Dialectical Behaviour Therapy (DBT) for Treatment-Resistant Depression (TRD): A Randomised Controlled Trial (RCT)

Treatment-resistant depression (TRD) is a major public health problem that, like other chronic conditions, substantially reduces health and well being (1). The World Health Organisation predicts that by 2020 depression will be the second most frequent cause of disability worldwide (2). In the UK the estimated cost of mood disorders was £25 billion in 2008 (some 1.5% of GDP) (3). The immense contribution of major depressive disorder (MDD) to this burden is due to its highly recurrent nature (2). In the absence of prophylactic treatment the rate of recurrence rises to about 80% (4), and chronic depression is harder to treat with both antidepressant medication (ADM) and psychotherapy (5–8). Given that most reports suggest only 30% to 40% of individuals treated with ADM achieve full remission, treatment resistance may be the most common outcome for individuals with MDD (9).

Important differences between acute, chronic and treatment-resistant forms of unipolar depression are still emerging. Broadly speaking, TRD is depression that does not respond to adequate intervention, whereas chronic depression lasts more than 2 years. TRD and chronic depression may therefore overlap, with many patients meeting both definitions; yet both reflect depression that is *unresponsive*. So our proposed trial focuses on TRD on the understanding that many patients will have comorbid chronic depression.

Risks for developing chronic depression include childhood adversity, environmental stress, and heightened stress reactivity (10). An estimated 40–60% of unipolar depressed patients meet criteria for comorbid personality disorder (PD), with even higher rates among those with chronic or TRD (e.g. 11–13). In common with Klein et al (13), data from our Dorset site show that more than 60% of TRD patients have some form of PD. The most common PDs among TRD individuals are Cluster-A (paranoid PD) and Cluster-C (obsessive-compulsive and avoidant PD) (8; 11; 14). In patients with long-standing depressive symptoms, Cluster-C personality disorders were the most predictive of chronic depression at follow-up (15). Thus TRD and chronic depression are prevalent, burdensome to sufferers, and hard to treat, yet understudied and poorly understood relative to acute depression (16).

Limitations of current research and treatments for TRD and chronic depression
There are 3 linked problems with existing research on TRD and chronic depression. First, research into the treatment of TRD is scarce. Relatively few interventions directly target TRD or chronic depression, and international registries reveal only 11 randomised controlled trials (RCTs) of psychotherapy for chronic depression. Most research has focused on pharmacological or somatic interventions, and a recent systematic review of RCTs of medication for TRD reported many conceptual and methodological problems (9). A recent review of psychotherapy for TRD (17) included only 4 RCTs among 12 studies; all of which with fewer than 25 participants, and thus lacked power to detect important effects.

Second, investigators have failed to adopt a consistent definition of TRD; both experimental and clinical studies vary widely in their interpretation of the concept. Many studies reported to include patients with TRD or chronic depression exclude patients who would usually be considered exemplars of either category. For example, most RCTs have excluded patients with comorbid personality disorder, suicidal behaviour, prior psychotherapy treatment, or frequent relapse (some with only 3 or more episodes). This limits the validity of current treatment research and means that most patients who would be characterised as treatment-resistant by GPs or psychiatrists are excluded from rigorous studies.

Third, most current treatments focus on acute unipolar depression and fail to account for the differences in the aetiology and persistence of TRD or chronic depression. One exception was the Cognitive Behavioural Analysis System of Psychotherapy (CBASP). However a recent large trial of CBASP showed that only 38% of participants experienced any response (18): although developed to treat chronic depression, CBASP was no better than brief supportive psychotherapy (BSP) and adding either psychotherapy (CBASP or BSP) to pharmacotherapy added nothing to pharmacotherapy alone. In contrast 4 months of group interpersonal psychotherapy plus medication plus occupational therapy (Re-ChORD; 19) compared with TAU demonstrated advantages for Re-ChORD in achieving remission. However the study did not collect enough data to study maintenance of gains, had significant drop-out, and was under-powered to investigate moderators of outcome. Finally, a large multi-site study designed to test switching and augmentation strategies showed no significant differences between approaches, with less than one third remitting following poor response to acute ADM treatment (including augmentation of ADM with cognitive therapy) (STAR*D 20; 21).

Thus, there are few promising candidates for the effective treatment of TRD or chronic forms of depression. The fact that existing trials in TRD patients have rather narrow inclusion criteria only aggravates the problem; patients with PD, for example, are known to respond less favourably to existing treatments of acute depression, such as cognitive behavioural therapy (8).

Accounting for poor outcomes in current psychosocial approaches to TRD and chronic depression We hypothesise that prior psychosocial therapies for TRD and chronic depression have been ineffective because they do not target features of PD. Personality disorder—particularly the emotionally-constricted Cluster-A & C PDs—are common in depressed patients, and can disrupt treatment (8). Developmental research shows that emotionally constricted, risk averse, and over-controlled children are more likely to develop into depressed and socially isolated adults (e.g. 22). Moreover, unresponsive depressed patients exhibit PD-like interpersonal difficulties; they pose greater challenges for therapists and are rated as more hostile and less 'friendly' than the acutely depressed (23). Finally, compared with non-chronic major depressive disorder (MDD), chronically depressed individuals show greater self-criticism, impaired autonomy, rigid internalised expectations, excessive control of spontaneous emotion, and inordinate fears of making mistakes

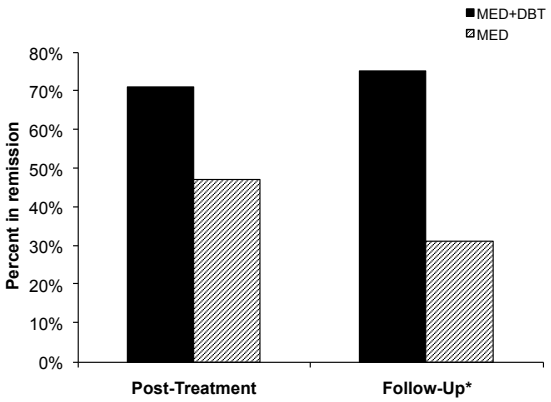


Figure 1: Percent of patients in remission (HAM-D $\leq$ 7) at post-treatment and at follow-up (Lynch et al 2003); black bars: DBT plus medication; grey bars: medication alone.

(10)—all maladaptive styles of coping characteristic of emotionally-constricted PDs.

These findings, and our own research, have led to the development of a theoretically derived and targeted therapy for TRD and chronic depression (24). Based on a biosocial theory for emotionally-constricted disorders (ECD) and TRD (24) we contend that individuals who develop treatment-resistant or chronic forms of depression are by nature highly sensitive to threat and insensitive to reward, have strong tendencies for constraint, and under stress prefer order and structure to novelty. These predispositions interact with a socio-biographic environment that values emotional control and avoiding mistakes. The individual acquires a coping style characterised by inhibited expression, risk avoidance, perfectionism, distress over-tolerance, and covert expression of hostility. This style of coping is intermittently negatively reinforced by reductions in arousal associated with avoidance of feared situations, and positively reinforced by achievement or performance. Unfortunately, rigid and over-controlled coping appears to result in poor interpersonal relationships and general difficulties with adapting to changing environmental circumstances, leading to depression and other related problems.

Our research has tested components of this theory: we have found that temperamental negative affectivity is linked with increased thought suppression and ambivalence towards emotional expression, which in turn lead to increased presence of hopelessness, depression and suicidal ideation (25–27). Furthermore, the presence of personality disorder and cognitions including guilt or sinfulness, contribute to the persistence or re-emergence of depressive symptoms and suicidal ideation (28; 29). We have also examined the role of biological vulnerabilities—such as reward insensitivity and risk aversion—which are key components of our model. Among depressed individuals we have found enhanced feedback-based decision-making and risk aversion using behavioural performance measures (30) and decreased activation during reward anticipation in the right caudate, supporting the hypothesis of hypo-responsivity in mesolimbic reward regions during reward anticipation (31).

We have also shown that obsessive-compulsive PD features (Cluster C-OCPD is the most common PD in TRD; 11), rather than depression, accounted for greater risk aversion in a depressed sample (32) suggesting that the lack of approach motivation common in TRD may be linked to an emotionally-constricted style of coping. Finally, we have verified the importance of social support and impaired dependency or autonomy in depression (33; 34), pointing to the importance of accounting for impaired interpersonal relationships when developing treatments for unresponsive depressed patients. Our theoretical approach is supported by a diverse literature: for example developmental research shows that over-control of emotion is likely to result in decreased social competence and internalising disorders (22); and experimental research suggests that suppressed expression may function as a socially contagious danger signal resulting in lowered social affiliation (35).

Proof of Concept: Dialectical Behaviour Therapy for TRD and chronic depression

Dialectical behaviour therapy (DBT) has proven efficacy in treating borderline personality disorder (36). In patients with BPD, among whom depression is common, trials have shown DBT to reduce depression, anxiety and suicidal behaviours (37). More recently DBT has been applied to TRD and chronic depression: 3 RCTs have piloted standard DBT as treatment for TRD or chronic depression (27; 38; 39), one specifically requiring TRD plus comorbid personality disorder (27).

Lynch (38) randomly assigned 34 chronically depressed individuals over 60 to either antidepressant medication alone or antidepressant medication plus a modified form of DBT. The main objective of this first study was to explore the feasibility of a group intervention for TRD. DBT treatment consisted of 28 weeks of a skills-training group, and weekly 30-minute phone contact with an individual therapist, followed by 3 months in which phone contact was every 2 weeks and 3 months in which it was every 3 weeks. Those receiving DBT showed significantly greater improvements than controls in self-rated and interviewer-rated depression. Post-treatment interviewer ratings showed that 71% of DBT recipients met criteria for remission, but only 47% of controls did so. After 6 months the corresponding percentages were 75% and 31% (Figure 1). DBT recipients had also improved significantly in adaptive coping and dependency, while controls did not.

The second RCT by Lynch (27) compared 24 weeks of both individual and group therapy plus ADM with ADM alone (both including clinical management by a study psychiatrist) in adults aged over 55 with personality disorder and comorbid depression. To be included participants had to demonstrate TRD prospectively via poor response to an 8 week course of researcher-controlled ADM. DBT recipients showed significantly greater decreases in interpersonal sensitivity and aggression than controls. At the end of the skills group 71% of DBT recipients were in remission, compared with only 50% of controls. Both groups showed significant reductions in clinician-rated depression; though the difference between groups was not significant, improvements were more rapid for DBT.

A third RCT, by Harley et al (39), used standard DBT group skills training to treat

Table 1: Elements of group DBT for TRD

Behaviours to increase	**Behaviours to decrease**
Core Mindfulness	Rigidity, Habitual responding
Radical Openness	Avoidance of risk, Emotion Inhibition
Interpersonal Effectiveness	Distrust, aloofness, Avoiding feedback
Emotion Regulation	Over-control, Envy and bitterness
Distress Tolerance (Self-soothing, acceptance)	Self-neglect, Rule-governance

major depressive disorder in adult outpatients for whom antidepressant medication had failed: 24 patients were randomly allocated to either group skills training or waiting list. DBT participants showed significantly greater improvements than controls in depressive symptoms. Feldman et al. (40) conducted secondary analyses showing that increases in emotional processing measured by the Emotional Approach Coping measure (EAC; 41) were associated with decreases in depressive symptoms in the DBT group, but with with increases in depression in controls, suggesting that DBT may facilitate adaptive processing of emotions.

