masking

Randomisation was stratified by site (Cambridge, Derby, or Nottingham) as per protocol and done through a registered centralised clinical trials unit (Queens Medical Centre, Nottingham, UK). Once baseline assessments of eligible participants were completed by the research staff, patient details were sent to the clinical trials unit by the trial administrator. Patients were assigned to either SDS treatment or TAU by use of a computer-generated pseudo-random code with random permuted blocks of varying sizes of two, four, or six, and was created by the clinical trials unit in accordance with their standard operating procedure and stored on a secure server.

Patients were allocated with equal probability (1:1) to each treatment group. Allocation of eligible patients to a treatment group was conveyed to the trial administrator, who relayed this information to a secretary supporting the SDS and the referring clinician who were expected to organise a patient's care if allocated to TAU. Only the trial administrator had password controlled access to the randomisation data. Research associates who completed outcome assessments did not have access to the patient's health service records, and were based at a site away from clinical care. Although patients and treating clinicians were aware of treatment allocation, outcome assessors and data managers were masked to study allocation at all time points throughout the 18 month follow-up. All unmasking events were recorded.

Procedures

The SDS intervention[9] consisted of NICE-recommended pharmacological and psychological treatments for depression as a specialty service within specialist mental health care, with a collaborative care approach between psychiatrists and cognitive behavioural therapists over 12 months, followed by a graduated transfer of care up to an additional 3 months (15 months in total) to primary care or usual specialist mental health care. Each participant concurrently received both specialist pharmacotherapy for depression and specialist Beckian cognitive behavioural therapy (CBT) for depression[14] (appendix). Optimised pharmacotherapy followed the principle of rapid switching if evidence showed non-response, using individually tailored next-step options from evidence- based alternatives.[15] We aimed to review participants at 1–2 week intervals during medication initiation, dose escalation, switching, or augmentation. Shared decision making between treatment clinicians underpinned drug treatment and, when indicated, efforts were made to enhance concordance with treatment by use of various methods of medication monitoring. Evidence of response was sought within a maximum of 4 weeks and non-response by this point led to discussion of treatment alternatives. Medication options included the full range of individual antidepressants, such as monoamine oxidase inhibitors; antidepressant combination strategies, such as SSRI and tricyclic antidepressants; and augmentation with lithium, liothyronine (tri-iodo-thyronine), antipsychotics, or modafinil. Medication choice was tailored to the degree of

response, tolerability, and complexity (associated medical and psychiatric comorbidity). Partly related to the generally high level of complexity, judicious use was made of anxiolytic and hypnotic medication during treatment and transitions in treatment, which included short-term use of benzodiazepines, pregabalin, and buspirone.

Collaborative discussions (between treating psychiatrists, CBT therapist, and patient) were about the effect of medication change within psychotherapy and in more severe cases the focus was on medication in the early part of treatment, since impaired cognitive executive function made engagement in psychotherapy difficult. Once medication had been optimised, medical review typically became less intense, with follow-up in this later phase of up to 3–4 months when psychotherapy was often a greater focus. When clinically indicated, the psychological interventions were augmented as part of relapse prevention (within the timeframe of the overall intervention) with mindfulness-based cognitive therapy.[16] Additionally, when the clinical presentations identified high levels of shame and self-criticism, compassion- focused therapy[17] integrated within a CBT framework was used as part of relapse prevention. Further, when clinically indicated, social inclusion initiatives were used in the context of CBT, which included vocational-based and occupational-based activities. The social inclusion services fostered and encouraged participants' access to self-help support groups.

Collaborative care incorporated the central principles outlined by Gunn and colleagues.18 Thus the psychiatrist and the CBT therapist did an initial joint or concurrent assessment lasting 1·5 h, after which a structured treatment plan was generated collaboratively. This initial assessment by the SDS team (before baseline assessments) was followed by up to two further assessment appointments completed separately by each clinician. Assessing clinicians also treated participants. At each meeting with a participant, psychotherapists and psychiatrists emphasised that both medication and psychological treatment should be adhered to. Joint review meetings were done at 3, 6, 9, and 12 months into treatment between participant, psychiatrist, CBT therapist, and any other relevant party whom it was felt clinically appropriate to include, such as a general practitioner, key worker from referring team, or family member. Each team had a dedicated administrator who acted as a central point of contact for both participants and team members. Enhanced inter-profession communication was facilitated by meetings once every 2 weeks to discuss cases and by a shared local intranet workspace.

Participants in the SDS group received 12 months of treatment within the service. Initially, each participant met once a week with the psychiatrist; once a suitable treatment regimen was established, and meetings were tapered as needed. Participants received CBT sessions once a week for up to 10 months and after then these were tapered to intervals of once every 2 weeks and once every 3 weeks. After 12 months of treatment, discharge was made to primary or secondary care teams as clinically indicated and was planned as a gradual transition with reviews once every 4 weeks with the participants for up to 3 months and explicit recommendations for continued care were relayed in writing, face to face, or by telephone to the clinician providing follow-up care, and to the participant.

The pharmacological interventions were delivered by four consultant psychiatrists with a special interest and high level of expertise in the treatment of depression. The psychological interventions were delivered by mental health professionals (six mental health nurses, one medical consultant psychotherapist, and one clinical psychologist) trained in cognitive and behavioural therapies on courses accredited by a national body for CBT practice, with extensive experience of treating complex depression with CBT (median 19 years' experience) and use of peer supervision of team members in the SDS. The Cambridge service also included a community psychiatric nurse as part of the team, who did the usual duties of the role.

For participants receiving TAU, this treatment was directed by a consultant psychiatrist, usually from a community mental health team. It usually consisted of individual work by the psychiatrist in secondary care and treatment with pharmacotherapy (often by use of augmentation and change strategies for depression not responding to trials of single antidepressants), sometimes shared reviews with primary

care, and sometimes used psychosocial interventions such as CBT, counselling, or community psychiatric nurse support. Little evidence was available of joint reviews of progress and regular professional meetings being used. Responses to medication and treatment were reviewed only by the provider of the treatment.

Participants in both the SDS and TAU groups had access to the full range of local NHS support services open to patients in secondary mental health care (e.g., access to inpatient and crisis teams, occupational therapy, and community psychiatric nurse input).

Outcomes

Primary outcome measures were changes in assessment scores of HDRS[17] and GAF,[12] from baseline to 6, 12, and 18 months, with a prespecified primary outcome assessment at 6 months and 12 months. However, the original protocol allowed SDS treatment for longer than 12 months so that planned transfer of care could be graduated for up to 3 months for these patients, from the SDS to a primary or secondary mental health clinical team. The median length of SDS treatment including the transfer of care was 13 months (IQR 12–15). We therefore anticipated an increase in the difference between the SDS group and the TAU group, in depression and function between 12 months and 18 months because not all of the SDS intervention was completed by 12 months. Consequently, we report the primary outcome at both 12 months and at 18 months after the completion of all SDS treatment and as recommended by NICE in assessment of the outcomes for service interventions for depression.[6] The protocol[9] specified that all participants would be assessed at baseline and at 3, 6, 9, 12, 18, and 24 months. However, with further funding of the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care (CLAHRC) and agreement of an independent scientific committee, participants were followed up for 36 months given their long duration of illness. We report results here for up to 18 months follow-up; results at 24 months and 36 months will be separately reported.

Secondary outcomes were changes in self-rated measures of depression on the Beck Depression Inventory version I (BDI-I),[19] the nine-item Personal Health Questionnaire (PHQ-9),[20] and the Quick Inventory of Depressive Symptomatology Self-Report (QIDS-SR),[21] from baseline to 3, 6, 9, 12, and 18 months. Self-report measures of function and quality of life—the modified Social Adjustment Scale (SAS-M)[22] and the EQ-5D-3L)[23]—were collected at baseline and at 6, 12, and 18 months. EQ-5D-3L index scores were estimated by use of the Measurement and Valuation of Health A1 tariff to obtain health state use values (i.e., to measure quality of life).[24] Quality-adjusted life-years (QALYs) were calculated on the basis of the area under the curve of the EQ-5D-3L scores over time, assuming linear interpolation between time points. Patient satisfaction with the treating team was rated by the patient with two questionnaires (the Patient Satisfaction Questionnaire and the Patient Doctor Relationship Questionnaire) at baseline and at 6, 12, and 18 months on a five-point scale.[25] Adverse and serious adverse events were death including suicide, reported self-harm, admission to hospital for physical health problems, or any act of harm to other people. Data for these safety endpoint were obtained by patient self-report, clinician report, and hospital records. The Patient Satisfaction Questionnaire was found to be confusing and burdensome to patients so we stopped collecting this questionnaire following their feedback.

