A randomised controlled trial of Intensive Short-Term Dynamic Psychotherapy for treatment resistant depression: the Halifax Depression Study

Joel M. Town[1, 2], Allan Abbass [1], Chris Stride[3], Denise Bernier[4]

[1] Centre for Emotions & Health, Department of Psychiatry, Dalhousie University, Halifax, Canada.

[2] Plymouth University Peninsula Schools of Medicine and Dentistry, Plymouth, UK.

[3] The Institute of Work Psychology, University of Sheffield, Sheffield, UK.

[4] Department of Psychiatry, Dalhousie University, Halifax, Canada.

Correspondence concerning this article should be addressed to Dr. Joel Town, Centre for Emotions & Health, Faculty of Medicine, Dalhousie University, 5909 Veterans' Memorial Lane, Halifax, NS, Canada, B3H 2E2. Email: joel.town@dal.ca

Key words: Treatment resistance, Depression, Short-term, Psychodynamic Psychotherapy, Randomised trial

Abstract

Background

While short-term psychodynamic psychotherapies have been shown effective for major depression, it is unclear if this could be a treatment of choice for depressed patients who have not sufficiently responded to existing treatments and commonly have chronic and complex health issues.

Method

This superiority trial used a single blind randomised parallel group design to test the effectiveness of time-limited Intensive Short-Term Dynamic Psychotherapy (ISTDP) for treatment resistant depression (TRD). Patients referred to secondary care community mental health teams (CMHT) who met DSM-IV criteria for major depressive episode, had received antidepressant treatment ≥6 weeks, and had Hamilton Depression Rating Scale (HAM-D) scores of ≥ 16 were recruited. The effects of 20 sessions of ISTDP were judged through comparison against secondary care CMHT treatment as usual (TAU). The primary outcome was HAM-D scores at 6 months. Secondary outcomes included dichotomous measures of both remission (defined as HAM-D score ≤ 7) and partial remission (defined as HAM-D score ≤ 12).

Results

Sixty patients were randomised to 2 groups (ISTDP=30 and TAU=30), with data collected at baseline, 3, and 6 months. Multi-level linear regression modelling showed that change over time on both depression scales was significantly greater in the ISTDP group in comparison to TAU. Statistically significant between-group treatment differences, in the moderate to large range, favouring ISTDP, were observed on both the observer rated (Cohen's d = 0.75) and self-report measures (Cohen's d = 0.85) of depression. Relative to TAU, patients in the ISTDP group were significantly more likely after 6 months to achieve complete remission (36.0% vs. 3.7%) and partial remission (48.0% vs. 18.5%).

Limitations

It is unclear if the results are generalizable to other providers, geographical locations and cultures.

Conclusions

Time-limited ISTDP appears an effective treatment option for TRD, showing large advantages over routine treatment delivered by secondary care services.

Introduction

Depression is among the largest single causes of disability worldwide, and the disease burden of depression is on the rise globally (Murray and Lopez, 1997; WHO, 2013). Anti-depressant treatment is considered a first line agent for major depression. However, up to half of patients do not show a satisfactory response, and up to 20% do not benefit from multiple treatment courses (Kubitz et al., 2013). With every failed treatment the patient's long-term prognosis deteriorates (Rush et al., 2006). Where adequate treatment dose and duration were applied, such cases are termed treatment resistant depression (TRD). Despite a lack of clear guidance on effective treatments when depression does not remit with first line medication (MacQueen et al., 2017; Malhi et al., 2009), it is more common to change medication than commence psychotherapy (Markowitz, 2008). This may be due to insufficient empirical evidence demonstrating the efficacy of psychotherapy for this complex population. A 2011 review of psychological treatments found only 6 trials (Trivedi et al., 2011) and 3 subsequent randomised controlled trials (RCTs) reported mixed findings. In one trial, cognitive behavior therapy as an adjunct to pharmacotherapy in primary care was found to be superior to general practitioner care as usual (Wiles et al., 2013). In another, long-term psychoanalytic psychotherapy was shown to be efficacious when compared to primary care treatment as usual (Fonagy et al., 2015) demonstrating that long-term therapies can contribute to the understanding of TRD. In contrast, a third RCT found no benefit of adding interpersonal psychotherapy to pharmacotherapy and clinical management (Souza et al., 2016).

Short-term psychodynamic psychotherapies (STPPs) are widely used, having been found as effective for reducing depressive symptoms as other first-line psychological treatments (Connolly-Gibbons et al., 2016; Driessen et al., 2013). While the effectiveness of STPPs has been demonstrated for depression (Connolly-Gibbons et al., 2016; Driessen et al., 2015) and in the setting of comorbid personality disorder (Abbass et al., 2011) the relevance of STPPs to patients with TRD is less clear. A pilot study of Intensive Short-Term Dynamic Psychotherapy (ISTDP), one of the STPP models, found preliminary evidence that the TRD sub-population of depressed patients can benefit from this approach (Abbass, 2006).

Given the paucity of empirical support for all psychotherapies with this population, this study formally examines the efficacy of ISTDP for TRD using a controlled trial. To provide an ecologically valid comparison, the effect of ISTDP was compared to that of secondary care Community Mental Health Teams

(CMHTs) which served as a treatment as usual (TAU) control. We hypothesized a linear decrease over time in patients' observer-rated and self-reported depression during the course of treatment and, critically, that this decrease would be greater in patients receiving ISTDP compared to TAU.