Development of the DBT for TRD manual Our feasibility trials and supporting theoretical work resulted in an adaptation of standard DBT (24). Unlike standard DBT, developed primarily for use with dramatic-erratic, under-controlled and impulsive disorders (e.g. BPD; 42), our new approach targets common problems in TRD and chronic depression, including over-control, rigidity, interpersonal aloofness, emotion inhibition and perfectionism. Treatment is informed by a biosocial theory that posits a biological predisposition for heightened threat sensitivity and diminished reward sensitivity, coupled with early childhood invalidation or maltreatment, resulting in an over-controlled coping style that limits opportunities to learn new skills and exploit positive social reinforcers. In addition, our new approach capitalises on recent findings showing the bi-directional influence of the autonomic nervous system (43) by introducing new treatment approaches designed to alter neuroregulatory responses by directly activating its antagonistic system; in other words, to "turn off" defensive emotional arousal by activating the calming parasympathetic nervous system. Table 1 shows the targets of the new group skills sessions, and Table 2 compares features of the new treatment for TRD with standard DBT for BPD. DBT for TRD is ready for rigorous evaluation: it is a fully manualised psychosocial intervention, with a defined individual treatment rationale and skills training sessions that specifically target TRD coping deficits.

2.2 Risks and benefits

Efficacious treatment for TRD has remained elusive: the area is under-researched, and treatments for acute depression have not proven efficacious for TRD; our approach is novel because it targets features of PD in TRD. Identifying an efficacious treatment for TRD has great potential for patients, the NHS and the wider economy. However all

Table 2: Differences between standard DBT and DBT for TRD

Mode/Target/Aspect	Standard DBT	DBT for TRD
Primary Treatment Target	The primary orientation is to reduce severe behavioural undercontrol and emotional dysregulation.	The primary orientation is to reduce behavioural over-control, rigidity, and emotional constriction and increase flexibility, openness to new experience, and encourage expression of emotions.
Motivation to Change	Motivation to change is a critical component of treatment	Places greater importance on attachment strategies that are designed to enhance motivation to change; similar to DBT for substance abuse.
Skills Training	Standard DBT skills include mindfulness, emotion regulation, distress tolerance, and interpersonal effectiveness.	Most of the standard DBT skills plus a new "Radical Openness" module focusing on problems of TRD: e.g., openness to new experiences and critical feedback, letting go of suspicious and emotionally constricted behaviours, inhibiting automatic avoidance of novelty.
Targeting Arousal and Emotion Vulnerability Directly	Skills designed to influence emotional vulnerability (e.g., PLEASE Master skills).	New skills to activate Parasympathetic Nervous System (PNS) & Social Engagement System.
Overcoming bitterness/grievances, & forgiving self/others.	Radical Acceptance and Opposite Action to Anger	New Loving-Kindness Forgiveness Meditation protocol. Standard DBT opposite action skills plus new skills for opposite action to envy and bitterness.
Mindfulness States-of-Mind	Emotion Mind, Wise Mind, & Reasonable Mind to identify when emotionally dysregulated or impulsive behaviour likely	Fixed, Fluid & Passive Mind to identify when rigid-emotionally constricted behaviour is likely
Behavioural Activation	Used as needed in standard DBT	Skills to enhance playful behaviour and honest expressions of affect critical for cooperative relationships.

research carries responsibility to address potential risks to participants. Fortunately review of large studies of DBT and the pilots with TRD patients (see 27; 38) shows no reports of adverse reactions to DBT. So we have identified and addressed 4 potential risks:

Suicide risk There is no reason to believe that the interventions or research procedures will increase suicidal risk: DBT and TAU are likely to reduce risk by reducing symptoms. NICE guidelines for BPD (44) recommend DBT for individuals with recurrent self-harm, and DBT is effective in reducing suicidal risk. Risk assessments will be conducted throughout the study by clinically-trained assessors. Serious risks will be discussed with participants and GPs when required by Mood Disorder Centre protocols. These protocols include specific methods for assessing and managing suicidal risk, and specific actions for researchers. DBT therapists are well trained in managing suicidal ideation and risk; our therapists already serve clients with these problems, and have access to risk-management resources. The TAU group will be subject to the same enhanced monitoring of suicidal risk. All participants will remain under the care of their GP or psychiatrist throughout the study.

Assessment Interviews and questionnaires may be upsetting if patients recall distressing events, but our previous work with depressed or personality-disordered patients shows that most people return to pre-assessment emotional arousal after assessment. As long or complicated assessment can tire participants, we shall minimise the length of assessments and make every effort to ensure the environment is comfortable. If a participant becomes distressed an on-call clinical supervisor will step in to manage risk.

Termination of Therapy Although this can be difficult for patients with TRD, our previous studies show that participants adjust to termination. We shall refer patients to their GP or psychiatrist if treatment is needed at the end of their DBT.

Risks to participant confidentiality All trial data will be identified only by trial number. All physical materials related to treatment and assessment will be kept locked separate from identifying information. Such data are sent to other professionals only when participants request it in writing for reporting serious risk to their GP or psychiatrist.

2.3 Rationale for the current study: Why is the trial needed now?

Our study is timely and appropriate for seven reasons: First, TRD is a chronic, disabling condition with few effective treatments, and severe patients are routinely excluded from evaluations of treatments. The results of our trial are urgently sought by both health professionals and patients. Second, we have enough evidence from preliminary trials to progress to the next stage of treatment development—a phase II/III RCT: DBT for TRD is the first treatment to target PD features that are common in TRD and may explain poor outcomes in previous studies (e.g. greater self-criticism, excessive control of spontaneous emotion, and inordinate fears of making mistakes

(10)). Third, the DBT for TRD treatment manual is grounded in a theoretical and experimental literature which posits a number of mechanisms of change. Fourth, our design provides a well-powered test of many of these important hypotheses in a large clinical population. Fifth, TL is one of the world's leading researchers in DBT, and acted as CI on the recently completed multi-centre RCT of DBT for BPD opiate addicts alongside the developer of DBT. Sixth, as DBT has become well established as a treatment for BPD in the UK, there are enough well-trained DBT therapists in the UK to disseminate the new DBT for TRD, if our proposed trial shows that it is efficacious. Seventh, in July 2010 the ISRCTN Register records no comparable recent or ongoing trials in the UK. In short this proposal is timely, builds on 14 years of research by TL and his team, and offers great potential for patients and NHS while fulfilling the remit of the EME Programme.

3. Research objectives

Our primary objective is to estimate the efficacy of DBT for TRD compared with TAU. We shall also extend current knowledge of the mechanisms of DBT treatment, and of moderators of treatment efficacy for this population, using cutting-edge statistical methods based on instrumental variables to minimise the bias from confounding that can distort conventional analyses. We shall also address the relative cost-effectiveness of DBT in comparison with TAU alone.

3.1 Efficacy

We shall estimate two measures of DBT efficacy: First, what is the effect of being *randomly allocated* to DBT rather than TAU? Second, what is the effect of *exposure* to specific 'doses' of DBT, where exposure is measured by adherence to DBT treatment protocols, and zero exposure corresponds to complete non-adherence (i.e. TAU). For both of these questions, the outcome measures will be measures of depressive symptoms (primary outcome HAMD), rates of remission, and measures of other symptoms including suicidal ideation or behaviour, PD symptoms, and global functioning (Section 7).

The first of these measures can be estimated by conventional analysis 'by treatment allocated' (45), previously known as 'by intention to treat'. Although that analysis will provide an unbiased 'pragmatic' estimate of the 'effectiveness' of DBT in clinical practice, this cannot be interpreted as 'efficacy' under ideal conditions because participants will vary in their adherence to the recommended course of treatment, and there will therefore be heterogeneity in DBT exposure. To account for this, we shall focus on the second measure and seek to estimate the causal effect of a specific exposure to DBT. In doing so, we must allow for the fact that attendance at therapy sessions occurs after randomisation, and may thus be subject to confounding. We shall address this potential bias by using instrumental variable (IV) techniques.

3.2 Mechanisms

We shall extend our efficacy analysis (which addresses questions relating to the effect of allocation and exposure to treatment), to ask questions about *how* DBT may be effective. Our approach includes both DBT-specific and trans-theoretical concepts, and recognises that important elements of psychotherapy may be common to many treatments. Our analyses will focus on 4 specific pathways between allocation and outcome: i) *Treatment exposure*, ii) *Therapeutic alliance*, iii) *Skill acquisition* and iv) *Expectancy*. These pathways are shown in Figure 2.

To estimate the *causal* effect of mediators, our trial will engage in manipulating selected mediators (46). We shall also measure variables which represent sources of variation in mediators that are unlikely to be contaminated by selection effects. In other words we shall identify instrumental variables (IVs), both experimental and observational, to facilitate causal analyses.

In addition to these primary pathways, we shall measure several potential modifiers of treatment outcomes. Based on our own and other research showing links between TRD and PD (11), temperamental risk aversion and reward insensitivity (30; 31; 47), and childhood adversity (12), we shall assess potential moderators of treatment response by measuring the following at baseline: (i) PD diagnosis (SCID-II), (ii) invalidating childhood experiences (ICES) and (iii) reward sensitivity or risk aversion (UPPS). We shall also conduct a complementary analysis of repeated measurements of outcomes and mediators using longitudinal models. This analysis will address theoretically-driven questions about the ordering of changes in key variables.

3.3 Causal analyses

In randomised experiments, post-randomisation 'intermediate' outcomes (e.g. mediators like treatment exposure and alliance scores) are influenced by the patient, the therapist, and other factors, and so are not under the control of the experimenter. Hence there is potential for confounding variables, associated with both the intermediate and study outcomes, to bias estimates of the effect of mediators. If these confounding variables are known and measured in the study, then suitable adjustments may be possible. However, it is more likely that confounders are unknown or unmeasured or both, and the intermediate outcome is therefore confounded. Thus conventional analyses often yield biased estimates of the effects of intermediate outcomes.

In econometrics, instrumental variable (IV) methods have long been used to estimate causal relationships from observational data. An instrument is a variable that is, by assumption, wholly mediated by other measured variable(s). For example, economists have used changes in tobacco taxation as an instrument for health outcomes: if the tax change has an effect on health it is assumed to be wholly mediated via increases or decreases in smoking behaviour, and not by other direct or indirect pathways (often termed the 'exclusion restriction'). If these assumptions hold, the instrument may be

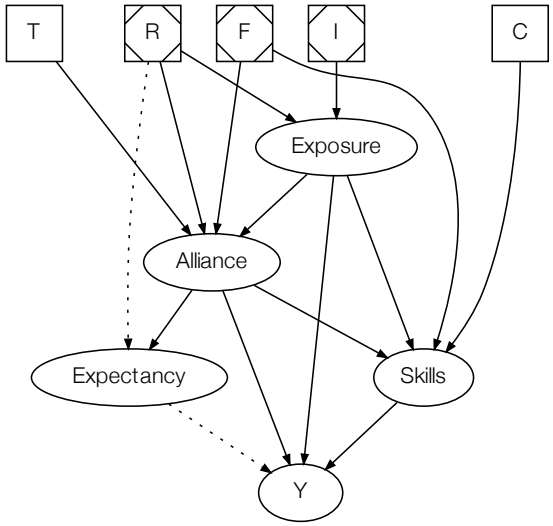


Figure 2: Causal pathways examined by the study. Nodes R, F and M indicate Random assignments to treatment (DBT/TAU), Feedback and Mode of reimbursement (from trial manager, or from therapist) respectively. Node Y indicates our outcome: depressive symptoms. The dotted lines indicate analyses we will perform to check assumptions relating to the exclusion restriction for treatment-assignment (see 3.3.4).

used to estimate the *causal* effect of smoking on health, without fear of confounding. (Note that we take "causal effect" to mean what happens on average if the experimenter intervenes and changes the value of each patient's intermediate outcome *while holding everything else constant*; the precise definition of the causal effect is specific to the analytical approach and the strength of the assumptions that the analyst is prepared to make.)