Statistical analysis

We first based our sample size calculation on improvement in global assessment of severity of depression in a study[26] that used a similar design, with 90% power at two-tailed 5% significance and 20% loss to follow-up, indicating that we needed 74 patients per treatment group (148 in total) at trial

entry. With further adjustments for the intraclass correlation of 0·051 due to level of variability in the community mental health team, we increased the final sample size to 174 patients by multiplying by a correction factor of 1·18, which was calculated by:[27]

$1 + p * n / (1 – p)$

where p is the intraclass correlation and n is the mean number of patients (3·5) per team.

The sample size was checked against a study[28] that compared inpatient combined psychotherapy and pharmacotherapy for depression treatment against TAU with HDRS17 as its primary outcome. At baseline, the mean $HDRS_{17}$ scores were 25·6 (SD 4·4) for combined treatment and 23·5 (4·8) for TAU.[28] At 12 months the mean $HDRS_{17}$ scores were 5·9 (SD 5·1) in the combined treatment group and 11·3 (10·5) in the TAU group. 102 patients were needed to detect such an effect size of 0·65 with 90% power at a two-sided 0·05 significance level.[28] If a 20% loss to follow-up was assumed, then 122 patients (61 in each group) were required. Using the correction factor of 1·18, as previously justified, we needed a sample size of 146 patients (73 in each group). In line with the more conservative estimate of the power of the study, our aim was to recruit 87 patients per treatment group (174 in total).[9]

We completed all analyses on an intention-to-treat basis. We report means and standard deviations of the mean. For the primary outcomes, we quantified the treatment effects on $HDRS_{17}$ score using multilevel modelling with time across all time points as a discrete variable, treatment group as a binary variable, and interaction between time and treatment group as fixed effects, baseline $HDRS_{17}$ score as a covariate, and patient as a level two unit. The treatment differences at every follow-up time point, together with their 95% CIs, were derived from the multilevel modelling. We used Markov chain Monte Carlo multiple imputation to impute missing data with REALCOM software under a missing-at-random assumption after exploration of the effect of the observed data on missing values.[29] To assess the robustness of our results and sensitivity to missing values, we also ran multilevel modelling with only observed scores. We did similar multilevel modelling to analyse GAF scores and all secondary outcome variables. We did covariate adjustment analysis by including site as a covariate in multilevel modelling. We ran a sensitivity analysis for only observed data (excluding missing data) for the primary outcomes and secondary depression outcomes to check how the results of the main analyse would be affected by exclusion of these variables and also by any participants without 6, 12, or 18 month outcomes.

As part of the economic analysis, therapists and psychiatrists completed bespoke forms to record their time dedicated to each intervention. We estimated the total staff costs associated with an intervention using published NHS staff earnings estimates.[30]

Data from an adapted version of the Client Services Receipt Inventory,[31] a self-completed resource-use questionnaire, were collected at baseline and at 3, 6, 9, 12, and 18 months' follow-up. If individuals had completed the questionnaire but did not report use of a particular service, we assumed that they had not used this service.

If duration was unknown, and information on average duration was not available from national data sources, we assumed a default duration of 15 min. We collected medication data using specifically designed medication collection proformas designed by the health economist (MJ) in conjunction with a psychiatrist (NN). These were constructed to enable easy recording and identification of the medications used by treating clinicians, the duration taken, and any changes made during the course of treatment to the drug regimen.

We used an NHS and personal social services cost perspective in accordance with NICE guidance, using 2014 prices.[32–34] Included costs are based on self-reported service use from the Client Services Receipt Inventory[31] and medication costs. The outcome measure for the economic evaluation was the number of QALYs. The analysis was based on the 18 month time horizon, and discounting was not applied to costs or outcomes.

We used the net monetary benefit framework and implemented a net benefit regression[35] to estimate the extent to which, and the probability that, the intervention was cost-effective at a range of threshold values for the willingness to pay per QALY. We used multiple imputation using chained equations and predictive mean matching to generate 90 imputed datasets separately for each intervention group, based on treatment centre, observed values of EQ-5D-3L index scores, and total costs at each time point. We used multilevel modelling with net monetary benefit at the individual level as the dependent variable, and ran this on the 90 imputed datasets.

The model estimated fixed effects associated with trial allocation, baseline EQ-5D-3L index score, and baseline costs. Differences between the three centres necessitated their inclusion as a level of observation in the model to account for centre effects. An independent CLAHRC scientific committee functioned as a trial steering committee and as a data monitoring committee, including a review progress of the study and statistical and health economics oversight. We used Stata (version 14) for multilevel modelling and used SPSS (version 21) for descriptive analysis. This trial is registered at ClinicalTrials.gov (NCT01047124) and the ISRCTN registry (ISRCTN10963342); the trial has ended.

Role of the funding source

The funders of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

307 patients were referred to the study and assessed for eligibility between Dec 21, 2009, and Oct 31, 2012, as part of the NIHR CLAHRC long-term follow-up of patient with persistent depression study. 228 patients (74%) completed a baseline interview, with non-progression at this stage attributable to various factors (figure 1). 41 patients (13%) of initial referrals were either screened out or withdrew after the baseline interview, so that 187 patients (61%) were randomly assigned to a treatment group. 21 patients (11%) were from the Derby site, 137 (73%) from the Nottingham site, and 29 (16%) were from the Cambridge site. One participant was randomly assigned at the Derby site but received treatment and was followed up by the Cambridge site because of the greater convenience of this arrangement to the patient; for the purposes of analysis, this patient is included as per randomisation in the Derby group of the study. For both groups, 144 patients (77%) were assessed at6 months' follow-up, 134 patients (72%) at 12 months, and 110 (59%) patients were assessed at 18 months' follow-up (figure 1). Unmasking events of outcome assessor were reported during follow-up in 71 patients (38%). Two participants, one in each treatment group, had baseline GAF scores higher than 60, a level identified in our protocol as an exclusionary criteria. These protocol errors were identified at 12 months' follow-up. The decision was taken by the study statistician before analysis was started, and agreed by the whole study team, to include these participants in the main analysis since they had been randomly assigned to a treatment, and had contributed their time to the study. In covariate adjustment analysis (including site as a covariate in multilevel modelling), the results showed almost identical treatment effects to those without adjustment for site effects.

At baseline for both groups the mean $HDRS_{17}$ score was in the moderate-to-severe depression symptoms range,[36] whereas the mean GAF score was in the range of serious impairment in social or occupational function[12] (table 1; appendix). Most participants had been ill with depression for many years, and had melancholia and psychosis in addition to other mental and physical comorbidities.

No significant differences were noted between groups at 6 months' follow-up in the mean change in $HDRS_{17}$ score (–1·01 [95% CI –3·30 to 1·28]; p=0·385) or GAF score (1·33 [–2·92 to 5·57]; p=0·538). No

significant differences were noted between groups at 12 months of follow-up in mean $HDRS_{17}$ score (14·8 [SD 7·9] in the SDS group vs 17·2 [7·3] in the TAU group; p=0·056; mean change in difference –2·45 [95% CI –5·04 to 0·14) and GAF score (60·4 [11·7] vs 55·8 [12·7]; p=0·064; mean change in difference 4·12 [95% CI –0·11 to 8·35]). By 12 months' follow-up, the change from baseline in $HDRS_{17}$ and GAF score did not significantly differ between treatment groups, but by 18 months' follow-up, the $HDRS_{17}$ score had improved by a significantly greater extent in the SDS group than in the TAU group; however, improvement in GAF score was still not significantly different between the groups at 18 months (figures 2 and 3; table 2). At 18 months, mean $HDRS_{17}$ score was lower (i.e., better) in the SDS group than in the TAU group (13·6 [SD 8·8] vs 16·1 [6·6], p=0·015), but the GAF score did not differ between the groups (61·2 [13·0] vs 57·7 [11·9], p=0·113). Between baseline and 18 months, the mean scores on the HDRS17 and GAF in the SDS group improved from moderate symptoms of depression and serious impairment in functioning to mild symptoms and mild impairment; in the TAU group, scores changed from moderate symptoms of depression to moderate severity, so patients remained in the moderate severity classification, and from serious impairment to moderate impairment in functioning.

Nearly all secondary outcome measures of the change in the severity of depression symptoms (BDI-I, PHQ-9, and QIDS-SR) had improved by a significantly greater extent by 9, 12, and 18 months with SDS treatment than with TAU, but the improvement in self-rated measure of social function (SAS-M) was not significantly different between groups at 6, 12, or 18 months (table 2). The observed-data only sensitivity analysis of the primary outcome and secondary outcome depression measures showed similar results to the intention-to-treat analyses (appendix). In terms of patient satisfaction with the treatment team, the change from baseline was significantly greater with SDS care than with TAU at 12 months, but not at the other time points (table 2). No deaths and no reported harm to other people were recorded. One (1%) patient was admitted to hospital for myocardial infarction that was judged to be unrelated to SDS treatment, and three episodes of self-harm (two [2%] patients in the TAU group, one [1%] patient in the SDS group) were recorded.