Methods

Study Design

This superiority trial used a single blind randomised parallel group design to examine the efficacy of ISTDP, in comparison with secondary care TAU provided by CMHTs, for depressed patients who were non-remitting following at least 1 course of antidepressant treatment. The primary indicator of effectiveness was the reduction of depressive symptoms, measured by the Hamilton Depression Scale (HAM-D) (Hamilton, 1967). The study protocol was registered with ClinicalTrials.gov (ID: NCT01141426) and approved by the Nova Scotia Health Authority Research Ethics Board (NSHA-RS/2013-049).

Sample Size Calculation

Pre-study power calculations utilized information from a pilot study (N=10) for TRD (Abbass, 2006), and meta-analysis (Driessen et al., 2010) with equivalent psychotherapy treatment (STPPs) for depression and using similar outcome instruments. The pilot yielded a very large within-group effect of ISTDP at post-treatment (Cohen's $d \geq 2.0$) but it was not compared to TAU. The meta-analysis suggested a conservative prediction was to estimate a large pre- to post-treatment effect measured by HAM-D (Cohen's $d \geq 1.0$) following ISTDP (Driessen et al., 2015) and small-to-medium (Cohen's $d \approx 0.30$) following TAU (Murray et al., 2010). Therefore, assuming an average between-group effect size of $d = 0.70$, a significance level of alpha = .05 and 2 groups of participants, a sample of 68 participants was required to achieve 80% power. However, due to delays in sample recruitment, only 60 participants were recruited.

Participant Eligibility and Recruitment

The study's CONSORT diagram is presented in figure 1. Between September 2012 and March 2015, 143 potential participants were referred to the study by clinicians from 4 secondary care outpatient CMHTs based in Halifax Regional Municipality, Nova Scotia, Canada. Patients eligible to be referred were outpatients, aged 18-65 years, with a primary diagnosis of major depressive disorder according to DSM-IV criteria.

Of 143 patients referred, 120 could be contacted to conduct a more detailed assessment of their eligibility for the study. DSM-IV diagnoses were assessed at baseline using the Mini Neuropsychiatric Interview-Plus (M.I.N.I. plus) (Sheehan et al., 1992) and the Structured Clinical Interview for DSM-IV Axis II Personality disorders (First et al., 1997) by a research assistant and research psychologist respectively. Both were trained for the study purposes and were blinded to treatment allocation. Patients met study criteria for TRD by having had at least one trial of antidepressants at the adequate recommended therapeutic dose; a current depressive episode duration of 6 or more weeks; inadequate response to treatment (assessed by 17-item HAM-D score ≥ 16); not having started further medication or changed dose of existing medication in the previous 6 weeks; and not having received treatment in the previous 2 years at any of the 4 CMHTs.

Patients were excluded if they had psychotic depression, bipolar depression, severe substance dependence or cognitive impairment, severe cluster A personality disorder, active suicidality or self-mutilating behavior that would require significant modifications to treatment; or if their depressive disorder was due to a general medical condition or secondary to a comorbid mental health or psychosocial condition; or they were unable to give informed consent to treatment.

Information on participants' antidepressant history was obtained from individual general practitioners and through review of medical electronic records. This information was reviewed by a psychiatrist to quantify level of treatment resistance using the Thase and Rush 5-Stage Model (TR-S) (Thase and Rush, 1997) and the Massachusetts General Hospital Staging Model (MGH-S) (Fava, 2003). Of the 120 patients assessed, 60 did not meet inclusion criteria or declined participation and were excluded.

Randomisation and Allocation

Screening assessments and enrollment were conducted by research assistants who remained blind throughout the randomisation and allocation process. Allocation was then conducted at the end of enrollment by an administrative assistant. Patients were allocated to ISTDP or secondary care TAU in a 1:1 ratio (i.e. 30 patients randomly assigned to each group). For purposes of randomisation a researcher external to the study team generated a permuted block randomisation sequence using a digital random number generator.

Intervention Protocol

Intensive Short-Term Dynamic Psychotherapy.

The ISTDP model is a brief psychotherapy format that helps the patient identify and address the emotional factors that culminate into, exacerbate and perpetuate depression. ISTDP is tailored to the patient's anxiety tolerance. Where necessary, the therapist uses a supportive "graded format" to build emotional capacities, and introduces elements to assist the cessation of emotional avoidant behaviors as seen in session (clarification and challenge to defenses). The ultimate goal is an active engaged therapy process that builds the patients' awareness and capacity to experience emotions that adversely affect their mood.

The number of ISTDP sessions provided for common mental health problems varies according to case severity (typically with an upper limit of 40 applied). Once the nature of the difficulties has been established in the extended trial therapy session, the therapist may estimate an approximate number of sessions required, though the exact duration of treatment is not agreed upon. For this study however, ISTDP was provided according to a 20-session time-limited, individual format to allow comparisons with other manualized time-limited psychotherapy models. The treatment was delivered according to a manual and published recommendations (Abbass, 2015; Davanloo, 2000). The first session was an extended 2-3 hour appointment; subsequent hourly sessions were scheduled every week. Any planned termination before 20 sessions was based upon agreement between therapist and patient.

Treatment was delivered by 4 therapists (1 psychiatrist, 3 clinical psychologists) on average treating 7 patients each. The distribution of number of cases seen by individual therapists was 2, 5, 8, 13. Allocation of cases was determined by therapist availability. All clinicians were licensed professionals with supervised experience practicing ISTDP (mean experience = 10.25 years, range= 4-20 years). All therapists participated in a half-day course on TRD and, for the duration of the trial, met on a weekly basis to review and discuss videotaped treatment sessions. Additionally, a study supervisor provided access to weekly case supervision using a review of video recordings of treatment: at least 1 session per quartile of treatment was reviewed.