Although conditions within a clinical trial are highly controlled, mediation analysis is still difficult when mediators are not selected by design. For example, unobserved variables (e.g. 'readiness for change') might be the cause of both high therapeutic alliance scores and positive outcomes; if 'readiness' were unmeasured, then researchers might incorrectly conclude that alliance is a cause of good outcomes. Unless researchers are confident they have measured *all* potential confounders then standard analyses of mediators and outcomes cannot have a causal interpretation. Thankfully, although finding IVs that satisfy the requisite conditions can be controversial in observational studies, the issue is less vexed for RCTs. Any randomly allocated exposure which influences a mediator is a promising candidate for an IV because, by design, it cannot be associated with unobserved confounding variables also affecting the mediator. The central issue that must be justified is that the IV cannot have a direct effect on the study outcome — the exclusion restriction. The effect of the randomised exposure must be wholly mediated by variables which are measured within the study. Estimators for causal effects based on IVs are now widely used (48–50); for example, treatment assignment and treatment location have been used successfully as instruments in controlled studies (51), but where multiple

mediators are hypothesised it is crucial to find additional strong instruments to facilitate more complex analyses.

3.3.1 Instrumental variables

In our study the key IV is random allocation of therapy. However, we have augmented our causal analysis by making several further design decisions that use random assignment of participants and therapists to create additional instruments. Furthermore, we have identified pre-allocation measurements which are strong candidates to be used as instruments, although we shall in due course need to test that these meet the criteria for IVs. Our proposed changes create IVs that will yield unconfounded estimates of the effect of our four specified mediators. We shall also attempt to model a direct pathway between treatment allocation and outcome, independent of receipt of treatment.

3.3.2 Instruments for exposure to therapy and the therapeutic alliance

• *Feedback* Research shows that, when participants provide regular structured feedback to therapists regarding their progress and their perceptions of the alliance, outcomes are improved (52). Current evidence suggests that feedback increases the number of sessions attended, and is likely to improve client-rated alliance. Consequently, we shall randomise clients between providing and not providing feedback about the alliance to therapists in each treatment session (Session Rating Scale; SRS) and their perceived progress (Outcome Rating Scale; ORS). Therapist expectancies for feedback will be assessed to provide a measure of *allegiance* to this additional procedure.

• *Mode of reimbursement* Because financial contingencies positively reinforce attendance at therapy sessions, we shall randomise participants between receiving vouchers to reimburse them for entering the study and undertaking assessments within therapy sessions and receiving cheques to the same value by post from the trial manager. However, all participants will receive the same total reimbursement across the study.

• *Treatment setting* Because treatment setting is known to affect both adherence and expectancy for complementary therapies (e.g. acupuncture), and may also have potential to influence impression formation (and thus the alliance), we shall allocate the setting in which treatment is delivered at random: half our participants will receive treatment in standard NHS consulting rooms; and the rest will receive treatment in enhanced rooms, simulating private consultation rooms used for many complementary therapies.

• *Therapists* Therapists are known to vary in their tendency to generate high or low alliance ratings from patients (53). So we shall randomise patients between therapists within centres, thus adopting 'therapist' as an IV.

We shall measure several other variables at baseline to provide potential IVs for mediation analyses of alliance and skills. For the alliance:

• *Therapist characteristics and client interactions* We shall measure several characteristics of therapists known to predict alliance within session, to improve the prediction of alliance from our random assignment: in therapists, warm versus cold interpersonal style (54), and childhood attachment (55) have been found to influence alliance scores. Congruence of client and therapist characteristics (e.g. personal values (56) or cognitive style (57) also predict alliance and outcomes. We shall measure these variables at baseline in both participants and therapists, to provide additional potential IVs for our causal models.

• *Pre-trial measures of alliance for trial therapists* To enter the study, therapists must demonstrate adherence with non-trial patients. Mean alliance ratings from these sessions will be used to predict in-trial alliance scores.

• *Travel distances from home to treatment centre* Because travel times may predict drop-out, this is another potential instrument. We shall condition first stage regressions on indicators of socio-economic status of the home postcode, to minimise the possibility that travel times are directly correlated with outcomes.

3.3.3 Instruments for skill acquisition

As the ability to learn and consistently apply new coping skills is hypothesised to play a crucial role in DBT treatment, we are keen to identify variables with the potential to act as instruments for analyses examining mediation of treatment via skill learning:

• *Pre-treatment measure of behavioural compliance* Before randomisation, we shall ask all patients to complete a simple homework assignment, consistent with interventions from the positive psychological literature, at a set time each day. We shall confirm adherence to this task via automated telephone calls each day. We shall also measure the personality trait of *conscientiousness* at baseline as a supplementary instrument for skill application.

• *Prospective memory* We plan to measure prospective memory capacity at baseline as a potential IV for skill acquisition.

3.3.4 Checking for a direct effect of allocation

Because it is possible that treatment allocation itself has a direct effect on outcomes (so-called 'resentful demoralisation' 58) we shall include variables which may allow us to model this pathway. We shall measure whether allocation has an effect on expectancies for outcome independent of actual exposure to treatment. We operationalise this direct effect as change between participants' hypothetical expectancies before allocation (e.g. "what will happen if you are assigned to DBT") and actual expectancies after allocation (e.g. "what will happen now you have been assigned to DBT"), where this change is different for patients assigned to DBT and TAU. These analyses are exploratory, but are important because they seek to check the validity of the exclusion restriction for treatment allocation, upon which existing

causal analyses in this area depend.

3.4 Analyses of temporal patterns and precedents of change

In conjunction with our causal analyses of the therapeutic alliance, we wish to examine patterns of change and cross-lagged effects among depressive symptoms, alliance ratings, and skills learnt in DBT. We will also examine whether there are differential rates of change in positive versus negative affect (PA/NA), and whether different patterns of temporal ordering exist for PA and NA. Previous research has highlighted the distinct and important role of PA in adaptive coping (59). PA is thought to broaden the individual's attentional focus and behavioural repertoire and, as a consequence, build social, intellectual, and physical resources (60). We expect change in PA to be more rapid and more closely associated with factors common to psychosocial interventions (e.g. expectancy and alliance) than in NA, and to precede improvements in coping strategy. Furthermore, ecological momentary assessment (EMA) will enable us to answer questions related to the *variability* of affect in treated versus untreated patients. We expect daily variability in affect for DBT patients to rise early in treatment (as a consequence of the difficult work clients undertake with therapists) but to decline relative to TAU patients by 6 months. Though these longitudinal analyses estimate temporal ordering or so-called 'Granger causality' rather than true causality, they complement our primary mediation analyses by providing a richer picture of patterns of change in response to treatment. Through the life of the trial we shall monitor the active methodological literature on causal analysis in RCTs to ensure we make best use of our valuable dataset.

4. Research design

We propose a 2-stage, 2-arm RCT in 3 centres – Dorset, Hampshire and North Wales. All participants will receive treatment as usual (TAU) in accordance with an explicit manual. On randomisation all patients will be on anti-depressant medication (ADM) prescribed by their GP or psychiatrist. The trial will not alter these prescriptions in any way, though switching, augmentation or supplementation may occur as part of normal tailoring of treatment by their ADM provider. Participants in the experimental arm will receive DBT treatment over 6 months. Adaptive randomisation will balance baseline depression severity (HAMD > 25, yes/no), PD status (meets SCID-II criteria, yes/no), and age at onset of depression (before 21 years old, yes/no) of depression across groups without risk of subversion. To maximise power to test explanatory hypotheses we shall allocate patients to DBT and TAU in the ratio 3:2 – with minimal loss of statistical power for our analysis of our primary outcome, but increased power to test hypotheses relating to mechanisms. To facilitate instrumental variables analysis (Section 3) within the DBT group we shall use a factorial design to allocate participants at random: (i) to provide their therapist with feedback or not; (ii) to receive

reimbursements from their therapist or in the post; (iii) to receive treatment in a standard or enhanced consulting room; and (iv) between all available trial therapists within their centre (the adaptive algorithm will balance case-loads between therapists).

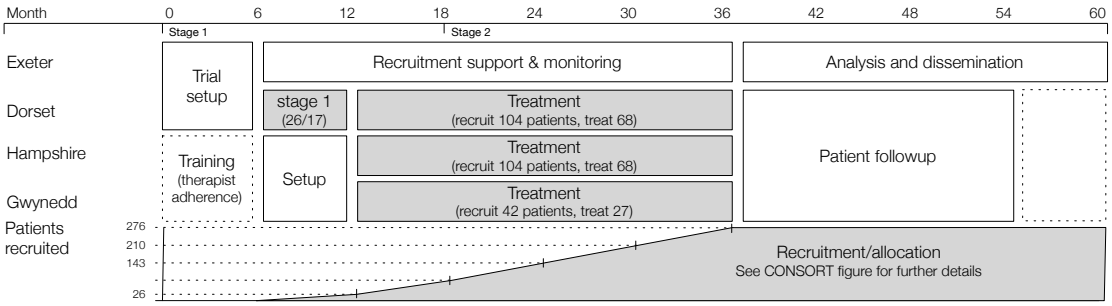
Two-stage design Before starting full evaluation of DBT we shall establish that we can deliver high-fidelity treatment and recruit enough participants. Thus we have divided the trial into two stages. In stage 1 our targets are: to achieve good adherence to the manual by trial therapists; to recruit at least 20 participants, ideally 26, in Dorset within 6 months; to achieve high response rates (at least 70%, ideally 80%) to the primary outcome HAM-D; and to show participant satisfaction with treatment (at least 80% scoring more than 16 on the Client Satisfaction Questionnaire–8, which shows indifference). Careful monitoring of these targets will enable us to refine our protocol before stage 2 begins, and the Trial Steering Committee (TSC) to approve or abort the second stage of the trial within 18 months. We shall begin preparation for, and recruitment to, stage 2 during this assessment of stage 1 to avoid incurring extra costs. In Stage 2 we shall allocate the remaining patients between the 2 arms across all 3 trial sites, and thus achieve our research objectives. We shall measure efficacy by comparing groups 6, 12 and 18 months after randomisation (i.e. after treatment, and at 6 and 12 months thereafter). The follow-up period of 18 months balances cost, loss to follow-up, and the diminishing returns of extended follow-up.