Response rates for the HDRS17 in the SDS group were 29% (27 of 93 patients) at 12 months and 40% (37 of 93) at18 months; in the TAU group were 20% (19 of 94) at 12 months and 25% (23 of 94) at 18 months. Remission rates in the SDS group were 20% (19 of 93 patients) at 12 months and 26% (24 of 93) at 18 months, and in the TAU group were 12% (11 of 94) at 12 months and 13% (12 of 94) at 18 months. Response and remission rates are reported descriptively and were not a prespecified outcome.[9]

The observed incremental QALY gain from all patients with complete follow-up data at all time points for the SDS group compared with the TAU group was 0·141. On the basis of the imputed data, and controlling for baseline differences in EQ-5D-3L by linear regression, the incremental benefit of SDS versus TAU was 0·079 QALYs. The incremental cost was £3446 (95% CI 1915–5180; table 3). Service use by patients and complete patient analyses are in the appendix. The mean cost of staff time associated with SDS was £2298 per patient (excluded from the analysis to avoid double-counting). The incremental cost-effectiveness ratio was £43 603. The results of the net benefit regression, which controlled for baseline differences and cluster effects, showed that SDS is more likely to be cost-effective than TAU above a willingness-to-pay threshold of £42 000 per QALY (appendix).

Discussion

In our RCT in patients with major depressive disorder, we found no differences in improvement in clinically rated depression symptoms and function by 6 months' and 12 months' follow-up between SDS treatment and TAU with specialist mental health care under the direction of a consultant psychiatrist in the NHS in the UK. At 18 months, for the outcome favoured by NICE for assessment of

the outcome of service interventions for depression,[2] the SDS group showed greater improvement in interview-rated depression symptoms than did the TAU group. Although the CIs between the two groups overlap substantially, the mean difference in $HDRS_{17}$ score reaches the minimum clinically important difference of three points.[36] A treatment effect of SDS care on depression symptoms is supported by significant or minimum clinically important differences between the SDS and TAU groups in all the self-rated depression symptom outcomes at 18 months' follow-up or earlier.[21,37,38] No significant improvements in function were reported at 18 months' follow-up.

The amount of care received in the TAU group, when patients were referred and accepted by a community mental health team (led by a consultant psychiatrist), was broadly similar in terms of psychiatrist, psychotherapy, and other community mental health team member contacts at two (Nottingham and Cambridge) of the three sites, according to a 12-month routine audit in people meeting similar entry criteria to those in our study, and who were not necessarily participants in the RCT. No routinely available data exist to show whether usual care at these three sites is similar to other sites in the UK. However, at some sites people with persistent moderate or severe depression might not always receive care in community mental health teams directed by a consultant psychiatrist. In this study, TAU often included complex pharmacology under the direction of a consultant psychiatrist, such as augmentation and change strategies, and selected participants also received courses of psychological treatment and other psychosocial interventions. The distinctive feature of SDS care, which was completely absent from TAU, was integrated psychiatric and psychological treatment sustained over 12 months with a collaborative care approach that systematically offered and reviewed specialist psychological and pharmacological treatment.

SDS treatment was associated with greater patient satisfaction at 12 months' follow-up than was TAU, but not at other time points. We recorded no adverse events attributable to SDS or TAU, although adverse events were reported in patients. Substantial improvements were noted in QALYs in patients receiving SDS compared with those receiving TAU, and the incremental cost per patient was £3446. However, at 18 months' follow-up SDS treatment was more expensive than the usual NICE willingness-to-pay threshold of £20 000 per QALY.[39]

Our RCT had broad inclusion criteria and few exclusion criteria, to reflect how services operate in usual NHS care. As a result, we systematically provided psychological treatment safely because it was integrated with psychiatric and other multidisciplinary care to people with long duration, seriously impairing depression with high levels of melancholia or psychosis, or other mental and physical comorbidity. A systematic review[15] of combined pharmacotherapy and psychotherapy versus pharmacotherapy alone for chronic depression reported no substantial differences in outcome, but these were single drug or psychotherapy approaches rather than the intensive tailored integrated combined treatment with various stepwise drug and psychotherapy treatments that we offered in our study. A recent, large RCT[40,41] showed sustained, cost-effective improvements in depression for up to 4 years with addition of CBT to antidepressant medication compared with antidepressant medication alone for primary care patients with chronic depression. However, in that study[40,41] the severity of depression and function, and duration of depression, were on average lower than in our study. The only previous RCT, in a small sample, of an outpatient specialist depression service that offered combined psychotherapy and pharmacotherapy for only 4 months versus community treatment, rather than specialist mental health care, showed that combined treatment had only limited benefits.[7] Evidence for serial offering of at least four different tailored pharmacological treatments for depression in 12 months[5] and the greater effectiveness of offering many psychological treatment sessions for chronic depression[8] suggested to us that at least 12 months of combined psychological and pharmacological treatment was required. The absence of any significant separation between the treatment groups for any of our primary or secondary outcomes until 9 months' follow-up suggests

that our approach was correct, and that brief integrated collaborative care within a specialist mental health-care context is unlikely to be either clinically efficacious or cost-effective.

Our study had limitations. We recruited a larger sample size than stated in our published protocol[9] because the number of patients in our follow-up at 12 months was lower than anticipated. Nevertheless, our study might have been underpowered to detect statistical improvement in symptoms and function at 12 months' follow-up. Patients were more severely ill and had been ill for longer than we had expected, thus requiring longer durations of SDS treatment and a graduated transfer to aftercare with clear recommendations for ongoing care that extended after the 12 months' primary outcome endpoint. As a result, and in line with NICE recommendations for research into service interventions for depression,[6] we reported outcomes at 18 months' and 12 months' follow up. We lost many patients during follow-up in this study (e.g., 72% at 12 months), which also occurred in other large, multicentre, secondary care RCTs for chronic depression—e.g., 70% of patients were lost to follow-up at 26 weeks in the REVAMP RCT.[42] An important limitation of both the clinical and economic outcomes is therefore the loss of participants to follow-up between 12 months and 18 months, which reduces the certainty of our results. However, in our study, imputed and original clinical outcomes were very similar. We did a multicentre study in the UK, which increases the generalisability of the results to the UK. Although recruitment was greater at one site than the other two, results were similar at all sites and no centre effects of treatment on outcome were reported. Another important limitation of our study was that full masking proved impossible to achieve. However, interview and self-report measures of outcome showed a similar pattern of results.

In accordance with the usual criteria for cost-effectiveness used by NICE,[39] our results at 18 months did not show that SDS is more cost-effective than TAU. Most of the incremental cost of SDS care is spent in the first 12 months, so this intervention might become more cost-effective after 2–3 years if patients receiving SDS treatment continue to improve quicker than do patients receiving TAU without additional cost. The incremental cost of SDS compared with TAU was £3446 per patient to improve patients with complex, moderately severe depression with a mean duration of more than 16 years to mild impairment in depression symptoms[34] and function[12] by 18 months. By contrast, cheaper, optimised, standard specialist mental health care, which might be better than routine usual care, could only improve patients' condition to a level of moderate impairment in depression symptoms[34] and function.[12]

Future research should attempt to replicate this study but include a larger sample size based on the effect sizes we have shown, and have a longer follow-up in view of the delayed treatment effect. Moreover, research could explore whether group psychological therapies might be of the same or more effective than individual psychotherapies at a reduced cost. Additional occupational support in selected patients might improve functional outcomes[7] and could be explored as an additional component of SDS treatment for depression.

Contributors

All authors wrote the manuscript. RMor was the chief investigator, obtained funding, designed the study, and interpreted the data. AG obtained funding, designed the study, supervised and delivered treatment, and interpreted the data. NN obtained funding, designed the study, delivered treatment, and analysed and interpreted the data. BG and CS analysed and interpreted the data. MJ designed the study and analysed and interpreted the data. CK-H was the trial manager, supervised the research assessments, and analysed the data. RMoo and RR obtained funding, designed the study, delivered treatment, and interpreted the data. TS designed the study, delivered treatment, and interpreted the data. TD obtained funding and interpreted the data.

Declaration of interests

We declare no competing interests.

Acknowledgments

Members of the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care (CLAHRC) Specialist Mood Disorder Study Group who made important contributions to the study were Ayesha Alrumaithi, Vijender Balain, Angie Balwako, Marcus Barker, Michelle Birkenhead, Paula Brown, Brendan Butler, Jo Burton, Isobel Chadwick, Adele Cresswell, Jo Dilks, Paige Duckworth, Heather Flambert, Richard Fox, Paul Gilbert, Emily Hammond, Joy Hodgkinson, Gail Hopkins, Valentina Lazarevic, Jane Lowey, Ruth MacDonald, Sarah E M Larson, Julie McKeown, Inderpal Panasar, Mat Rawsthorne, Kathryn Reeveley, Jayne Simpson, Katie Simpson, Sandra Simpson, Kasha Siubka-Wood, Gemma Walker, Samm Watson, Shirley Woolley, Nicola Wright, Min Yang, Ian Young. We acknowledge the help and support of the NIHR Clinical Research Network Mental Health in the East Midlands and East of England, and the University of Nottingham for providing sponsorship and Clinical Trials Unit support for this trial. This study was funded by the NIHR CLAHRC East Midlands (previously CLAHRC Nottinghamshire, Derbyshire, and Lincolnshire, and CLAHRC Leicestershire, Northamptonshire, and Rutland); NIHR CLAHRC East of England (previously the NIHR CLAHRC Cambridgeshire and Peterborough); the UK Medical Research Council; Nottinghamshire Healthcare NHS Foundation Trust; and the Derbyshire Healthcare Foundation Trust. The views expressed by the authors do not necessarily reflect those of the NIHR, NHS, the UK Medical Research Council, or the Department of Health in England.