Medication Management

The study sought to minimize confounding effects of pharmacotherapy treatments that may alter the comparative impacts of ISTDP and TAU, whilst balancing the need to promote treatment adherence. According to the study inclusion criteria of non-remittance of depressive symptoms following at least 1 course of antidepressant, a patient remains symptomatic despite the use of any existing medications. A

medication washout phase was therefore not specified in the protocol prior to treatment allocation to avoid possible withdrawal and subsequent treatment effects.

For participants receiving ISTDP it was requested that medications remain stable. Participants' general practitioners (GP) were contacted to monitor medication, and participants were advised to discuss medication with their GP. Participants in the ISTDP group unable to follow the medication protocol were not excluded, but detailed information was collected on medication changes. For participants receiving TAU, pharmacotherapy was described as one of the possible treatment approaches that could be offered by a CMHT. Pharmacotherapy treatment strategies consisted of individualized regimes informed by Canadian-based clinical guidelines for managing major depression (Lam et al., 2009).

Secondary care treatment as usual (TAU)

At each CMHT site, TAU consisted of a multidisciplinary team approach including pharmacotherapy and clinical management, supportive or structured activities focused around symptom management and, in some cases, individual or group psychotherapy. TAU was not regulated, thus facilitating a naturalistic assessment of standard secondary care treatment delivery - except that trial participants were not offered ISTDP during the trial. Therapeutic interventions were expected to be heterogeneous: hence, details on doses and approaches delivered to each participant were obtained and documented.

Outcome Measures

Outcome assessments were conducted at baseline, mid-treatment (prior to session 11) and post-treatment in the ISTDP group. TAU group patients likewise followed this assessment schedule; however, this occurred at pre-defined time-points selected to match the measurement schedule in the ISTDP group, specifically baseline, 3 months, and 6 months following enrolment.

The primary measure of participants' depressive symptomology was the observer-rated 17-item HAM-D (Hamilton, 1967). A research assistant was trained to use a standardized clinical interview to administer this structured rating system designed to assess symptom severity of individuals with a clinical diagnosis of depression disorder (Bech et al., 2003). During training 3 clinical interviews were rated by the evaluator and ratings compared with an expert's ratings. Research assistants were blind to treatment condition; and to maintain concealment patients were instructed to refrain from discussing their treatment during assessments.

All interviews were videotaped to assess rating reliability. At regular intervals during the study, the evaluator met with an independent experienced evaluator (clinical psychologist) to review and discuss videotaped interviews to prevent rater drift. In total, 26 HAM-D assessments were independently rated by both evaluators and the average intraclass correlation coefficient was 0.96. Internal consistency of the HAM-D was high for 2 of the 3 time-points (time 1, Cronbach's alpha = .617; time 2, alpha = .841; time 3, alpha = .849).

Our secondary outcome was patient-rated depression, measured by the scale sum score from the 9-item Patient Health Questionnaire (PHQ-9) (Kroenke and Spitzer, 2002). Internal consistency of this measure was high at all 3 time-points (alpha = .783, alpha = .899, alpha = .912 respectively). For a series of supplementary analyses we also converted the sum score versions of the HAM-D and PHQ-9 scales to binary outcomes representing full and partial remission. For the purposes of this study, 'being in remission' was defined as achieving a HAM-D score of ≤ 7, and 'at least partial remission' being reflected by a HAM-D score of ≤ 12. For the PHQ-9 measure, remission was defined as achieving a scale sum score of ≤ 4, and at least partial remission by achieving a score ≤ 9. The results from additional secondary outcome measures will be published in a follow-up manuscript.

Treatment Integrity

To assess treatment integrity of the ISTDP intervention, the Comparative Psychotherapy Process Scale (CPPS) (Hilsenroth et al., 2005), a validated measure with distinct psychodynamic-interpersonal (PI) and cognitive behavioral (CB) subscales, was selected for characterizing therapist interventions according to a psychodynamic model compared to a cognitive-behavioral model. For every participant treated using ISTDP, sessions 1, 4, 10 and 16 (representing trial therapy, early, mid and late treatment phases) were watched and evaluated in their entirety by 2 independent researchers (evaluators). Evaluators were blind to the session number. All evaluators attended 24 hours of training in using the CPPS to establish their reliability in scoring; during training, 10 psychotherapy tapes were independently rated by evaluators and scores compared with pre-established expert ratings. All evaluators demonstrated satisfactory inter-rater reliability prior to data collection. In order to minimize the drift in the accuracy of scoring of treatment sessions, evaluators discussed coding with the CPPS instructor at regular intervals. Having independently rated sessions, evaluators participated in a consensus discussion to generate an agreed rating for each session: a third rater was consulted where necessary.

Treatment provided by secondary care CMHTs was selected to provide an ecologically valid TAU condition. Secondary care TAU was therefore unregulated to provide a naturalistic comparison group and therapeutic interventions were expected to be heterogeneous. Due to the range of interventions available, formal treatment integrity ratings were therefore not conducted; however, treatment received was well documented.

Statistical Analyses

We first investigated whether the demographic and clinical characteristics of the sample differed between intervention and control groups, using independent groups t-tests, Mann-Whitney U-tests, or chi-square tests as appropriate, enabling us to identify if we needed to control for any potential confounding demographic effects in the subsequent hypothesis testing.

To compare the impact of ISTDP and TAU on participants' depressive symptoms, as measured by the HAM-D and PHQ-9 scale sum scores across the 3 time-points, longitudinal multilevel linear regression modelling was used (Singer and Willet, 2003). The multilevel modelling approach to longitudinal data, in which data is arranged at the observation (i.e. time-point) level, is advantageous in maximising the sample size, and enabling the distinction between outcome variance to be explained at subject and observation levels as well as modelling within-subject auto-correlations across time. For modelling diagnostic versions of the HAM-D and PHQ-9, logistic generalized estimating equations (GEEs) (Twisk, 2004; Zeger and Liang, 1986), which also model the data at the observation level were used. GEEs were preferred to logistic multilevel models due to their greater stability (Driessen et al., 2013).