5. Study population

Inclusion criteria Patients must: (i) be at least 18 years; (ii) have a current diagnosis of major depressive disorder (MDD); (iii) have a HAMD score of at least 15; and (iv) have TRD. We define current MDD by SCID-I diagnosis (Section 7). We define TRD as having had two or more previous episodes of depression or meeting the criteria for chronic depression and, in the current episode, to have taken an adequate dose of ADM for more than 6 weeks without symptom relief. Participants may have Cluster A or C personality disorder, but this is not required. **Rationale for inclusion criteria** The definition of *treatment-resistant* depression in the literature is blurred, primarily because of the varied priorities of researchers and clinicians. While lay definitions of treatment resistance expect long periods of failure to respond to multiple ADMs or psychosocial interventions, researchers and trialists have typically adopted less stringent inclusion criteria for trials of TRD. Indeed, a recent systematic review (9) of 29 randomised trials for TRD concluded that the average minimum period of poor response to ADM was five weeks.

Thus, although the 2009 update of NICE guidelines for depression recommend a combination of ADM and high-intensity psychological intervention (Cognitive Behaviour Therapy, CBT, or Interpersonal Therapy, IPT) for moderate or severe depression, we shall not require a poor response to high-intensity psychotherapy for inclusion in our trial, for 3 reasons. First, failure to achieve symptom relief with an adequate ADM dose within 6 weeks is in itself a powerful indicator of poor long-term

Figure 3: Project timeline



prognosis (20). Second, unlike ADM, CBT or IPT are not always available to participants in the UK; so requiring a poor response to these treatments would reduce our ability to recruit participants. Third, individuals with chronic depression often prefer ADM treatment to psychotherapy (18) and emotionally-constricted PDs are less likely to seek treatment or consider their personality style problematic (61), making it likely that they prefer biological explanations and seek treatment only when symptoms are severe. So requiring poor response to psychotherapy may exclude participants that DBT was designed to treat.

Exclusion criteria We shall exclude patients: who have IQs less than 70 or insufficient English to complete treatment and assessment; who meet DSM-IV criteria for dramatic-erratic PD (borderline, histrionic, antisocial or narcissistic PD), bipolar depression or psychosis; who have a primary diagnosis of substance dependence or substance abuse disorder; who are currenlty receiving standard DBT; are on a waiting list for standard DBT; have received standard DBT within the last six months. We exclude BPD and other dramatic-erratic under-controlled PDs (Cluster B) because: (i) DBT has proven efficacy in reducing depression and suicidal behaviour in BPD (Lynch et al., 2007) and (ii) DBT for TRD is designed to treat specific problems associated with behavioural over-control common in TRD, chronic depression and Cluster A and C personality disorders, e.g. cognitive and behavioural rigidity and emotional constriction.

Contrasts with previous studies of TRD Importantly, we shall NOT exclude patients with antecedent dysthymic disorder, or with long-standing or intractable depression, previous or current suicidal behaviour, or co-morbid Cluster A or C personality disorder. Indeed, we expect about 60% of the sample to meet criteria for diagnosis of Cluster A or C personality disorder. Variation in PD status at baseline will allow us to study the moderating role of PD diagnosis on treatment outcome. We shall also include older adults (two pilot studies of DBT for TRD were with middle-aged or older adults; 27; 38).

Recruitment methods

Based on experience of other large trials of psychological therapy at Exeter (COBALT, PREVENT) and of the Oxford-based MBCT trial, of which ITR is methodological lead, we shall recruit participants from primary and secondary care. Indeed we have already identified many potential participants in secondary care centres run by the clinical PIs, to supplement the majority recruited via primary care.

Our primary care recruitment procedure, tried and tested in previous multi-centre trials run from Exeter and North Wales, comprises 6 steps: (1) Search of general practice databases by trained Clinical Studies Officers to identify potentially eligible patients. (2) Consult medical records to check whether these patients meet inclusion criteria. (3) GPs screen these patients for suitability. (4) GPs sign pre-prepared letters describing the study and inviting patients to opt out or consider participating in the trial. (5) Unless patients opt out, they will be contacted by telephone to discuss the study and, with their oral consent, screened for eligibility. (6) Potential participants who are eligible and willing attend for trial assessment, when they are invited to sign a formal consent form. GPs and practice nurses can also refer patients from routine consultations. Where possible we shall use computer prompts for GPs who write repeat ADM prescriptions to consider the trial. We plan to screen some 9000 potentially-eligible patients across 35 GP practices, and we are confident of recruiting at least 200 patients. Several recent trials which have recruited successfully have given ITR and our clinical leads strong links with general practices in our centres, with the result that recruitment at these practices could soon begin.

Recruitment of patients from secondary care will repeat steps 4 to 6, but letters will be signed by clinicians responsible for secondary care services. We have already identified many patients waiting for treatment who may be eligible for the trial. In Dorset, 277 patients meet criteria for a current depressive episode, of whom 152 are known to have had at least 1 course of ADM. In North Wales 80 TRD patients are on waiting lists for secondary care for depression, of whom over 80% have undergone ADM treatment.

Based on the demographic characteristics of our sites, the prevalence of TRD, the number of patients already identified as potentially eligible by clinical PIs, and the experience of other large trials at Exeter and elsewhere, we expect to recruit between 5 and 7 patients per month in stage 1, rising to between 10 and 12 patients per month in stage 2, when all sites will be active. Figure 3 gives more detail on predicted recruitment rates.

6. Planned interventions

Our design compares anti-depressant 'treatment as usual' delivered in accordance with an explicit manual (TAU) with TAU and psychotherapy (DBT). Both TAU and DBT groups will receive ADM prescribed by their GP or psychiatrist in accordance

with the TAU manual, but independently of the study team.

6.1 Treatment as Usual (TAU)

In our study all patients will receive ADM prescribed by their GP or psychiatrist in accordance with an explicit manual including up-to-date prescribing information. NICE guidelines advocate combining ADM with CBT for persons who have not responded to either pharmacological or psychological interventions (36). However, our TAU manual is based on knowledge that (i) evidence for the efficacy of CBT in this severely depressed population is weak, and that trials of promising therapies for TRD have reported poor results; (ii) access to psychotherapy suitable for this severely ill population is very restricted in the UK; and (iii) our TAU comparator is helpful in estimating the efficacy of DBT relative to a manualised version of the modal treatment for TRD patients in the UK. We shall also gather data on adherence to medication and receipt of concurrent psychotherapy from patients' reports and medical records. Our TAU manual includes procedures for identifying and responding to poor medication adherence, including providing additional information to participants, GPs or psychiatrists. We shall include ADM adherence and type and frequency of concurrent psychotherapy as covariates in our analyses.

Rationale for choice of ADM-TAU control

Relevance Though NICE guidelines suggest combined approaches (e.g. ADM and high-intensity CBT) for moderate-to-severe depression or depression not responding to first-line treatments, ADM alone is more readily available in the UK and often the preferred treatment for those with unresponsive depression (18). Moreover, there are few good data to compare augmentation with psychotherapy with switching or supplementing ADM. Recommendations are primarily based on one multi-centre study designed to compare switching and augmentation following poor response to acute ADM treatment; results showed no significant differences between approaches with less than one third remitting following poor response to acute ADM treatment (STAR*D 20; 21). Thus, psychotherapy augmentation strategies cannot yet be considered standard care for TRD, underlining the importance of further research.

Feasibility Although comparison between DBT and another psychotherapy might provide useful information, high-intensity psychotherapies (CBT or IPT) are not readily available in the UK, and would require many more resources to develop extra treatment sites, train therapists and recruit a much larger sample of patients to detect the likely smaller effect size. Moreover, the optimal psychotherapy comparator for studies examining TRD or chronic depression has yet to be established. A trial comparing one multi-faceted therapy with another, without knowledge of, or control for, overlapping components is unlikely to provide value for money. In short, at this stage of treatment development, we believe that a comparison of DBT with another psychotherapy would be premature.

Ecological validity We provide ADM through participants' GPs or psychiatrists rather than researchers to protect the ecological validity of our trial; keeping treatment costs low was an important secondary consideration. In the UK almost all ADM prescriptions are issued by GPs. Although prescribing by the trial team would increase consistency, our participants' experience of ADM would differ from current practice, and could change motivation or non-specific benefits of ADM. Using patients' GPs or psychiatrists to provide ADM-TAU also ensures that potential drug interactions are monitored by a doctor familiar with the patient's medical history. In addition, findings from the recent multi-centre STAR*D study (20) suggest that stepwise prescribing protocols lead to higher relapse rates in those who required more treatment steps. So study-provided interventions may fare no better than community-provided treatment, particularly when patients have failed to respond to their first course of ADM. Hence we shall monitor prescribing and medication adherence in both groups, so that our analyses can control for medication use.

6.2 Dialectical Behaviour Therapy (DBT) plus TAU

DBT will comprise weekly 1-hour individual DBT sessions and weekly 2-hour group DBT sessions over 24 weeks. All DBT participants will receive TAU, including ADM prescribed by their GP or psychiatrist. Though differential adherence is a potential risk whenever usual treatment includes medication, we shall train DBT therapists to avoid discussion of medication adherence; they will advise patients who raise this issue during therapy to talk to their ADM prescriber. Our previous research using an ADM comparator did not find significant differences in medication use between TAU and DBT plus TAU (27; 38). Following the treatment manual, DBT therapists will strongly discourage concurrent psychotherapy. DBT for TRD is a psychosocial intervention with a manual that outlines in full the rationale for individual and skills training sessions (24). DBT for TRD includes most of the components of standard DBT for BPD. Individual DBT is designed to rectify motivational and behavioural flexibility deficits, and group skills training to rectify skill deficits. Brief phone contact with the patient's therapist in crises is an adjunct to treatment. Our pilot trials showed that only about half of patients used such phone contact and no patient called more than 3 times over the study (27).

DBT Individual Therapy Participants will meet their individual therapist for weekly 50-minute sessions over 24–26 weeks. The weekly agenda depends on the current maladaptive behaviour to be stopped or reduced or the adaptive behaviour to be introduced or increased. Treatment targets follow the hierarchy: (1) reduce life-threatening behaviours (2) reduce therapy-threatening behaviours (e.g. missing sessions) (3) reduce wellbeing-threatening behaviours, notably depression, and increase openness and flexibility. **DBT Group Skills Training** The skills training manual is tightly structured and defines the content and format of each session. The training is didactic, with strong emphasis on skill use, behavioural rehearsal, feedback, coaching and homework. Training includes 5 skill modules (Table 1). **DBT**

Team Consultation As in standard DBT, a weekly team meeting is part of the treatment. This serves several important functions, including preventing therapist 'burn-out', providing support for therapists, both general and about specific participants, and improving empathy for patients. **DBT Therapists** Sixteen DBT therapists (6 in Dorset, 6 in Hampshire and 4 at North Wales) employed by their local NHS Trust (Health Board in Wales) but partially seconded during the trial will spend an average 20% of their time providing DBT therapy in the trial over 2 years (Hampshire and North Wales) or 2.5 years (Dorset). **Therapist Training** We chose our 3 sites because they have existing NHS-based DBT programmes and experienced DBT therapists. Each site's clinical lead is internationally recognised as a senior DBT trainer. Therapists will be mental health professionals who have been, or will be, trained in a standard 10-day intensive DBT course. During trial set-up, therapists will be trained to adherence in DBT for TRD. During Stage 1 all DBT therapists will be supervised by their local clinical lead for 1 hour per week until they reach adherence. Thereafter, in addition to normal team supervisions, therapists will receive individual supervision from TL or local clinical leads if their scores fall below 3.9 on the DBT Adherence Rating Scale (DBT-ARS) (62).