References

1 Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2015; 368: 743–800.

2 McCrone P, Dhanasiri S, Patel A, Knapp M, Lawton-Smith S. Paying the price: the cost of mental health care in England to 2026. London: King's Fund, 2008.

3 Mueller T, Leon A, Keller M, et al. Recurrence after recovery from major depressive disorder during 15 years of observational follow-up. Am J Psychiatry 1999; 156: 1000–06.

4 Arnow BA, Constantino J. Effectiveness of psychotherapy and combination treatment for chronic depression. J Clin Psychol 2003; 59: 893–905.

5 Rush A, Trivedi M, Wisniewski S, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. Am J Psychiatry 2006; 163: 1905–17.

6 NICE. Depression: the treatment and management of depression in adults. National Clinical Practice Guideline 90. London: National Institute for Health and Care Excellence, 2009.

7 Murray G, Michalak E, Axler A, et al. Relief of chronic or resistant depression (Re-ChORD): a pragmatic, randomized, open-treatment trial of an integrative program intervention for chronic depression. J Affect Disord 2010; 123: 243–48.

8 Cuijpers P, van Staten A, Schuurmans J, van Oppen P, Hollon SD, Andersson G. Psychotherapy for chronic depression and dysthymia: a meta-analysis. Clin Psychol Rev 2010; 30: 51–62.

9 Morriss R, Marttunnen S, Garland A, et al. Randomised controlled trial of the clinical and cost effectiveness of a specialist team for managing refractory unipolar depressive disorder. BMC Psychiatry 2010; 10: 1–11.

10 First MB, Gibbon M, Spitzer RL, Williams JBW, Benjamin LS. Structured clinical interview for DSM-IV Axis I, (SCID-I). Washington, DC: American Psychiatric Press Inc, 1997.

11 Williams JBW, Kobak KA, Bech P, et al. The GRID-HAMD: standardization of the Hamilton Depression Rating Scale. Int Clin Psychopharmacol 2008; 23: 120–29.

12 American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 4th edn. Washington, DC: American Psychiatric Association, 1994.

13 Guo B, Kaylor-Hughes C, Morriss R, James M, Sampson C. Randomised controlled trial of the clinical and cost effectiveness of a specialist mood disorders team for refractory unipolar depressive disorder: statistics and economic plan. 2014. https://figshare.com/articles/SAP_mood_disorder_clahrc_feb_8_docx/2076910 (accessed July 20, 2016).

14 Moore R, Garland A. Cognitive therapy for chronic and recurrent depression. Chichester: Wiley, 2003.

15 von Wolff A, Hölzel LP, Westphal A, Härter M, Kriston L. Combination of pharmacotherapy and psychotherapy in the treatment of chronic depression: a systematic review and meta-analysis. BMC Psychiatry 2012; 12: 61.

16 Kuyken W, Hayes R, Barrett B, et al. Effectiveness and cost-effectiveness of mindfulness-based cognitive therapy comparedwith maintenance antidepressant treatment in the prevention of

depressive relapse or recurrence (PREVENT): a randomised controlled trial. Lancet 2015; 386: 63–73.

17 Gilbert P, Procter S. Compassionate Mind Training for people with high shame and self-criticism: overview and pilot study of a group therapy approach. Clin Psychol Psychother 2006; 13: 353–79.

18 Gunn J, Diggens J, Hegarty K, Blashki G. A systematic review of complex system interventions designed to increase recovery from depression in primary care. BMC Health Serv Res 2006; 6: 1–11.

19 Beck AT, Ward CH, Mendelson M, Mock JE, Erbaugh JK. An inventory for measuring depression. Arch Gen Psychiatry 1961; 4: 561–71.

20 Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: the validity of a brief depression severity measure. J Gen Intern Med 2001; 16: 606–13.

21 Rush AJ, Trivedi MH, Ibrahim HM, et al. The 16-item Quick Inventory of Depressive Symptomatology (QIDS) Clinician Rating (QIDS-C) and Self-Report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol Psychiatry 2003; 54: 573–83.

22 Cooper P, Osborn M, Gath D, Feggetter G. Evaluation of a modified self-report measure of social adjustment. Br J Psychiatry 1982; 141: 68–75.

23 Euroqol Group. EuroQol: a new facility for the measurement of health related quality of life. Health Policy 1990; 16: 199–208.

24 Measurement and Valuation of Health Group. The measurement and valuation of health: final report on the modelling of valuation tariffs. York: University of York, 1995.

25 Van der Feltz-Cornelis CM, Van Oppen P, Van Marwijk HWJ, De Beurs E, Van Dyck R. A patient-doctor relationship questionnaire (PDRQ-9) in primary care: development and psychometric evaluation. Gen Hosp Psychiatry 2004; 26: 115–20.

26 Guthrie E, Moorey J, Margison F, et al. Cost-effectiveness of brief psychodynamic-interpersonal therapy in high utilizers of psychiatric services. Arch Gen Psychiatry 1999; 56: 519–26.

27 Gulliford MC, Adams G, Ukoumunne OC, Latinovic R, Chinn S, Campbell MJ. Intraclass correlation coefficient and outcome prevalence are associated in clustered binary data. J Clin Epidemiol

2005; 58: 246–51.

28 Schramm E, Schneider D, Zobel I. Efficacy of interpersonal psychotherapy plus pharmacotherapy in chronically depressed inpatients. J Affect Disord 2008; 109: 67–73

29 Carpenter J, Kenward MG. Multiple imputation and its application. Chichester: John Wiley, 2013.

30 UK Department of Health. NHS staff earnings estimates to April 2014. London: Health and Social Care Information Centre, 2014.

31 Beecham J, Knapp M. Costing psychiatric interventions. In: Thornicroft G, ed. Measuring mental health needs. London: Gaskell, 2001.

32 Curtis L. Unit costs of health and social care. 2014. http://www. pssru.ac.uk/project-pages/unit-costs/2014/ (accessed July 20, 2016).

33 UK Department of Health. NHS reference costs 2013 to 2014. 2014. https://www.gov.uk/government/publications/nhs-referencecosts- 2013-to-2014 (accessed July 20, 2016).

34 British Medical Association, Royal Pharmaceutical Society. British National Formulary. http://www.medicinescomplete.com/mc/bnf/current/factsheet.pdf (accessed July 20, 2016).

35 Hoch JS, Briggs AH, Willan AR. Something old, something new, something borrowed, something blue: a framework for themarriage of health econometrics and cost-effectiveness analysis. Health Econ 2002; 11: 415–30.

36 Zimmerman M, Martinez JH, Young D, Chelminski I, Dalrymple K. Severity classification on the Hamilton Depression Rating Scale. J Affect Disord 2013; 150: 384–88.

37 NICE. Depression: management of depression in primary and secondary care. London: National Institute for Clinical Excellence, British Psychological Society, Gaskell, 2004.

38 Löwe B, Unützer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the Patient Health Questionnaire-9. Med Care 2004; 42: 1194–201.