For both HAM-D and PHQ-9 sum score outcomes, five models were fitted in sequence, to quantify extent and variation in change in depression over time, and then assess how this variability was explained by treatment differences. First, an unconditional model (with no predictors) assessed baseline levels of within and between-subject variation, and model fit (Model Deviance). Second, time-point (coded 0 = start of treatment, 1 = 3 months after baseline, 2 = 6 months after baseline) was added as a predictor, testing our initial hypothesis of a linear decrease in depression across the study period (having controlled for demographic variables associated with study group, if any were found). Third, a random effect of time-point was added, to estimate between-subject variation in any such change. The potential correlation between intercept and slope was examined and retained if it improved model fit, though given the random allocation into treatment groups there was no expectation that starting level of depression would covary with any

change due to treatment differences. Fourth, the main effect of study group was added. Finally, to assess whether any between-subject variation in change in depression was due to treatment (our primary hypothesis), the interaction between time-point and group was added. An autoregressive type1 (AR1) autocorrelation structure was tested in conjunction to this final model in an attempt to explain any further variation.

As well as testing the fixed interaction effect of time by group (i.e. whether change in depression differs between ISTDP and TAU groups), simple slopes estimates from the final model were tested to ascertain whether the impact of each treatment was different from 'no change' in symptoms. For the supplementary analyses of diagnostic outcomes, a single GEE model was fitted for each outcome, simultaneously testing the main effects of, and interaction between time and treatment group.

SPSSv21 was used, with multilevel analyses performed using the MIXED function, with Maximum Likelihood estimation used to fit the models; sequential models were compared by testing the change in the Deviance i.e. -2*log-likelihood. GEE analyses were performed using the GENLIN function. Mplusv7.4 software was used to calculate simple slopes tests, estimated marginal means, corresponding post-hoc tests and effect sizes. For all analyses, the $p < 0.05$ level of statistical significance was applied; using two-tailed tests, and Bonferroni-corrected where appropriate given the testing of multiple related outcomes. 95% confidence intervals and effect sizes (variance explained at each level) are reported. Analyses and reporting are in accordance with Consolidated Standards of Reporting Trial guidelines (Moher et al., 2010).

Adverse Events

Adverse events during the trial occurred in two participants in the TAU group who reported increases in depressive symptoms. None were judged to be related to study intervention. Statistical tests of difference were not conducted due to the small number of participants with adverse events.

Results

Sample Description and Treatment Completion

The majority of participants in the sample scored within the severe range on the HAM-D alongside comorbid mental health disorders and chronic physical illness; 96.7% had a comorbid Axis I disorder; 89.7%

fulfilled criteria for Axis II personality disorder; 90.9% had current longstanding physical illness or disability. A relatively high rate of participants (63.8%) had failed at least two pharmacotherapy trials from two distinct antidepressant classes for the current major depressive episode. Table 1 summarises demographic and baseline clinical characteristics of the participants. No significant differences in demographic and clinical characteristics were found between the treatment conditions. The 60 randomised participants were included in our primary 'intention to treat' analysis sample. Of the 30 participants randomised to ISTDP, the mean number of sessions completed was 16.1 (SD = 6.68). Twenty-four (72%) participants received at least 15 sessions; 1 completed treatment after 5 sessions due to symptom remission; 2 dropped out of the study immediately after randomisation due to work commitments; 1 dropped out after 1 session, and 2 more dropped out after 8 sessions, again due to work commitments. Table 2 summarises the treatment received within the experimental and comparison condition. In the control group (TAU), 29 participants (97%) received at least one session of talking therapy (mean number of sessions = 7.6, SD = 3.7). The one-to-one therapy delivered by CMHTs was primarily counseling (57%) or CBT (40%). In addition, 15 participants (50%) received group therapy within a structured 10-12 two-hour session CBT format. Participants receiving TAU were more likely ($X^2 = 13.017$, $p < .001$) to have pharmacotherapy increased or changed than those in the ISTDP group (53% vs. 10%). Participants receiving ISTDP were more likely ($X^2 = 26.447$, $p < .001$) to stop or reduce pharmacotherapy treatments (67% vs. 3%)

Due to the aforementioned drop-out and a handful of non-completed measures, our analysis sample consisted of 167 responses from 60 participants, with complete data at time-point 1 (baseline), 55 responses at time-point 2, and 52 at time-point 3. Forty-nine participants responded at all 3 time-points. The missing data was distributed equally between the two groups and there were no significant differences in attrition rates between the groups.

Treatment Integrity

To examine the treatment integrity of the ISTDP intervention, a total of 100 sessions were rated using the CPPS (Hilsenroth et al., 2005); recordings were not available for 3 sessions. Mean inter-rater reliability values (as measured by the intraclass correlation coefficient type 1) for the PI and CB scales were 0.92 and 0.82 respectively, falling in the 'excellent' range (≥ 0.75) according to standardized recommendations (Shrout and Fleiss, 1979).

The consensus rating across treatment sessions on the PI scale (mean = 2.44, SD = 0.89) was greater than for the CB scale (mean = 0.72, SD = 0.26): i.e. consistent with the treatment protocol, ISTDP videotapes were rated as more characteristic of a psychodynamic therapy than a cognitive-behavioral model ($t = 21.40$, $p < .0001$). Furthermore, in line with our a priori expectation that ISTDP sessions would be better characterized by specific PI interventions focused on emotional experiencing rather than interpretative intervention, the highest item mean scores were recorded by PI item-8 ("The therapist encourages the patient to experience and express feelings in the session"); and PI item-1 ("The therapist encourages the exploration of feelings regarded by the patient as uncomfortable e.g., anger, sadness)".