Therapist Monitoring and Adherence to Treatment During the first 3 months of the trial, TL will review the DBT-ARS and its manual to modify it for DBT-TRD in collaboration with Prof Marsha Linehan (ML), the principal author of DBT for BPD. We shall videotape all trial therapy sessions and sample a random 10%, stratified by patient-therapist pair, of individual sessions for adherence rating by reliable British raters in collaboration with ML. The DBT-ARS generates a global index of DBT adherence and sub-indices for the 12 DBT domains from 66 items, each operationalised with behaviourally defined anchor points in the manual (62). Inter-rater reliabilities of indices range from 0.78 to 0.83. Correlations between sub-indices and the global index range from 0.89 to 0.99. Similarly we shall sample a stratified random 10% of group skills-training sessions for adherence rating. DBT therapists will be trained not to target ADM usage but instead to refer patients to their GP or psychiatrist to minimise differences in adherence between DBT and TAU; compliance with this will be monitored through the modified DBT-ARS.

7. Proposed outcome measures

Clinical interview measures (months 0, 6, 12, 18) The primary outcome is the *Hamilton Rating Scale for Depression* (HAMD; 17-item version 63), which measures depressive symptoms. We shall infer remission from a HAMD score of less than 8 (64) combined with minimal functional impairment as assessed by the *LIFE-RIFT* (semi structured interview 65). We will measure suicidal ideation and behaviour by the *Scale for Suicide Ideation* (SSI; 66) and the *Suicidal Behaviour Questionnaire* (SBQ; 67).

Health economics (months 0, 6, 12, 18) Health-related quality of life and cost effectiveness will be assessed in terms of quality-adjusted life years (QALY), using the

EuroQol (EQ–5D; also assessed at month 3) (EQ-5D; 68) , which generates a single standardised index for health status from a simple descriptive profile in which clients indicate their health status on five dimensions. We shall collect complementary data on service and other resource use through the Adult Service User Schedule (AdSUS) (see 69).

Client Satisfaction (month 6 of stage 1 only) In stage 1 we shall use the *Client Satisfaction Questionnaire–8* (CSQ-8; 70) to assess the acceptability of treatment in both arms of the trial.

Moderator or control variables measured only at baseline We shall ask participants to complete a brief questionnaire asking about basic demographics. The *Structured Clinical Interview for Diagnostic Disorders Axis I and II* (SCID-I & II) is a semi-structured diagnostic interview for Axis I & II disorders used to verify the presence of MDD at study entry and to estimate rates of cluster A & C PD in the sample. The *Invalidating Childhood Experiences Scale* (ICES; 71) asks participants to rate negative childhood experiences ($\leq$18 years) in relation to each parent. The *Urgency Premeditation Perseverance Sensation Seeking* scale (UPPSS; 72) measures reward sensitivity and risk aversion. We shall also use the *Frost Multidimensional Perfectionism Scale* (FMPS; 73). To control for social desirability we will ask participants to complete the *Balanced Inventory of Desirable Responding-Brief version* (74). To investigate under-control and resiliency we shall ask participants to complete the *Ego-Undercontrol and Ego-Resiliency* scales (75). The *Personal Need for Structure* scale measures over-control (76). Also at baseline both clients and therapists will complete: *Conscientiousness* (subscale of the NEO 77); *Prospective memory* (CAM-PROMPT 78); and *Personal Values* (Schwartz Values Scale; SVS 79).

Online outcome and mediator assessments in months 0,3,6,12 and 18 The *IIP-PD* (80) is a dimensional personality disorder measure. The *DBT-CCL* (81) is based on the earlier Revised Ways of Coping Checklist and is an inventory of emotional coping skills taught in DBT. The *Ambivalence Over Emotional Expression Questionnaire* (AEQ; 82) is a self-report measure indexing emotional constriction. The *White Bear Suppression Inventory* (WBSI; 83) indexes suppression and avoidance of unwanted thoughts. *Social network size* will be estimated with the 3-item SSQ (84). The *Acceptance and Action Questionnaire-II* measures psychological inflexibilty (85). **Online assessments in months 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 18** The *PHQ–9* (86) is a short but psychometrically valid assessment of depressive symptoms. The *PANAS* (20 item version; 87) provides independent estimates of positive and negative affect. The *Emotional Approach Coping scale (EAC)* (41) is a self-reported index of emotional experience that is also sensitive to depressive rumination. Among DBT participants only the *Credibility-Expectancy Questionnaire* (CEQ; 88) indexes positive expectancies for treatment; and the patient-rated version of the *California Psychotherapy Alliance Scales* measure therapeutic alliance (DBT condition only; 89).

Ecological Momentary Assessments (EMA) The 10-item *PANAS* (90) provides a brief measure of positive and negative affect. To measure *use* of coping skills taught in DBT

we shall use selected items from both the DBT-CCL and the EAC to identify recent skills rehearsal in everyday life. In addition *conscientiousness* — patients' ability to complete simple homework assignments — will be assessed via ecological momentary assessment (EMA) before randomisation (Para 2 of Subsection 3.3.3).

8. Assessment and follow up

Our protocol balances the need to minimise burden on participants with 3 research priorities: (i) estimating primary and secondary treatment outcomes accurately; (ii) identifying and characterising moderators of these outcomes; (iii) identifying potential mediators and modelling causal pathways. Each participant will complete process and outcome assessments at baseline and over the 18 months of treatment and follow-up.

8.1 Assessment of efficacy

Assessment schedule Section 7 gives assessment times for each instrument. We shall measure our primary outcome (HAMD) and secondary outcomes at months 0, 6, 12 and 18; and online self-reported outcomes more frequently. **Observer-rated assessments** Our assessment of MDD, PD (SCID-I&II) and depression (HAMD) will be by experienced raters trained in the use of these instruments. To minimise observer bias, they will do so blind to group allocation. A blind independent rater will assess audio-recordings of the interviews. Low inter-rater agreement (Kappa less than 0.7) for an individual observer will generate extra training and independent ratings. **Online self-report measures** A trial-specific website will collect self-reported secondary outcomes and mediator measures. Participants will complete assessments from home, or via computers at trial centres. This website will also collect data on expectancy and therapeutic alliance. Printed questionnaires will always be available to overcome technical difficulties and improve accessibility for patients with disabilities; we shall enter data from these questionnaires twice to maintain accuracy. **Ecological Momentary Assessment of mood and coping** Telephone-based sampling will initiate short measurements of mood and coping. A computerised system developed and tested by one of the PIs (BW) will make automated calls to participants during the 6-month treatment period; patients will respond by phone keypad. **Medication adherence and healthcare costs** To supplement AdSUS scores, experienced staff from local research networks will abstract GP and psychiatric records, notably whether ADM prescriptions have been collected on time. To maintain blindness data will be collected in sealed envelopes at each follow-up. We shall also collect data from therapy files at trial end, distinguishing in this efficacy trial between treatment and service support costs. **Use of clinicians' notes** We shall derive attendance at psychotherapy from clinicians' notes of individual therapy and group skills training. **Other techniques to minimise assessment bias** Assessors will conduct assessments away from therapy, using methods developed in the current MBCT trial at the Exeter

MDC, and assess their own blindness at each follow-up for use as a covariate. If blinding is compromised, a new assessor will conduct further follow-ups when possible. The use of self-reported measures will provide a check against observer bias for secondary outcomes.

8.2 Assessment of health economics

The economic evaluation will take the NHS-Personal Social Services perspective preferred by NICE, and will also estimate productivity losses resulting from time off work or reduced productivity at work due to illness. Data on therapist contacts in the DBT group will be collected from therapist records to avoid patients revealing their treatment group to the research assessors. We shall collect data on indirect time costs, including preparation and supervision, directly from trial therapists. Data on the use of other health and social services will be collected using the AdSUS. Productivity losses will be measured using the absenteeism questions of the World Health Organisation's Health and Work Performance Questionnaire (91). Cost of DBT will be directly calculated from salaries via micro-costing approach, and national UK unit costs applied. Productivity losses will be valued using the human capital approach. Sensitivity analysis will use the friction cost approach to address concerns that the human capital approach overestimates these losses.

8.3 Assessment of safety

The safety of participants will be paramount. However there is no reason to believe that either intervention or any research procedure will increase risk to participants. Indeed both treatments should reduce risk. During the treatment period telephone sampling will regularly monitor patient mood and coping. The clinician responsible will act on clear changes by providing skills coaching (DBT) or initiating case review and referral (TAU). All participants will have access to the support they were receiving before the research study (e.g. GP or CMHT or both). Serious Adverse Events (SAEs) are untoward events (in DBT or TAU) including admissions, suicide attempts, self-harm, overdoses and reactions to medication. Local PIs will be responsible for identifying SAEs, assessing expectedness and causality, and reporting to the CI immediately. The CI will report SAEs to the DMEC regularly and in expedited fashion when the SAE is a Suspected Unexpected Serious Adverse Reaction (SUSAR). SAEs and serious risks identified by the trial will be discussed with the participant. With their permission, we shall inform their GP in line with the Exeter Mood Disorders Centre protocol for assessing and reporting risk.

Table 3: Effect sizes for DBT for TRD from pilot studies (for group differences in change on the HAMD from baseline to end of treatment)

Study	N	Group Diff.	D
Lynch et al. (2003)	31	2.75	0.71
Lynch et al. (2007)	31	5.47	0.85
Harley et al. (2008)	19	3.32	1.45

9. Proposed sample size

This section distinguishes between the number of *patients randomised*; the number of patients who complete assessments required for primary comparisons between DBT and TAU; and the *effective sample size*, that is the equivalent number of completing patients available for comparison after adjusting for clustering effects due to therapists and groups. Thus we plan to randomise enough participants to yield the required power to compare DBT and TAU after accounting for clustering and loss to follow-up.

Completers required for the primary outcome Considering the effect sizes achieved by the 3 original trials of DBT for TRD in the US (Table 3), including two by TL (designer of DBT for TRD), we judge that a standardised difference of 0.4 between groups (half that achieved by TL) is both feasible and desirable. It is also likely to be important to the National Institute of Health and Clinical Excellence (NICE). The HAMD is the primary outcome for which we have estimated statistical power at six months after the end of treatment. The previous data suggest that the population SD for TRD is at most 5. Simulations of the random effects models described below suggest that, if there were no relevant intra-class correlation coefficients, a sample of 200 analysable participants from the 3 centres would yield 80% power to detect a mean difference of 2 points on the HAMD (ie a standardised difference of 0.4) using a significance level of 5%. Although this difference is not clinically significant for an *individual patient* (the reliable change index for HAMD is about 8 points), the aim of our efficacy analysis is to test whether this intervention has a standardised effect that will benefit the TRD population on average, rather than to show significant benefit for *individual patients*. Nonetheless, we shall report the estimated proportion of participants in both groups experiencing clinically meaningful change to guide clinicians.