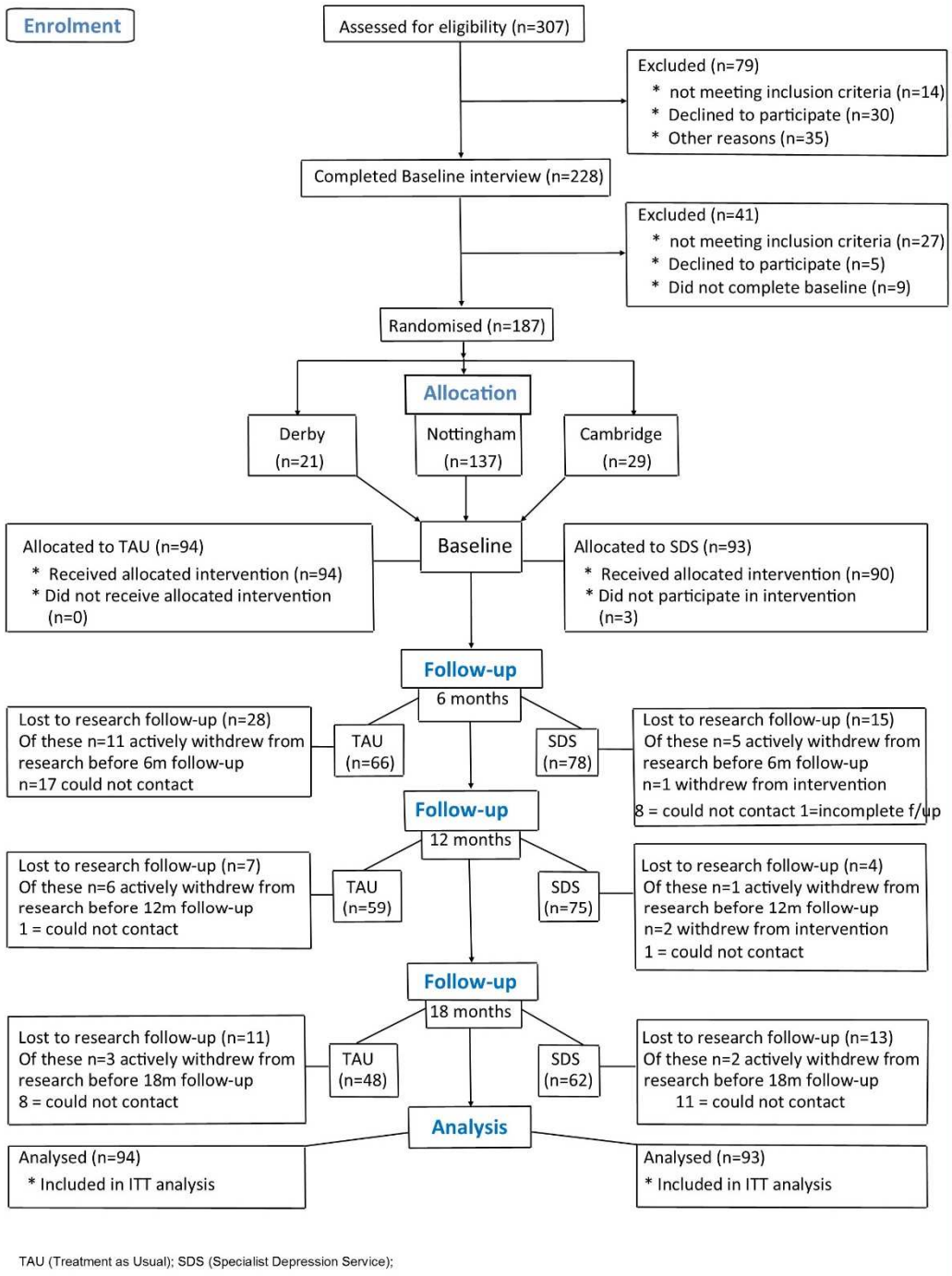
39 NICE. Guide to the methods of technology appraisal 2013. London: National Institute for Clinical Excellence, 2013.

40 Wiles N, Thomas L, Abel A, et al. Cognitive behavioural therapy as an adjunct to pharmacotherapy for primary care based patients with treatment resistant depression: results of the CoBalT randomised controlled trial. Lancet 2013; 381: 375–84.

41 Wiles NJ, Thomas L, Turner N, et al. Long-term effectiveness and cost-effectiveness of cognitive behavioural therapy as an adjunct to pharmacotherapy for treatment-resistant depression in primary care: follow-up of the CoBalT randomised controlled trial.Lancet Psychiatry 2016; 3: 137–44.

42 Kocsis J, Gelenberg A, Rothbaum B, et al. Cognitive behavioural analysis system of psychotherapy and brief supportive psychotherapy for augmentation of antidepressant nonresponse in chronic depression: the REVAMP trial. Arch Gen Psychiatry 2009; 66: 1178–88.

Figure 1. Flow of participants into the study



TAU (Treatment as Usual); SDS (Specialist Depression Service);

Table 1. Baseline Demographic and Clinical Characteristics of Treatment as Usual versus Specialist Depression Service.

	Treatment as Usual **(n = 94)**	**SDS** **(n=93)**
Age, mean (sd, range)	46 (11.3, 20-71)	47 (11.6, 20-84)
Gender, female, n (%)	60 (64)	54 (58)
Employment status, n (%)	[n=91]	[n=90]
full-time employment	22 (26)	17 (19)
Other employment[1]	11 (12)	10 (11)
Retired	10 (11)	16 (18)
Unemployed	37 (41)	36 (40)
Receipt of benefits: n (%)	[n=90] 63 (70)	[n=91] 61 (67)
Education, n (%)	[n=94]	[n=93]
before 16	7 (7)	3 (3)
up to 18 or apprenticeship	38 (40)	43 (46)
Highest qualification - Advanced levels	22 (23)	18 (19)
Highest qualification - degree or post-degree	27 (29)	29 (31)
Married or co-habiting, n (%)	50 (53)	42 (45)
Children, 1 or more, n (%)	58 (62)	61 (66)
Baseline HDRS$_{17,}$ mean (sd, range)	23.2 (5.8, 16-40)	22.0 (4.5, 16-33)
Baseline GAF, mean (sd, range)	47.7 (9.4, 21-61)	49.3 (6.8, 31-65)
Baseline BDI-I, mean (sd, range)	35.6 (9.0, 16-55)	35.9 (8.8, 11-56)
Baseline PHQ9, mean (sd, range)	19.3 (5.1, 5-27)	19.9 (4.6, 7-27)
Baseline QIDS-SR, mean (sd, range)	27.4,(7.3, 10-48)	27.6 (7.1,11-41)
SAS-M, mean (sd, range)	2.1 (0.6, 0.3- 3.6)	2.0 (0.7, 0.4-3.4)
PDRQ, mean (sd, range)	64.0 (16.8, 26-90)	60.9 (16.7, 24-90)
Baseline EQ-5D-3L index score, mean (range)	[n=85] 0.337 (-0.349-0.848)	[n=90] 0.361 (-0.239-0.848)
Years since first diagnosis of depression mean (sd, range)	16.9 (11.6, 0.5-49.0)	16.5 (11.1, 0.5 -49.0)
Depressed > 1 year, n (%)	82 (87)	80 (86)
Years since first diagnosis of current depression episode, median (IQR)	5.7 (2.1-20.1)	7.3 (2.6-16.0)

SDS = Specialist Depression Service

[1] Other employment includes part-time, sheltered and voluntary employment and higher education

$HDRS_{17}$ = 17 item Hamilton Depression Rating Scale

GAF = Global Assessment of Function

BDI-I = Beck Depression Inventory-I

PHQ-9 = 9 item Personal Health Questionnaire

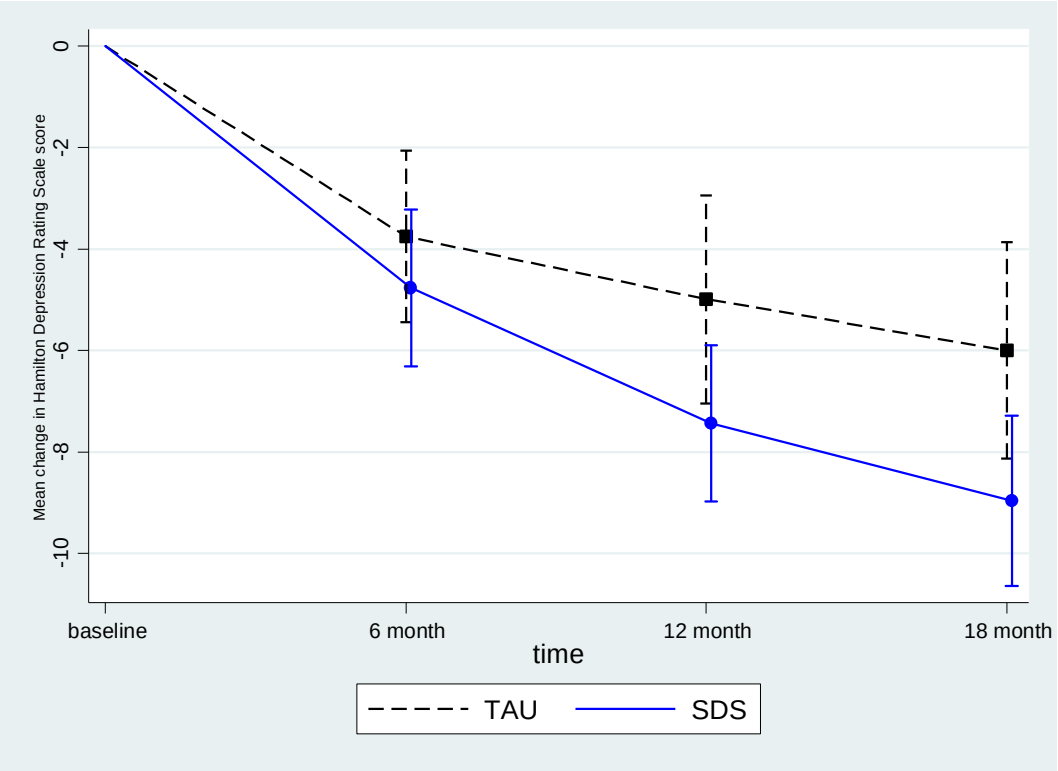
QIDS-SR = Quick Inventory for Depression Scale Self-Rated