Outcomes

Continuous HAM-D and PHQ-9 measures

Sample mean scores for HAM-D and PHQ-9 sum scores of patients in each study group are provided in figures 2 and 3. For both of these outcomes, introducing a linear effect of time to the model produced a statistically significant improvement in model fit compared to the unconditional model (table 3, model 2). The effect of time was significant and negative (HAM-D: $B = -3.876$, $p < 0.05$; PHQ-9: $B = -3.258$, $p < 0.05$) indicating that the incidence of depressive symptoms decreased over time. Allowing the effect of change over time to vary between subjects (table 3, model 3) further improved the model fit, indicating that participants were showing different rates of improvement in depressive symptoms; however there was no indication that improvement was significantly related with initial level of symptoms (hence the intercept-slope covariance was fixed at zero for subsequent analyses).

Adding the interaction effect between time and study group again significantly improved the model fit (table 3, model 5; adding an autoregressive type1 within subjects correlation structure did not improve model fit, hence this was not retained). Unexplained variation in change over time was reduced by 18.5% for HAM-D, and by 25.4% for PHQ-9. The interaction effect was significant and negative (HAM-D: $B = -2.704$, $p < 0.05$; PHQ-9: $B = -2.677$, $p < 0.05$). For both outcomes, a simple slopes test indicated that the decrease in symptoms was significant for both groups; however, decreases for the ISTDP intervention group was just over twice as steep as for the TAU control group (HAM-D: -5.313 vs. -2.609; PHQ-9: -4.659 vs. -1.982). Post-hoc tests, simple slopes tests and simple effect sizes relating to the predicted between-group differences at each time-point from this final model are presented in table 4.

These results support our primary hypothesis that ISTDP yields a greater improvement in patient's well-being relative to standard treatments. In addition, the assumption of sustained linear change appears more clearly justified for the ISTDP group: the sample mean scores (figures 2 and 3) demonstrate the efficacy of a linear model for decrease in depressive symptoms over the course of treatment with ISTDP, even over this short period of treatment.

Diagnostic HAM-D and PHQ-9 measures

The results found for the continuous versions of the HAM-D and PHQ-9 are mirrored by their diagnostic counterparts, albeit with effects less pronounced (due to the use of blunter dichotomous diagnostic measures as opposed to continuous scales). The sample percentages achieving full and partial remission using the HAM-D or PHQ-9-based diagnostic thresholds varied widely by group (e.g. HAM-D full remission, ISTDP 36% vs. TAU 3.7%; HAM-D partial remission, ISTDP 48.0% vs. TAU 18.5%; PHQ-9 full remission, ISTDP 32.0% vs. TAU 4.3%; PHQ-9 partial remission, ISTDP 60.0% vs. TAU 8.7%). When modelling these diagnostic outcomes, we found significant interactions between treatment group and time-point, with the odds of partial and full remission diagnoses showing a significantly greater increase over time for patients treated with ISTDP as opposed to TAU, thus supporting our hypothesis. Specifically, the odds of partial remission from HAM-D diagnosed symptoms per each additional time-point were over 2.5 times greater for the ISTDP group ($\exp(B) = 2.799$, $p < 0.05$); for partial remission, the odds increased to over 3 times greater ($\exp(B) = 3.373$, $p < 0.05$). The estimated odds of partial remission from PHQ-9 assessed symptoms were almost 5 times greater for the ISTDP treated group for each additional time-point ($\exp(B) = 5.278$, $p < 0.05$). Only for the full remission criterion from PHQ-9 assessed symptoms was a significant difference between treatments not found. Full details of these models are presented in table 5.

Discussion

The Halifax Depression Study RCT examined the efficacy of time-limited ISTDP for secondary care patients presenting with TRD. Improvements in both observer and patient rated depression scores for those receiving ISTDP were greater than for patients receiving TAU. The observed benefits of individuals with TRD receiving ISTDP as a front line alternative to routine treatment also included patients being more likely to reach remission on the primary measure of depression.

The definition of TRD requires participants to have tried at least 1 adequate antidepressant dose for 6 weeks. Like other recent psychotherapy trials (Souza et al., 2016; Wiles et al., 2013), we chose this pragmatic definition to be representative of a heterogeneous and difficult-to-treat population commonly requesting care at secondary care CMHTs. While there is no standard universally accepted definition or criteria of TRD, the number and type of prior antidepressant medication failures may be considered an indication of degree of TRD. Nevertheless, using the less conservative inclusion threshold of at least 1 failed antidepressant, we found that two-thirds of the recruited sample had failed to respond to an adequate trial of pharmacotherapy from two distinct antidepressant classes (Thase and Rush, 1997) and over half (56%) had received past psychotherapy treatment. When the intensity and optimization of each prior failed antidepressant treatment was taken into consideration using the MGH-S staging model (Fava, 2003), degree of resistance score indicated multiple unsuccessful treatment efforts. Furthermore, we were struck by the level of chronicity and complexity found in the current sample: for example, participants reported on average 4 previous depressive episodes, alongside extremely high rates of severe depression (80%), long-term medical illness (90.7%), axis I comorbidity (96.6%) and axis II comorbidity (89.7%).

These observations suggest that in cases of unremitting clinical depression presenting to secondary care services, treatment resistance simply defined based on a categorical approach of at least 1 medication failure, is likely to include patients with multiple risk factors that predict poorer treatment prognosis (Thase et al., 2001). Of note, a significant majority of patients in the current sample fulfilled criteria for personality disorder, a proportion (15%) of which were cluster B personality disorder. The relative efficacy of the treatment protocols evaluated in this study could therefore have been confounded by the use of a mood disorder approach for some participants with symptoms of affective instability and rejection sensitivity (Choi-Kain and Rodriguez-Villa, 2015). These findings highlight the challenges faced in clinical practice and have implications for the design of healthcare services as it suggests that treatment approaches for major depression need also be tailored to personality-disordered patients.