Loss to follow-up The 3 previous trials, and 2 similar trials of treatment for depression we are conducting in the UK (Staying Well After Depression and FolATED), suggest we can collect analysable data from 83% of participants. We therefore increase our initial target to 240. To increase the power of our analysis of DBT mechanisms we shall randomise at a ratio of 3:2, allocating 144 patients to DBT and 96 to TAU (92).

Clustering—the intra-class correlation (ICC) Because DBT combines sessions for individual participants by individual therapists with sessions for ever-changing groups of participants, ICC is difficult to estimate. So we have conservatively assumed

that ICC will resemble that in individual psychotherapy. Kim et al (93) reported an ICC of 0.015 when analysing HAMD outcomes from the NIMH Treatment of Depression Collaborative Research Program by 'intention to treat'. So we postulate an ICC of 0.015, equivalent to a design effect of 1.25 if 16 therapists each treat an average of 9.375 analysable participants. To maintain power we have therefore increased the number of participants starting DBT to 180 (of whom 150 will be analysable, thus yielding the same power as 120 independent DBT participants).

Final N to be randomised Thus we shall randomise 276 patients, viz. 180 to DBT and 96 to TAU. These will comprise 26 in the pilot phase (all from Dorset), and 250 in the definitive phase: 104 from Dorset, 104 from Hampshire and 42 from North Wales. Accounting for our conservatively estimated design effect of 1.25 due to therapist and group clustering, 180 experimental participants subject to clustering will yield the same statistical power as 180 / 1.25 viz. 144 unclustered participants. However the 96 control participants are not subject to any clustering. Hence the 'effective sample size' (in the sense that it generates an accurate estimate of statistical power) is (144 + 96) x 0.83, viz. 200, which is sufficient to power the study for the target effect size of 0.4 for DBT.

10. Statistical and economic analysis

PC and the trial statistician will develop an analysis plan consistent with WWORTH SOP25, itself consistent with ConSORT recommendations. Both Data Monitoring and Ethics Committee (DMEC) and Trial Steering Committee (TSC) will review and approve this plan. Our causal analyses of efficacy and mechanisms will be developed by BW and PC in conjunction with the trial statistician. In his capacity as a member of the Mental Health Research Network (MHRN) Methodology Group, PC will present the proposed causal analysis plan to the group to obtain feedback. We plan to publish this analysis plan with the trial protocol as soon as feasible.

Efficacy For our primary measure of depression (HAMD), we will perform both analysis 'by treatment allocated' (45), previously known as 'by intention to treat' (ITT), and causal analyses using instrumental variables (IV). Analysis by treatment allocated will use repeated HAMD measures at 6, 12 and 18 months after randomisation; contrasts for the effect of DBT versus TAU will be estimated using subject-specific multi-level random effects models (94); linear regression will be used to analyse the final outcome. The causal analysis must account for the 'dose response' relationship between adherence (and consequent exposure to DBT) and HAMD. The primary IV for adherence will be the treatment allocated. Two types of IV estimators will be used to analyse the final outcome: 2-stage least squares and G-estimation of structural nested mean models (SNMMs) (48; 95) will be used to estimate the dose effect and assess sensitivity to different assumptions about the dose-response relationship. Each model will adjust for baseline factors to minimise the extent of unobserved confounding; the effect of (pre-randomisation) moderators will be estimated by including interactions

with the relevant baseline factors. Moreover, we will investigate the sensitivity of these results to other IVs, namely Feedback, Mode of reimbursement, Setting and Therapist, both separately and in combination. The analysis of HAMD measures (Section 3.4) at 0, 6 and 12 months will be based on models for the repeated measures.

Mechanisms To understand the mechanisms through which DBT affects the outcome we will conduct a series of mediator analyses using recently developed models (49; 50). These models use IVs to account for the effect of unobserved confounding on the mediators. The basic analysis will follow others in using treatment allocation and its interaction with baseline measures as IVs. We will first consider each potential mediator in turn (Expectancy, Exposure, Alliance and Skills) and estimate both the direct effect of treatment (via exposure) and the indirect effect via the mediator. The estimates will be compared with the naive analysis in which we assume no unobserved confounding; moreover, the assumptions under which the causal estimates are unbiased and consistent will be clearly stated and subject to critical scrutiny. Finally, we shall extend these pairwise analyses to a joint model of the mediators using the IVs described above. We will consider two-stage estimators and structural equation models and assess identification (96). We shall also explore sensitivity to different choices of dose-response model, different IVs (Section 3.3) and identification using multiple IVs.

Temporal ordering To examine patterns of temporal ordering we shall use multivariate growth curve models including autoregressive and lagged terms. These analyses are longitudinal rather than causal and hence do not require the use of IVs.

Therapist effects Additional models will estimate clustering of individual outcomes due to therapists and group treatment. Although we have too few therapists for a precise estimate of variability, additional estimates of intra-cluster correlation coefficients in this population will provide a useful addition to a sparse literature.

Missing data We will investigate the use of multiple imputation via iterative chained equations and data augmentation to adjust for missing data under the usual missing at random (MAR) assumption (97). For structural equation modelling, we shall use full information maximum likelihood estimation, as implemented in software packages like Mplus (98). In general we shall use sensitivity analyses to test whether relaxing the MAR assumption affects our models; in structural equation modelling we shall do so by joint modelling of data and non-response (97).

Economic evaluation We shall analyse differences in mean costs by parametric t-tests and confirm the validity of findings by bias-corrected, non-parametric bootstrapping (i.e. repeat re-sampling). Secondary analysis will include productivity losses. Cost-effectiveness will be assessed by estimating incremental cost-effectiveness ratios (ICERs), both for HAMD and quality-adjusted life-years using the EQ–5D measure of health-related quality of life. Uncertainty around the cost and effectiveness estimates will be represented by cost-effectiveness acceptability curves.

Timetable & reporting procedures We plan analyses of the entire dataset after treatment, and again six months and 1 year later. We shall not analyse subgroups

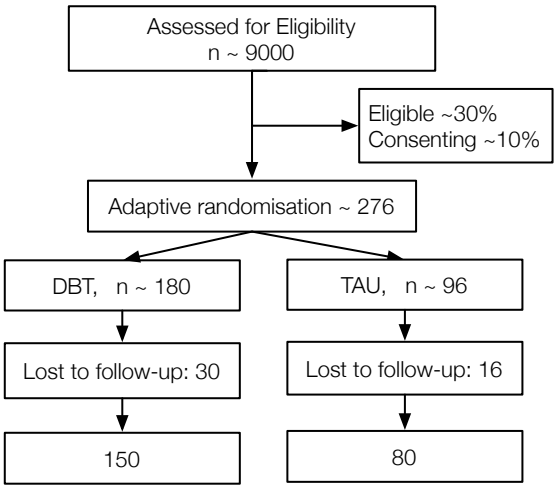


Figure 4: Study flow diagram

(although some planned analyses will adjust for baseline covariates). We shall report Serious Adverse Events (SAEs) to each meeting of the DMEC. We shall monitor data quality by following WWORTH SOPs 17 and 18, consistent with the International Conference on Harmonisation Good Clinical Practice guidelines.

11. Ethical arrangements

We shall submit the trial protocol via IRAS for multi-centre research ethical approval and then to the 3 clinical centres for site-specific assessment. The main ethical issues are:

Consent We shall tell eligible participants about the design, process of randomisation and both treatments before seeking consent. We shall give them opportunity to clarify details before consenting. Participants will be informed that they may withdraw at any time without standard treatment being affected in any way; treatment and collection of follow-up data will be completely separate so that participants may withdraw from either or both at any point. **Treatment phase** Risks to patients in both DBT and TAU groups are outlined and addressed in Section 2.2 above.
Follow-up phase During the follow-up phase participants will not receive active treatment from the research team, but will still complete assessments. Hence we shall continue to use our standard protocol for managing suicidal behaviours and alerting other professionals. **Payment** Participants will not receive payment for participating in the trial, but will receive appropriate reimbursement of expenses incurred in completing assessments.

12. Research Governance

The University of Southampton will sponsor this trial. The TSC and DMEC will follow MRC Guidelines on Good Clinical Practice in Clinical Trials. Professor Mark Williams from Oxford has accepted our invitation to chair the TSC. We shall recruit the other independent members of these committees immediately after funding. The TSC, the quarterly Trial Management Committee (TMC) and the monthly Trial Research Team (TRT) will each meet regularly and we propose that both TSC and TMG include service users as permanent, active members. We shall keep all trial documents for at least 5 years after the main publication from the trial; we (or successors) shall then (or earlier) deposit an anonymised data set in an appropriate databank.

13. Project timetable and milestones, 14. Gantt chart, & 18. Flow chart

See Figure 3 above (page 8) for our project Gantt chart and recruitment schedule. See Figure 4 for the project flow chart. Our project milestones are as follows (project month in parentheses):

- Establish a Trial Steering Committee; initiate staff recruitment processes; prepare ethics application (0)
- Trial Steering Committee meet to finalise study protocol (2)
- All therapists in Dorset achieve DBT adherence (4)
- Recruitment of Clinical Studies Officer (1) and assessor (3) in Dorset
- Start recruitment of patients in Dorset (4)
- First 26 patients finished treatment in Dorset (12)
- Clinical Studies Officer (6) and assessor (12) posts filled in Hampshire and North Wales; Start recruitment (12)
- TSC meeting convened to discuss continuation (15-16)

- **Decision to continue with stage 2 (see section 4) (<18)**
- See Fig. 3 for detailed recruitment targets in the treatment period
- Last patient finishes treatment (36)
- All adherence ratings of therapy sessions complete (48)
- Last follow-up data collected (54)
- Data integrity checks; data cleaning complete; analysis commences (40)
- Results presented at international conference (months 40-60)
- Draft of follow-up data submitted for publication (60)