SAS-M = Modified Social Adjustment Scale

PDRQ = Patient Doctor Relationship Questionnaire

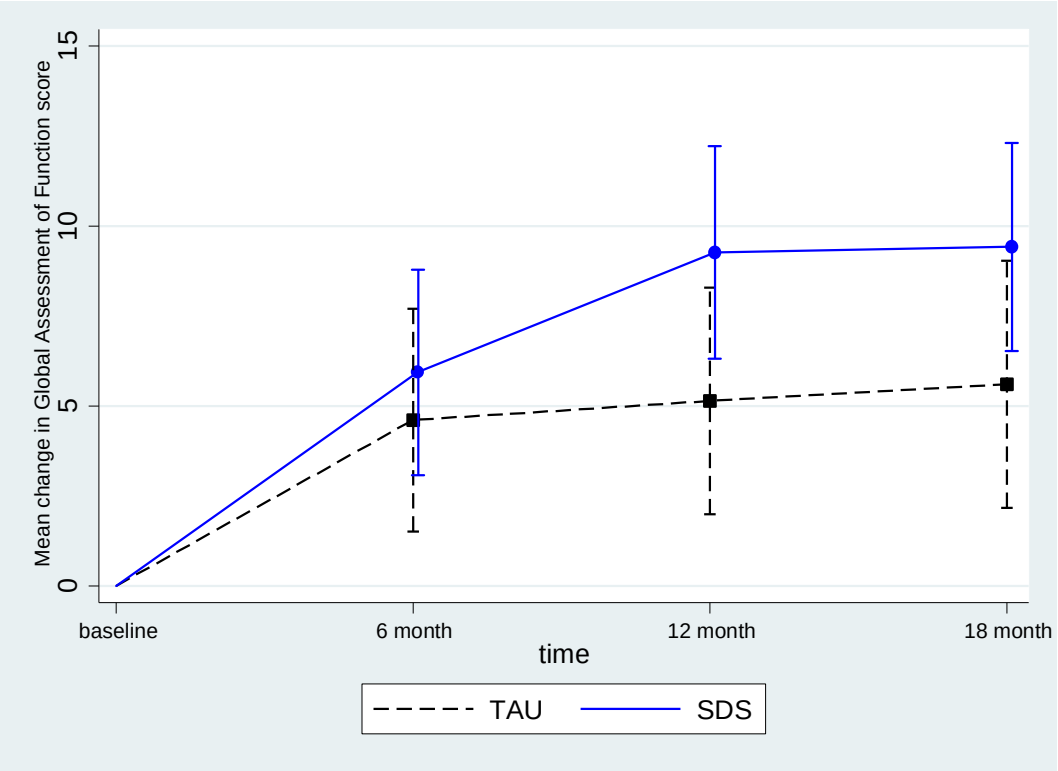
EQ-5D-3L

Figure 2: Mean change in 17 item Hamilton Depression Rating Scale score (95%CI) by group, intention-to-treat analysis (n=187) across time, Treatment as Usual versus Specialist Depression Service.



TAU = treatment as usual (n=94), SDS = specialist depression service (n=93)

Figure 3: Mean change in Global Assessment of Function (95% CI) by group intention to treat analysis (n=187) across time, Treatment as Usual versus Specialist Depression Service.



TAU = treatment as usual (n=94), SDS = specialist depression service (n=93)

Table 2. Results of multilevel modelling of change score (95% CI) from baseline and group difference (95% CI) Treatment as Usual versus Specialist Depression Service.

	TAU (n=94) Mean change from Baseline (95% CI)	**SDS (n=93)** Mean change from baseline (95% CI)	**Group comparison** Change difference (95%CI)	P value
$HDRS_{17}$				
6 month	-3.76(-5.45, -2.07)	-4.77(-6.32, -3.22)	-1.01(-3.30, 1.28)	0.385
12 month	-4.99(-7.04, -2.94)	-7.44(-8.98, -5.90)	-2.45(-5.04, 0.14)	0.064
18 month	-6.00(-8.13, -3.87)	-8.96(-10.64, -7.28)	-2.96(-5.33, -0.59)	0.015
GAF				
6 month	4.61(1.51, 7.70)	5.93(3.08, 8.79)	1.33(-2.92, 5.57)	0.538
12 month	5.14(1.99, 8.28)	9.26(6.31, 12.21)	4.12(-0.11, 8.35)	0.056
18 month	5.60(2.17, 9.03)	9.42(6.53, 12.31)	3.82(-0.93, 8.57)	0.113
PHQ-9				
3 month	-0.91(-2.63, 0.81)	-2.66(-4.03, -1.30)	-1.75(-3.84, 0.33)	0.099
6 month	-3.56(-5.18, -1.94)	-4.71(-6.35, -3.07)	-1.15(-3.21, 0.90)	0.269
9 month	-3.15(-4.55, -1.75)	-6.61(-8.20, -5.03)	-3.46(-5.69, -1.23)	0.003
12 month	-3.89(-5.53, -2.26)	-6.88(-8.27, -5.50)	-2.99(-4.97, -1.01)	0.003
18 month	-3.32(-5.58, -1.05)	-7.57(-9.32, -5.81)	-4.25(-6.94, -1.56)	0.003
BDI-1				
3 month	-3.32(-5.67, -0.97)	-5.17(-7.74, -2.60)	-1.85(-5.36, 1.66)	0.299
6 month	-4.28(-7.09, -1.46)	-6.97(-9.52, -4.42)	-2.69(-6.77, 1.38)	0.191
9 month	-5.98(-9.06, -2.91)	-11.00(-13.35, -8.66)	-5.02(-8.81, -1.24)	0.010
12 month	-6.13(-8.50, -3.75)	-11.51(-14.07, -8.95)	-5.39(-8.88, -1.89)	0.003
18 month	-7.48(-10.64, -4.31)	-12.23(-14.77, -9.69)	-4.75(-8.83, -0.68)	0.023
QIDS-SR				
3 month	-1.40(-3.52, 0.72)	-2.47(-4.31, -0.62)	-1.06(-3.88, 1.75)	0.455
6 month	-3.17(-5.31, -1.03)	-5.71(-7.67, -3.75)	-2.54(-5.41, 0.33)	0.082
9 month	-4.44(-6.94, -1.95)	-7.85(-9.74, -5.96)	-3.41(-6.71, -0.10)	0.044
12 month	-4.92(-7.40, -2.44)	-7.81(-9.55, -6.06)	-2.89(-5.89, 0.12)	0.059
18 month	-5.33(-8.10, -2.56)	-9.30(-11.25, -7.36)	-3.97(-7.60, -0.35)	0.033
SAS-M				
6 month	0.09(-0.05, 0.22)	0.16(0.04, 0.28)	0.07(-0.11, 0.26)	0.431
12 month	0.20(0.05, 0.34)	0.24(0.12, 0.37)	0.04(-0.16, 0.25)	0.669
18 month	0.08(-0.09, 0.26)	0.26(0.12, 0.40)	0.18(-0.02, 0.38)	0.080
PDRQ				
6 month	0.00 (-5.22, 5.21)	6.88(2.21, 11.55)	6.88(-0.18, 13.95)	0.056
12 month	-3.37 (-8.87, 2.14)	6.78(1.76, 11.81)	10.15(2.65, 17.65)	0.008
18 month	-1.14 (-7.62, 5.35)	2.89(-3.14, 8.92)	4.03(-4.89, 12.94)	0.376

TAU = treatment as usual, SDS = specialist depression service

$HDRS_{17}$ = 17 item Hamilton Depression Rating Scale

GAF= Global Assessment of Function

PHQ-9 = 9 item Personal Health Questionnaire

BDI-1 = Beck Depression Inventory-I

QIDS-SR= Quick Inventory for Depression Scale

SAS-M = Modified Social Adjustment Scale

PDRQ = Patient Doctor Relationship Questionnaire

Table 3. Total and incremental costs Quality-Adjusted Life Years (QALYS) for Treatment as Usual versus Specialist Depression Service.