There are few published RCTs examining the effectiveness of psychotherapy for TRD (Trivedi et al., 2011). To our knowledge this is the first using a STPP model. ISTDP outcomes clearly outperformed those from the TAU condition, in which over half of patients changed or augmented antidepressant medications, and many received talking therapy in combination. This result is particularly notable based upon the relatively intensive interventions received in the TAU condition. These findings contribute to the evidence for

psychodynamic treatments for depression. The Halifax Study replicates previous results (Wiles et al., 2013) suggesting that time-limited psychotherapy is an alternative first line treatment strategy for TRD that significantly improves the odds of remission compared to routine CMHT treatments provided. Based on the chronic and often relapsing nature of TRD, we are planning to examine long-term outcomes to assess whether improvement is sustained and maintained. The study conclusions should also consider that the between-group difference in efficacy could be a non-specific treatment effect related to ISTDP participants receiving on average a greater dose of therapy. When the effectiveness of different manualized psychotherapy treatments for TRD has been examined, where the number of available sessions, training and supervision were comparable, no differences been treatments were found (Kocsis et al., 2009). The same is true for other psychiatric disorders when bona-fide psychotherapies are directly compared (Wampold and Imel., 2015). Future research therefore remains necessary in naturalistic settings that compare the delivery of manualised psychotherapies to TAU involving comparable doses of psychotherapeutic services.

With the exception of the Sequenced Treatment Alternative to Relieve Depression (STAR*D) study (Thase et al., 2007), evidence from clinical trials for TRD have previously tested the effects of psychotherapy treatment in combination with antidepressant augmentation/changes. In contrast, the Halifax Study attempted to isolate and study the additional benefit of psychotherapy as a treatment alternative by requesting that existing antidepressant medications remained stable in the ISTDP group. This appeared acceptable to participants receiving ISTDP in that only 1 person requested antidepressant augmentation during treatment and, compared to TAU, significantly more participants stopped or reduced medications during treatment. Furthermore, the 36% full remission rates at the end of ISTDP in the Halifax Study is comparable to the remission rates reported in other trials (21.0-38.5%) (Kocsis et al., 2009; Souza et al., 2016; Thase et al., 2007; Wiles et al., 2013). To test the effectiveness of ISTDP as an augmentation strategy for TRD it should be delivered alongside antidepressants and compared against a control group receiving the same medication algorithm.

Short-term psychodynamic models, such as ISTDP, overlap with Long-Term Psychoanalytic Psychotherapy (LTPP) in the use of common techniques such as focusing on the relationship (transference) between the therapist and patient, and an emphasis on memories and feelings related to the patient's early experiences. The Tavistock Depression Study (TADS) tested LTPP as an adjunct to TAU for TRD (Fonagy et al., 2015). Although remission at the end of the 18-month treatment was uncommon in both groups (e.g., LTPP

remission 9.4%), at 2 years post treatment the observed decrease in depression scores had continued and was significantly greater in the LTPP group (LTPP remission 14.9% at 42-month follow-up). In comparison to the percentage of participants who achieved complete remission after a 6-month ISTDP treatment (36%), these remission rates suggest that symptom change may be possible after relatively fewer sessions of psychodynamic therapy. Such variation in observed rates of response may have implications for service utilization and thus, for the cost effectiveness of treatment. It will be important that future research can assist stakeholders for public health providers in understanding which patients can benefit from brief therapy and which require long-term therapy.

We considered the Halifax Study design a relatively stringent test of the real world effectiveness of the experimental treatment, based on several factors including: the recruitment of refractory patients determined to need secondary care services; random allocation to an active treatment comparison provided by secondary care CMHTs; use of blinded outcome ratings; and a time-limited treatment protocol that also requested no augmentation of current antidepressant medications in the experimental group. While ISTDP therapists' adherence to general psychodynamic technique was measured by trained independent assessors, focal adherence to model specific techniques was not formally examined. Therapists had however all received substantial training in ISTDP and there was ongoing supervision using audio-visual session recordings of treatment in order to promote treatment credibility. The quality of treatment was likely representative of what could be expected from a small but established clinic specializing in this form of talk therapy.

Study limitations include the following issues. First, this is a single blind RCT because therapists and patients could not be blinded to treatment allocation. The knowledge alone of being in receipt or not of the experimental treatment could have influenced outcomes through an expectancy effect. Second, this was a single centre study, with the experimental treatment delivered by one clinical team; thus, it is unclear if the results are generalizable to other providers, geographical locations and cultures. It is noticeable that there is no ethnic diversity within the recruited sample. Third, this project was conducted by proponents of the experimental treatment which increases the likelihood of an allegiance effect (Luborsky et al., 1999), a phenomenon that has been shown to lead to an overestimation of treatment effects in psychotherapy effectiveness research (Baldwin and Imel, 2013). We attempted to minimize this through selecting the observer-rated HAM-D as the primary outcome measure and ensuring administration by an independent assessor blinded to treatment allocation.

Conclusions

Moderate to large effects, favouring ISTDP relative to secondary care mental health team care, were observed following 6 months of treatment in a sample with TRD. Like other treatments for TRD, an unsatisfactory number of patients who received time-limited ISTDP did not reach the desired goal of full remission and required further treatment. Future research needs to better understand why this was the case and if therapy could have been optimized, extended or medically augmented to facilitate remission. It is our intention to explore the degree to which presumed processes underlying therapeutic change in ISTDP (e.g., patient emotional processing) were present in sessions and second, if variation in putative variables predicted change in depressive symptoms.