15. Expertise.

Professor Thomas R. Lynch (TL) TL is one of the world's leading researchers of DBT and, based at the University of Southampton, will lead the trial team. TL has been PI or CI on 7 RCTs, including a large multi-site trial of DBT, and has conducted an

extensive programme of scientific research in mood- and personality-disordered patients in the USA and UK. *Contribution* Chief Investigator; oversight of trial research, design and methodology, training and supervision of DBT therapists and clinical protocols, supervision of trial manager. **Professor Ian Russell (ITR)** ITR, one of the UK's most experienced trialists and now head of WWORTH, the Registered Clinical Trials Unit at Swansea, has run more than 30 large RCTs including 2 current multi-centre trials of depression (MBCT for relapse prevention; folic acid to augment anti-depressive medication). *Contribution* Lead trialist; oversight of research governance and data safety; supervision of trial statistician and data manager. **Dr Ben Whalley (BW)** BW is Lecturer in Health Psychology at the University of Plymouth and affiliated with the University of Bristol Centre for Multilevel Modelling (CMM). BW has particular expertise in the use of online and telephony-based data collection, and analysis of longitudinal data. *Contribution* Design and methodological issues; mediation analysis of longitudinal data. **Dr Paul Clarke (PC)** PC is a statistician based at the Centre for Market & Public Organisation in the Department of Economics at the University of Bristol. His expertise is developing and applying statistical methods, especially those for causal analysis using instrumental variables, and for incomplete data. He is a member of the MRC's Mental Health Research Network Methodology Group, and affiliated to the University of Bristol Centre for Multilevel Modelling. **Dr Sarah Byford (SB)** SB is Reader in Health Economics with expertise in the economic evaluation of mental health services. SB is CI on an NIHR Clinical Trials Board grant evaluating the prevention of depressive relapse, an NIHR HTA programme grant evaluating psychodynamic psychotherapy, CBT and TAU in adolescents with moderate-to-severe depression, and an MRC grant evaluating joint crisis plans for people with PD. *Contribution* Oversight and publishing of economic evaluation of the trial; supervision of junior health economist. **Dr Roelie Hempel (RH)** RH is currently a Post-Doctoral Research Fellow at the Mood Disorders Centre (MDC) of the University of Exeter. She is manager of the Bio-Behavioural Lab at the MDC and has extensive experience in recruiting and testing psychiatric patients. *Contribution:* Trial Manager, Southampton. **Professor Susan Clarke (SC)** SC is the Foundation Chair of Mental Health at Bournemouth University and lead Consultant Clinical Psychologist at an NHS Beacon Service for patients with a personality disorder. She has 13 years experience of delivering DBT in NHS settings, and 12 years experience as a UK DBT trainer. SC has been CI of 4 small RCTs, and supervised 4 successful ESRC PhD studentships on the treatment of PD. *Contribution* Clinical oversight of Dorset centre; supervision of local research staff; input into design and methodological issues; training in adherence ratings. **Professor David Kingdon** DK is Professor of Mental Health Care Delivery at the University of Southampton, UK, and Director of R&D and honorary consultant adult psychiatrist for the Hampshire Partnership NHS Trust. His research interests are in cognitive therapy of severe mental illness in which he has conducted definitive efficacy and effectiveness RCTs. He has collaborated on grants funded by MRC, NIHR and US NIMH. *Contribution* Clinical oversight of Hampshire centre; supervision of local research staff; psychiatric and ADM advice to study; input

into design and methodological issues. **Dr Michaela Swales (MS)** MS is a Consultant Clinical Psychologist with 15 years experience of delivering DBT in NHS settings. She is the Director of the UK's national training programme in DBT that has seeded over 200 DBT programmes nationally and trained over 1000 professionals. MS has a joint NHS-University appointment between BCUHB and the School of Psychology of Bangor University. MS recently completed a 3- year ESRC-KTP grant examining the implementation of DBT in the NHS. *Contribution* Clinical oversight of North Wales centre; supervision of local research staff; training of DBT therapists. **Dr Heather O'Mahen (HO)** HO is an expert on psychosocial treatments for depression, and Deputy Director of the Exeter Mood Disorders Centre. Her primary research is on the treatment of depression and improving engagement and adherence to treatment. She is CI on two RCTs. *Contribution* Advisory and oversight role on the treatment for depression. **Professor RE Remington (RER)** RER is Deputy Head of School of Psychology at the University of Southampton. He has 30 years experience of clinical and educational research and most recently was CI on RCT for autistic children (2007). *Contribution* Advisory and oversight role to the Trial Management Team at the University of Southampton and to the Dorset centre in collaboration with Prof Clarke.

16. Service Users

Following WWORTH SOP09, we shall work with *Involve* and *Cynnwys Pobl* to recruit 4 patients who have experience of TRD. Service users will contribute to trial development, conduct, analysis, interpretation, reporting and dissemination, and receive the information and help they need to contribute in these ways. To reflect this priority we have allocated a budget of £3000 to cover honoraria and expenses.

17. References

[1] Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B. Depression, chronic diseases, and decrements in health. The Lancet. 2007;370(9590):851–858.

[2] Murray C, Lopez A. Alternative projections of mortality and disability by cause 1990-2020. The Lancet. 1997;349(9064):1498–1504.

[3] Layard R, Clark D, Knapp M, Kerr N. Implementing the NICE guidelines for depression and anxiety. A cost-benefit analysis. 2006;.

[4] Frank E, Kupfer DJ, Perel JM, Cornes C. Three-year outcomes for maintenance therapies in recurrent depression. Archives of General Psychiatry. 1990;47(12):1093–1099.

[5] Blom MBJ, Jonker K, Dusseldorp E, Spinhoven P, Hoencamp E, Haffmans J, et al. Combination treatment for acute depression is superior only when psychotherapy is added to medication. Psychotherapy and Psychosomatics. 2007;76(5):289–297.

[6] Joyce PR, Mulder RT, Luty SE, Sullivan PF, McKenzie JM, Abbott RM, et al. Patterns and predictors of remission, response and recovery in major depression treated with fluoxetine or nortriptyline. Australian and New Zealand J of Psychiatry. 2002;36(3):384–391.

[7] Sotsky SM, Glass DR, Shea MT, Pilkonis PA. Patient predictors of response to psychotherapy and pharmacotherapy: Findings in the NIMH Treatment of Depression Collaborative Research Program. The American J of Psychiatry. 1991;148(8):997–1008.

[8] Fournier JC, DeRubeis RJ, Shelton RC, Hollon SD, Amsterdam JD, Gallop R. Prediction of response to medication and cognitive therapy in the treatment of moderate to severe

depression. J of Cons and Clin Psych. 2009;77(4):775–787.

[9] Berlim MT, Turecki G. What is the meaning of treatment resistant/refractory major depression (TRD)? A systematic review. European Neuropsychopharmacology European Neuropsychopharmacology. 2007;17(11):696–707.

[10] Riso LP, Blandino JA, Penna S, Dacey S, Grant MM, Du Toit PL, et al. Cognitive aspects of chronic depression. J of Abnormal Psychology. 2003;112(1):72–80.

[11] Fava M, Farabaugh AH, Sickinger AH, Wright E, Alpert JE, Sonawalla S, et al. Personality disorders and depression. Psychological Medicine. 2002;32(6):1049–1057.

[12] Riso LP, Miyatake RK, Thase ME. The search for determinants of chronic depression: a review of six factors. J of Affective Disorders. 2002;70(2):103–115.

[13] Klein DN, Riso LP, Donaldson SK, Schwartz JE, Anderson RL, Ouimette PC, et al. Family study of early-onset dysthymia. Archives of General Psychiatry. 1995;52(6):487–496.

[14] Candrian M, Schwartz F, Farabaugh A, Perlis RH, Ehlert U, Fava M. Personality disorders and perceived stress in major depressive disorder. Psychiatry Research. 2008;160(2):184–191.

[15] Hayden EP, Klein DN. Outcome of dysthymic disorder at 5-year follow-up: The effect of familial psychopathology, early adversity, personality, comorbidity, and chronic stress. The American J of Psychiatry. 2001;158(11):1864–1870.

[16] McCullough J James P. Treatment for chronic depression: Cognitive Behavioral Analysis System of Psychotherapy (CBASP). New York, NY US: Guilford Press; 2000.

[17] McPherson S, Cairns P, Carlyle J, Shapiro D, Richardson P, Taylor D. The effectiveness of psychological treatments for treatment-resistant depression: a systematic review. Acta Psychiatrica Scandinavica. 2005;111(5):331–340.

[18] Kocsis JH, Gelenberg AJ, Rothbaum BO, Klein DN, Trivedi MH, Manber R, et al. Cognitive behavioral analysis system of psychotherapy and brief supportive psychotherapy for augmentation of antidepressant nonresponse in chronic depression: The REVAMP trial. Archives of General Psychiatry. 2009;66(11):1178–1188.

[19] Murray G, Michalak EE, Axler A, Yaxley D, Hayashi B, Westrin Å, et al. Relief of Chronic or Resistant Depression (Re-ChORD): a pragmatic, randomized, open-treatment trial of an integrative program intervention for chronic depression. Journal of Affective Disorders. 2009;.

[20] Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D. Acute and Longer-Term Outcomes in Depressed Outpatients Requiring One or Several Treatment Steps: A STAR*D Report. Am J Psychiatry. 2006;163(11):1905–1917.

[21] Thase ME, Friedman ES, Biggs MM, Wisniewski SR, Trivedi MH, Luther JF. Cognitive therapy versus medication in augmentation and switch strategies as second-step treatments: A STAR*D report. Am J of Psychiatry. 2007;164(5):739–752.

[22] Eisenberg N, Fabes RA, Guthrie IK, Reiser M. Dispositional emotionality and regulation: Their role in predicting quality of social functioning. J of Personality & Soc Psych. 2000;78(1):136–157.

[23] Constantino MJ, Manber R, DeGeorge J, McBride C, Ravitz P, Zuroff DC, et al. Interpersonal styles of chronically depressed outpatients. Psychotherapy: Theory, Research, Practice, Training. 2008;45(4):491–506.

[24] Lynch T, Cheavens J. Dialectical Behaviour Therapy for Treatment Resistant Depression. New York: Guilford Press; in press.

[25] Lynch TR, Cheavens JS, Morse JQ, Rosenthal MZ. A model predicting suicidal ideation and hopelessness in depressed older adults: The impact of emotion inhibition and affect intensity. Aging & Mental Health. 2004;8(6):486–497.

[26] Lynch T, Robins C, Morse J, Krause E. A mediational model relating affect intensity, emotion inhibition, and psychological distress. Behavior Therapy. 2001;32(3):519–536.

[27] Lynch TR, Cheavens JS, Cukrowicz KC, Thorp SR, Bronner L, Beyer J. Treatment of older adults with co-morbid personality disorder and depression: A dialectical behavior therapy approach. International J of Geriatric Psychiatry. 2007;22(2):131–143.

[28] Morse JQ, Lynch TR. A preliminary investigation of self-reported personality disorders in late life: Prevalence, predictors of depressive severity, and clinical correlates. Aging & Mental Health. 2004;8(4):307–315.

[29] Lynch T, Mendelson T, Robins C, Krishnan K, George L, Johnson C, et al. Perceived social support among depressed elderly, middle-aged, and young-adult samples: cross-sectional and longitudinal analyses. J of Affective Disorders. 1999;55(2-3):159–170.

[30] Smoski MJ, Lynch TR, Rosenthal MZ, Cheavens JS, Chapman AL, Krishnan RR. Decision-making and risk aversion among depressive adults. J Behav Ther Exp Psychiatry. 2008;39(4):567–576.

[31] Smoski MJ, Felder J, Bizzell J, Green SR, Ernst M, Lynch TR, et al. fMRI of alterations in reward selection, anticipation, and feedback in major depressive disorder. J of Affective Disorders. 2009;118(1-3):69–78.

[32] Chapman A, Lynch T, Rosenthal M, Cheavens J, Smoski M, Krishnan K. Risk aversion among depressed older adults with obsessive compulsive personality disorder. Cognitive Therapy and Research. 2007;31(2):161–174.

[33] Lynch TR, Johnson CS, Mendelson T, Robins CJ, Krishnan KRR, Blazer DG. New onset and remission of suicidal ideation among a depressed adult sample. J of Affective Disorders. 1999;56(1):49–54.

[34] Robins CJ, Bagby MR, Rector NA, Lynch TR, Kennedy SM. Comparing dimensional and categorical approaches to the relations of sociotropy and autonomy to symptoms in patients with major depression. Cognitive Therapy and Research. 1997;21:43–58.