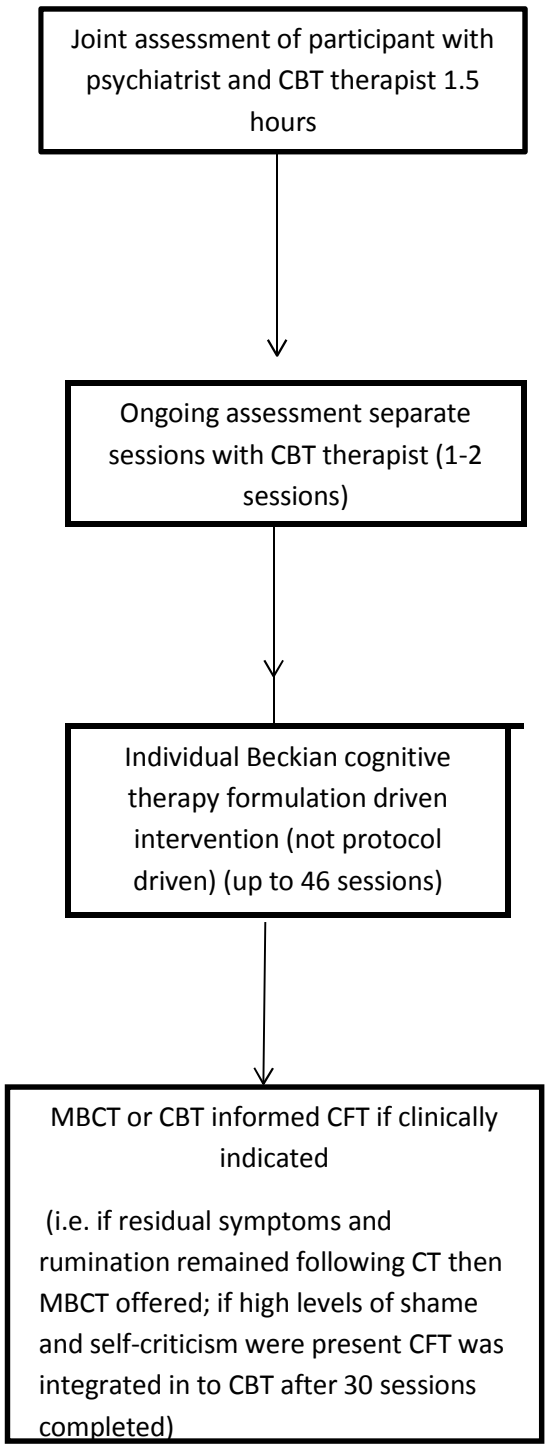
	TAU (95% CI)	SDS (95% CI)	Incremental (Bias-corrected bootstrapped 95% CI)
Costs	£5869 (£4501, £7238)	£9315 (£7547, £11084)	£3446 (£1915, £5180)
QALYs	0.646 (0.538, 0.754)	0.753 (0.659, 0.847)	0.079* (0.007, 0.149)

TAU = treatment as usual, SDS = specialist depression service

*Incremental QALYs controls for baseline EQ-5D-3L index score by linear regression

Appendix

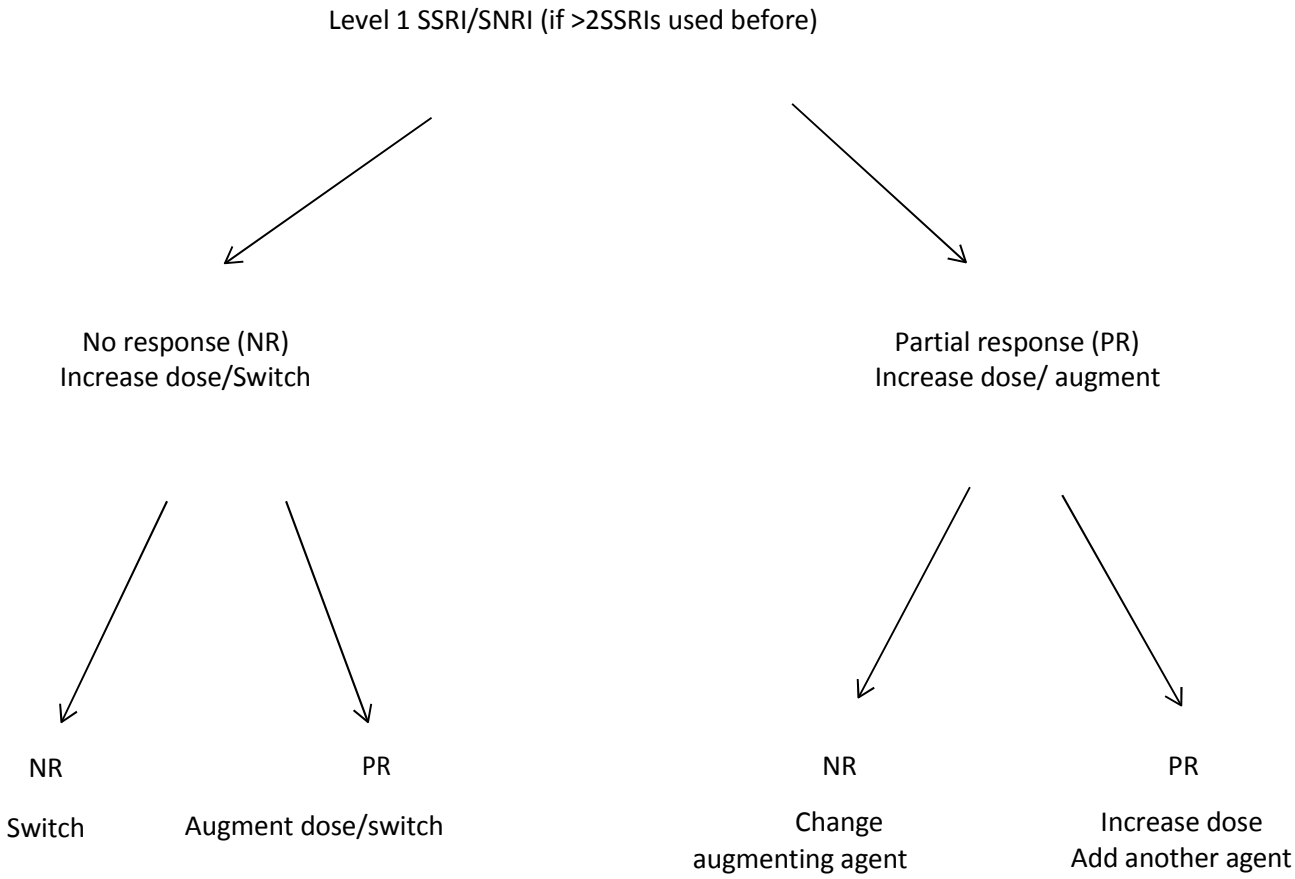
Figure A1: Psychological Treatment Pathway in Specialist Depression Service



CBT = cognitive behaviour therapy, MBCT= mindfulness based cognitive therapy, CFT = compassion focussed therapy.

Joint follow up reviews between psychiatrist and CBT therapist at 3,6, 9 and 12 months. In Derby group treatment offered after initial individual treatment. At Cambridge assessment meeting not held jointly with participant but each participant discussed jointly by psychiatrists and psychological therapist soon after individual assessment.

Figure A2: Guide to pharmacotherapy used in Specialist Depression Service



Change and augmentation agents outlined in main text of paper.

For bipolar depression first line treatment was lamotrigine or quetiapine. For severe bipolar depression other classes of antidepressants (with the exception of tricyclics) were used along with mood stabilisers. Principles of treatment with regard to switching and augmentation remained the same.

Table A1. Additional clinical characteristics of sample at baseline (n=187).

Characteristic	Number (%)
Current unipolar major depressive episode[1]	179 (95.7)
Current bipolar 2 major depressive episode	8 (4.3)
Past major depressive episode	156 (83.4)
Current melancholia	105 (56.1)
Current psychotic symptoms (delusions and/or hallucinations)	49 (26.2)
With dysthymia ("double depression")	17 (9.1)
Any other comorbid anxiety, substance use or eating disorder	151 (80.3)
Substance use disorder (alcohol and/or drug abuse or dependence)	32 (17.1)
Eating disorder (anorexia nervosa, bulimia nervosa, binge eating disorder)	22 (11.8)
Anxiety disorder:	146 (78.1)
Panic disorder or agoraphobia	86 (46.2)
Generalised anxiety disorder	85 (45.7)
Simple phobia	48 (25.8)
Social phobia	44 (23.7)
Obsessive compulsive disorder	37 (19.9)
Post-traumatic stress disorder	30 (16.1)
Somatoform disorder (hypochondriasis or other somatoform disorder)	31 (16.6)
Current active physical illness:	120 (64.2)
One current active physical illness	77 (41.2)
Two or more active physical illnesses	25 (13.4)
Current rheumatological or orthopedic problem	43 (23.4)
Current cardiovascular disorder (including diabetes mellitus)	33 (17.1)
Current respiratory disorder	26 (13.5)
Current neurological disorder	18 (9.4)

[1] Referral to RCT made as a unipolar major depressive episode but using standardised psychiatric interview diagnosed as bipolar 2 major depressive episode.