Role of funding source

This study was funded by the Dalhousie University Department of Psychiatry, a charitable donation provided through Nova Scotia Research Foundation, and ongoing support from Nova Scotia Healthcare Authority. These funding sources had no involvement in the design, implementation, analysis, interpretation of data, or writing of the report.

Conflict of interest

The authors declare that they have no conflicts of interest.

Acknowledgements

The authors thank all of the participants, therapists, research assistants and administrative staff who participated in the study. In particular, we thank Dana Jakeman for her work as lead research assistant. The authors also thank Ellen Driessen, Ph.D (VU University of Amsterdam) for commenting on an earlier draft of this manuscript.

Figure 1. CONSORT Diagram of patient flow through Halifax Depression Study

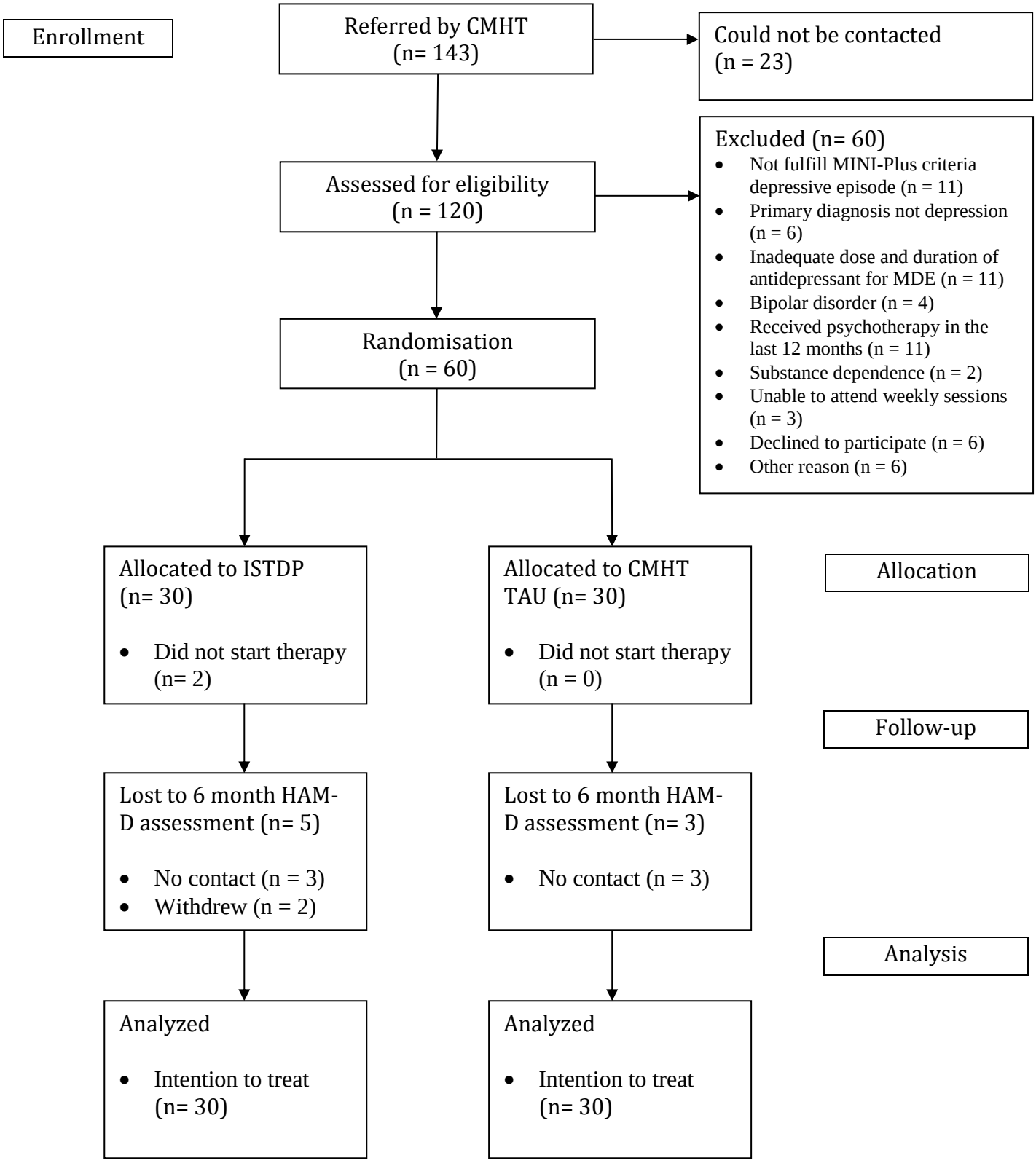


Table 1: Demographic variables and clinical characteristics by experimental group