[35] Butler EA, Egloff B, Wlhelm FH, Smith NC, Erickson EA, Gross JJ. The social consequences of expressive suppression. Emotion. 2003;3(1):48–67.

[36] NICE. Depression in Adults (update): Depression: the treatment and management of depression in adults. London: National Institute for Health and Clinical Excellence; 2009.

[37] Lynch TR, Trost WT, Salsman N, Linehan MM. Dialectical behavior therapy for borderline personality disorder. Annual Review of Clinical Psychology. 2007;3:181–205.

[38] Lynch TR, Morse JQ, Mendelson T, Robins CJ. Dialectical behavior therapy for depressed older adults: A randomized pilot study. American J of Geriatric Psychiatry. 2003;11(1):33–45.

[39] Harley R, Sprich S, Safren S, Jacobo M, Fava M. Adaptation of dialectical behavior therapy skills training group for treatment-resistant depression. J of Nervous and Mental Disease. 2008;196(2):136–143.

[40] Feldman G, Harley R, Kerrigan M, Jacobo M, Fava M. Change in emotional processing during a dialectical behavior therapy-based skills group for major depressive disorder. Behaviour Research and Therapy. 2009;47(4):316–321.

[41] Stanton A, Danoff-Burg S, Cameron C, Bishop M, Collins C, Kirk S. Emotionally expressive coping predicts psychological and physical adjustment to breast cancer. J of Cons and Clin Psych. 2000;68(5):875–882.

[42] Linehan MM. Cognitive Behavioral Treatment of Borderline Personality Disorder. New York: Guilford Press; 1993.

[43] Porges SW. The polyvagal theory: Phylogenetic substrates of a social nervous system. International J of Psychophysiology. 2001;42(2):123–146.

[44] NICE. Borderline Personality Disorder: Treatment and Management: National Clinical Practice Guideline Number 78. The British Psychological Society and The Royal College of Psychiatrists; 2009.

[45] Schulz K, Altman D, D M, CONSORT. CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomised Trials°. PLoS Medicine. 2010;7(3):e1000251.

[46] Mattei A, Mealli F. Augmented Designs to Assess Principal Strata Direct Effects. 2010;.

[47] Kasch KL, Rottenberg J, Arnow BA, Gotlib IH. Behavioral activation and inhibition systems and the severity and course of depression. J of Abnormal Psychology. 2002;111(4):589–597.

[48] Robins J. Correcting for non-compliance in randomized trials using structural nested mean models. Communications in Statistics - Theory & Methods. 1994;23:2379–2412.

[49] Ten Have T, Joffe M, Lynch K, Brown G, Maisto S, Beck A. Causal mediation analyses with rank preserving models. Biometrics. 2007;63(3):926–934.

[50] Emsley R, Dunn G, White I. Mediation and moderation of treatment effects in randomised controlled trials of complex interventions. Statistical Methods in Medical Research. 2009;p. 096280209105014v1.

[51] Bloom HS. Learning More from Social Experiments: Evolving Analytic Approaches. New York: Russell Sage Foundation; 2005.

[52] Anker MG, Duncan BL, Sparks JA. Using client feedback to improve couple therapy outcomes: a randomized clinical trial in a

naturalistic setting. J Consult Clin Psychol. 2009;77(4):693–704.
[53] Wampold BE. The Great Psychotherapy Debate. Mahwah, NJ: Lawrence Erlbaum Associates; 2001.
[54] Hersoug A, Hoglend P, Monsen J, Havik O. Quality of working alliance in psychotherapy: Therapist variables and patient/therapist similarity as predictors. J of Psychotherapy Practice and Res. 2001;10(4):205.
[55] Hilliard R, Henry W, Strupp H. An interpersonal model of psychotherapy: linking patient and therapist developmental history, therapeutic process, and types of outcome. J of Cons and Clin Psych. 2000;68(1):125–133.
[56] Herman S. Therapist-client similarity on the multimodal Structural Profile Inventory as a predictor of early session impact. J of Psychotherapy Res and Practice. 1997;6:139–144.
[57] Hunt DD, Carr JE, Dagadakis CS, Walker EA. Cognitive match as a predictor of psychotherapy outcome. Psychotherapy. 1985;22(4):718–721.
[58] Dunn G, Maracy M, Dowrick C, Ayuso-Mateos J, Dalgard O, Page H, et al. Estimating psychological treatment effects from a randomised controlled trial with both non-compliance and loss to follow-up. The British J of Psychiatry. 2003;183(4):323.
[59] Folkman S, Moskowitz J. Positive affect and the other side of coping. American Psychologist. 2000;55(6):647–654.
[60] Fredrickson BL. The Role of Positive Emotions in Positive Psychology: The Broaden-and-Build Theory of Positive Emotions. American Psychologist. 2001;56(3):218–226.
[61] Tyrer P, Mitchard S, Methuen C, Ranger M. Treatment rejecting and treatment seeking personality disorders: Type R and Type S. Journal of Personality Disorders. 2003;17(3):263–268.
[62] Linehan MM, Korslund KE. Dialectical Behavior Therapy Adherence Manual; 2003.
[63] Williams J, Kobak K, Bech P, Engelhardt N, Evans K, Lipsitz J, et al. The GRID-HAMD: standardization of the Hamilton Depression Rating Scale. International clinical psychopharmacology. 2008;23(3):120–129.
[64] Keller MB, McCullough JP, Klein DN, Arnow B, Dunner DL, Gelenberg AJ. A comparison of nefazodone, the cognitive behavioral-analysis system of psychotherapy, and their combination for the treatment of chronic depression. NEJM. 2000;342(20):1462–1470.
[65] Leon A, Solomon D, Mueller T, Turvey C, Endicott J, Keller M. The Range of Impaired Functioning Tool (LIFE–RIFT). Psychological medicine. 1999;29(04):869–878.
[66] Beck A, Kovacs M, Weissman A. Assessment of suicidal intention: The Scale for Suicide Ideation. J of Cons and Clin Psych. 1979;47(2):343–352.
[67] Linehan MM. Suicidal Behaviors Questionnaire (SBQ). Seattle: Unpublished work available at the Department of Psychology, Unviersity of Washington; 1981.
[68] Group TE. EuroQol-a new facility for the measurement of health-related quality of life. Health Policy. 1990;16(3):199–208.
[69] Kuyken W, Byford S, Taylor RS, Watkins E. Mindfulness-based cognitive therapy to prevent relapse in recurrent depression. J of Cons and Clin Psych. 2008;.
[70] Attkisson CC, Zwick R. The Client Satisfaction Questionnaire. Evaluation and Program Planning. 1982;5(3):233–237.
[71] Mountford V, Corstorphine E, Tomlinson S, Waller G. Development of a measure to assess invalidating childhood environments in the eating disorders. Eating Behaviors. 2007;8(1):48–58.
[72] Whiteside SP, Lynam DR, Miller JD. Validation of the UPPS impulsive behaviour scale: a four factor model of impulsivity. European J of Personality. 2005;19(7):559–574.
[73] Frost RO, Marten P, Lahart C, Rosenblate R. The dimensions of perfectionism. Cognitive Therapy and Research. 1990;.
[74] Hart CM, Ritchie T, Hepper EG, Gebauer J. The Balanced Inventory of Desirable Responding Short Form: Reliability, validity, and factor structure. submitted manuscript. 2012;.
[75] Letzring TD, Block J, Funder DC. Ego-control and ego-resiliency: Generalization of self-report scales based on personality descriptions from acquaintances, clinicians, and the self. Journal of Research in Personality. 2005;39(4):395–422.
[76] Neuberg SL, Newsom JT. Personal need for structure: Individual differences in the desire for simpler structure. Journal of Personality and Social Psychology. 1993;65(1):113.
[77] McCrae RR, Costa PT. A contemplated revision of the NEO Five-Factor Inventory.

Personality and Individual Differences. 2004;36:587–596.
[78] Wilson EHFJSAWPHkKGYEJ B A. The Cambridge Prospective Memory Test. London: Harcourt; 2005.
[79] Schwartz SH. Universals in the content and structure of values: Theoretical advance and empirical tests in 20 countries. In: Zanna M, editor. Advances in Experimental Social Psychology Vol. 25. San Diego: Academic Press; 1992. p. pp. 1–65.
[80] Hughes J, Barkham M. Scoping the inventory of interpersonal problems, its derivatives and short forms: 1988-2004. Clinical Psychology & Psychotherapy. 2005;12(6):475–496.
[81] Neacsiu AD, Rizvi SL, Vitaliano PP, Lynch TR, Linehan MM. The Dialectical Behavior Therapy Ways of Coping Checklist (DBT-WCCL). J of Clinical Psychology. in press;.
[82] King LA, Emmons RA. Conflict over emotional expression: Psychological and physical correlates. J of Personality & Soc Psych. 1990;.
[83] Wegner D, Zanakos S. Chronic thought suppression. J of Personality. 1994;62(4):615–640.
[84] Sarason I, Sarason B, Sheerin E, Pierce G. A brief measure of social support: practical and theoretical implications. J of Social and Personal Relationships. 1987;4:497–510.
[85] Bond FW, Hayes SC, Baer RA, Carpenter KC, Guenole N, Orcutt HK, et al. Preliminary psychometric properties of the Acceptance and Action Questionnaire–II: A revised measure of psychological flexibility and acceptance. Behavior Therapy. 2011;p. 1–38.
[86] Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: Validity of a brief depression severity measure. J of General Internal Medicine. 2001;16(9):606–613.
[87] Watson D, Clark LA, Tellegen A. Development and Validation of Brief Measures of Positive and Negative Affect - the Panas Scales. J of Personality & Soc Psych. 1988;54(6):1063–1070.
[88] Devilly G, Borkovec T. Psychometric properties of the credibility/expectancy questionnaire. J of Behavior Therapy and Experimental Psychiatry. 2000;31(2):73–86.
[89] Marmar C, Gaston L, Gallagher D, Thompson L. Towards the validation of the California Therapeutic Alliance Rating System. J of Cons and Clin Psych. 1989;1:46–52.
[90] Thompson ER. Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS). J of Cross-Cultural Psychology. 2007;38(2):227–242.
[91] Kessler RC, Barber C, Beck A, Berglund P, Cleary PD, McKenas D, et al. The World Health Organization Health and Work Performance Questionnaire (HPQ). J of Occupational and Environmental Medicine. 2003;45(2):156–174.
[92] Pocock S. *Clinical trials*. Chichester: Wiley; 1983.
[93] Kim DM, Wampold BE, Bolt DM. Therapist effects in psychotherapy. Psychotherapy Research. 2006;16(2):161–172.
[94] Goldstein H. Multilevel Statistical Models. London: Edward Arnold; 2003.
[95] Wooldridge J. Econometric Analysis of Cross Section and Panel Data. Cambridge, MA.: MIT Press; 2002.
[96] Pearl J. Causality. Cambridge: CUP; 2000.
[97] Little RJA, Rubin DB. Statistical Analysis with Missing Data,. Hoboken, NJ: Wiley; 2002.
[98] Muthén L, Muthén B. Mplus user's guide. Los Angeles: Muthén & Muthén; 2007.