Table A2. Results of observed data only sensitivity analysis of change score (95% CI) from baseline and group difference (95% CI) for primary outcomes and secondary depression outcomes

	TAU		SDS		Group comparison	P value
	n	Mean change from Baseline (95% CI)	n	Mean change from baseline (95% CI)	Change difference (95%CI)	
$HDRS_{17}$						
6 month	66	-3.59(-5.28, -1.90)	78	-4.45(-6.02, -2.89)	-0.86(-3.17, 1.44)	0.463
12 month	59	-4.78(-6.55, -3.02)	75	-6.98(-8.56, -5.39)	-2.19(-4.57, 0.18)	0.07
18 month	48	-5.94(-7.86, -4.03)	62	-8.54(-10.25, -6.83)	-2.60(-5.16, -0.03)	0.047
GAF						
6 month	64	6.31(3.44, 9.18)	78	6.09(3.52, 8.66)	-0.22(-4.07, 3.64)	0.912
12 month	57	7.74(4.72, 10.76)	75	11.21(8.60, 13.83)	3.47(-0.52, 7.47)	0.088
18 month	49	9.04(5.81, 12.26)	67	12.08(9.24, 14.93)	3.05(-1.25, 7.35)	0.165
PHQ-9						
3 month	63	-0.85(-2.36, 0.66)	78	-2.62(-3.98, -1.26)	-1.77(-3.81, 0.26)	0.088
6 month	54	-3.57(-5.16, -1.98)	63	-4.83(-6.29, -3.37)	-1.26(-3.42, 0.90)	0.253
9 month	46	-3.22(-4.92, -1.52)	67	-6.70(-8.13, -5.27)	-3.48(-5.70, -1.26)	0.002
12 month	45	-4.10(-5.79, -2.41)	57	-7.04(-8.55, -5.52)	-2.94(-5.21, -0.67)	0.011
18 month	40	-3.49(-5.25, -1.72)	51	-7.86(-9.45, -6.28)	-4.38(-6.75, -2.01)	<0.001
BDI-1						
3 month	66	-2.95(-5.47, -0.42)	78	-4.98(-7.34, -2.62)	-2.03(-5.49, 1.42)	0.249
6 month	57	-4.26(-6.92, -1.60)	67	-6.85(-9.32, -4.38)	-2.59(-6.22, 1.04)	0.162
9 month	49	-5.52(-8.31, -2.73)	64	-10.72(-13.22, -8.21)	-5.20(-8.95, -1.45)	0.007
12 month	48	-6.26(-9.07, -3.44)	61	-11.40(-13.96, -8.84)	-5.14(-8.95, -1.34)	0.008
18 month	42	-7.46(-10.40, -4.51)	54	-12.58(-15.25, -9.92)	-5.12(-9.10, -1.15)	0.011
QIDS-SR						
3 month	64	-1.35(-3.26, 0.56)	78	-2.22(-3.98, -0.47)	-0.87(-3.46, 1.73)	0.511
6 month	57	-3.38(-5.38, -1.39)	67	-5.86(-7.71, -4.02)	-2.48(-5.20, 0.24)	0.074
9 month	48	-4.61(-6.73, -2.49)	64	-7.70(-9.58, -5.83)	-3.09(-5.92, -0.26)	0.032
12 month	48	-4.92(-7.04, -2.80)	62	-7.63(-9.54, -5.73)	-2.71(-5.57, 0.14)	0.063
18 month	38	-5.16(-7.47, -2.85)	51	-9.35(-11.39, -7.30)	-4.19(-7.28, -1.10)	0.008

TAU = treatment as usual, SDS = specialist depression service

$HDRS_{17}$ = 17 item Hamilton Depression Rating Scale

GAF= Global Assessment of Function

$HDRS_6$ = 6 item Hamilton Depression Rating Scale

PHQ-9 = 9 item Personal Health Questionnaire

BDI-1 = Beck Depression Inventory-I

QIDS-SR= Quick Inventory for Depression Scale

Table A3. Service use at baseline, 6, 12 and 18 months for TAU and SDS groups.

	TAU					SDS				
	n	**Mean**	**Sd**	**Min**	**Max**	**n**	**Mean**	**sd**	**Min**	**Max**
Baseline										
Total hospital inpatient days	91	0.187	0.942	0	7	90	1.556	10.025	0	91
Psychiatric outpatient attendances	91	1.615	1.504	0	8	91	2.253	6.276	0	60
Other outpatient and A&E attendances	91	1.341	2.837	0	14	89	0.640	1.618	0	10
GP surgery / home attendances	91	3.484	5.555	0	40	91	3.846	3.703	0	24
Practice / District / community psychiatric nurse	91	1.451	2.861	0	12	91	3.560	11.286	0	75
Psychotherapist	Not collected									
3 month										
Total hospital inpatient days	58	1.914	9.643	0	66	60	0.033	0.258	0	2
Psychiatric outpatient attendances	58	1.966	1.685	0	6	60	5.233	6.207	0	36
Other outpatient and A&E attendances	58	2.328	8.134	0	50	60	2.000	5.499	0	30
GP surgery / home attendances	57	1.298	1.742	0	8	61	2.361	2.921	0	12
Practice / District / community psychiatric nurse	58	1.483	4.185	0	24	61	2.836	4.495	0	18
Psychotherapist	56	0	0	0	0	60	.872	2.732	0	13
6 month										
Total hospital inpatient days	64	0.297	1.136	0	6	74	0.966	5.321	0	35
Psychiatric outpatient	64	1.984	2.622	0	16	76	4.355	3.874	0	20

attendances										
Other outpatient and A&E attendances	64	0.703	1.933	0	12	76	0.395	1.609	0	12
GP surgery / home attendances	64	1.844	2.176	0	10	76	2.303	2.713	0	18
Practice / District / community psychiatric nurse	64	0.797	2.125	0	12	75	2.413	4.268	0	20
Psychotherapist	64	1.344	2.907	0	12	75	7.507	4.512	0	18
9 month										
Total hospital inpatient days	40	0.350	2.214	0	14	55	0.455	2.768	0	20
Psychiatric outpatient attendances	40	1.65	1.777	0	7	56	3.768	5.194	0	18
Other outpatient and A&E attendances	40	0.35	1.122	0	6	55	2.800	6.346	0	36
GP surgery / home attendances	40	1.15	1.889	0	7	56	1.679	2.133	0	11
Practice / District / community psychiatric nurse	40	.55	1.280	0	6	56	1.5	2.886	0	10
Psychotherapist	40	.613	1.961	0	10	54	2.341	4.503	0	12
12 month										
Total hospital inpatient days	55	0	0	0	0	69	3.174	15.438	0	97
Psychiatric outpatient attendances	56	1.393	1.702	0	11	63	2.540	2.729	0	13
Other outpatient and A&E attendances	55	0.527	2.008	0	13	63	.984	3.019	0	18
GP surgery / home attendances	55	1.636	2.422	0	12	62	1.613	2.250	0	12
Practice / District / community psychiatric nurse	56	.643	1.151	0	5	63	.905	2.123	0	10

Psychotherapist	55	1.691	3.532	0	13	62	6.355	4.192	0	12
18 month										
Total hospital inpatient days	48	0	0	0	0	60	0.367	2.712	0	21
Psychiatric outpatient attendances	48	1.229	1.259	0	6	60	1.033	1.461	0	5
Other outpatient and A&E attendances	48	0.146	.583	0	3	59	0.525	2.409	0	14
GP surgery / home attendances	48	1.104	1.491	0	6	60	1.467	1.944	0	10
Practice / District / community psychiatric nurse	48	.917	2.491	0	12	60	.917	3.185	0	23
Psychotherapist	48	.762	2.418	0	14	59	1.255	2.660	0	12

Table A4: Complete case cost cost-effectiveness analysis at 6, 12 and 18 month follow-up

Assessment	EQ-5D-3L index score [n] Mean (s.d.)		Self-reported NHS service use costs [n] Mean (s.d.)		ICER for complete cases [n] Mean (Bias-corrected bootstrapped 95% CI)
	TAU	SDS	TAU	SDS	SDS vs TAU
Baseline	[85] 0.337 (0.343)	[90] 0.361 (0.329)	[90] £1410 (1970)	[89] £1898 (4227)	-
3 month	-	-	[56] £1742 (3405)	[60] £2194 (2389)	-
6 month	[56] 0.489 (0.360)	[64] 0.500 (0.360)	[58] £1088 (1446)	[70] £1839 (2006)	[51] £2699 (-£159271, £619730)
9 month	-	-	[40] £803 (940)	[54] £1900 (2118)	-
12 month	[49] 0.399 (0.390)	[62] 0.536 (0.349)	[49] £995 (1481)	[54] £1744 (3067)	[31] *SDS dominates* (-£7914105, £43288)
18 month	[42] 0.423 (0.357)	[51] 0.583 (0.372)	[45] £639 (865)	[54] £742 (1578)	[24] *SDS dominates* (-£277242, £208765)

Figure A3 Cost-effectiveness acceptability curve of TAU versus SDS over 18 months