	ISTDP (N= 30)		TAU (N=30)		TOTAL (N=60)	
Demographic Variables	Mean	SD	Mean	SD	Mean	SD
Age (years)	38.900	11.868	44.200	12.240	41.550	12.248
	N	%	N	%	N	%
Male	13	43.3%	9	30.0%	22	36.7%
Caucasian	30	100.0%	28	93.3%	58	96.7%
Married	14	46.7%	9	30.0%	23	38.3%
Living with one or more person	23	76.7%	21	70.0%	44	73.3%
In employment	21	70.0%	18	60.0%	39	65.0%
University Education	22	73.3%	19	63.3%	41	68.3%
Clinical Characteristics	Mean	SD	Mean	SD	Mean	SD
HAM-D-17 sum score, baseline	23.500	5.309	24.033	5.169	23.767	5.202
PHQ-9 sum score, baseline	19.167	4.639	20.300	4.843	19.733	4.737
N of previous episodes of depression	3.655	1.717	4.200	1.400	3.932	1.574
MGH score	4.810	2.667	4.089	2.077	4.456	2.402
	Median	IQR	Median	IQR	Median	IQR
Duration present depressive episode (mths)	24.0	75.0	36.0	87.0	30.0	81.0
N of different antidepressants previously taken in lifetime	3.0	2.0	3.0	2.5	3.0	2.0
	N	%	N	%	N	%
HAM-D-17 severe/very severe depression (sum score > 18)	24	80.0%	24	80.0%	48	80.0%
PHQ-9 severe depression (sum score > 19)	14	46.7%	18	60.0%	32	53.3%
Any current co-morbid Axis I disorder	28	93.3%	30	100.0%	58	96.7%
Any current co-morbid Axis II disorder†	25	86.2%	27	93.1%	52	89.7%†
Any current co-morbid Cluster B disorder	4	13.3%	5	16.7	9	15.0%
Any current longstanding physical illness or disability††	23	88.5%	27	93.1%	50	90.9%
Any current co-morbid substance use	4	13.3%	3	10.0%	7	11.7%
Lifetime history of receiving psychotherapy	14	46.7%	18	60.0%	32	53.3
Two or more failed antidepressants for current episode‡	12	40%	8	27.6%	20	33.9%
TR-S score‡‡: Stage I	10	33.3%	11	39.3%	21	36.2%
Stage II	17	56.7%	11	39.3%	28	48.3%
Stage III-V	3	10.0%	6	21.4%	9	15.5%

HAM-D = Hamilton Depression Scale. PHQ-9 = Population Health Questionnaire for Depression. MGH = Massachusetts General Hospital staging method (1 point per trial of adequate dose of antidepressant for ≥ 6 weeks; 0.5 point per trial per optimization strategy). TR-S = Thase and Rush staging method (Stage 1: failure of at least 1 adequate trial of 1 major class of antidepressant; Stage 2: Stage 1 resistance plus

failure of adequate trial of an antidepressant in a distinctly different class from that used in Stage I; Stage III-V: Stage II resistance plus at least failure of an adequate trial of a TCA). Axis II disorder data collected using Structured Clinical Interview for DSM-IV Axis II personality disorders
Total N/percentage base = 60, Group N/percentage base = 30, 30, except for:
† N = 58, 29, 29; †† N = 55, 29, 26; ‡ N = 59, 29, 30; ‡‡ N = 58, 28, 30

Table 2: Treatment Received Summary

	ISTDP (N=30) N (%)	TAU (N=30) N (%)	Statistic	p-value
Talking Therapy				
ISTDP (Mean N sessions)	16.1 (6.6)	0	$\chi^2 = 13.473$	<.001***
Other Therapy (Mean N sessions)	0	7.6 (3.7)	$\chi^2 = 11.340$	<.001***
Other Therapy (%)	0	29 (96.7)	$\chi^2 = 56.129$	<.001***
CBT (%)	0	12 (40.0)	$\chi^2 = 15.000$	<.001***
Counselling (%)	0	17 (56.7)	$\chi^2 = 23.721$	<.001***
Group CBT (%)	0	15 (50.0)	$\chi^2 = 20.000$	<.001***
Other Interventions				
Guided Self-help (%)	0	2 (6.7)	$\chi^2 = 2.069$	.492
Peer support group (%)	0	4 (13.3)	$\chi^2 = 4.286$	.112
Meditation group (%)	0	2 (6.7)	$\chi^2 = 2.069$	.492
Healthy living plans (%)	0	3 (10.7)	$\chi^2 = 3.158$	.237
A&E urgent care (%)	0	3 (10.0)	$\chi^2 = 3.158$	.237
Psychiatry (%)	1 (3.3)	8 (26.7)	$\chi^2 = 6.405$	.011*
Medications				
Medications at T1:				
Antidepressants (%)	29 (96.7)	28 (93.3)	$\chi^2 = .351$	.554
Anxiolytics (%)	13 (43.3)	11 (36.7)	$\chi^2 = .278$	.792
Antipsychotics (%)	2 (6.7)	3 (10.0)	$\chi^2 = .218$	.640
Hypnotics (%)	5 (16.7)	4 (13.3)	$\chi^2 = .131$	.718
Medications T1 to T3:				
Medications added/changed (%)†	3 (10.3)	16 (55.2)	$\chi^2 = 13.228$	<.001***
Medications stopped/reduced (%)†	20 (66.6)	1 (3.3)	$\chi^2 = 26.947$	<.001***
T1 N of medications (mean/SD)	1.80 (1.10)	1.77 (.90)	t = .129	.989
T3 N of medications (mean/SD) †	1.24 (1.12)	2.28 (.88)	t = 3.901	.001***

Note: Statistical comparisons of differences between groups were conducted on proportions using the chi-square test and of means with Student t-tests. CBT- cognitive behavioral therapy; ISTDP- intensive short-term dynamic psychotherapy; TAU- treatment as usual
*p<0.05, **p<0.01, ***p<0.001
Total N/percentage base = 60, Group N/percentage base = 30, 30, except for: † N = 58, 29, 29

Table 3: Longitudinal multilevel models for HAM-D and PHQ-9 sum scores - model fit and improvement, fixed and random effects coefficients

		Outcome	
Model	Model fit and improvement (bold); Fixed Effects and 95% CIs; Random Effects and 95% CIs (italics),	HAM-D continuous	PHQ-9 continuous
1	**Model Deviance (-2*Log-likelihood)**	1142.446	1065.665
	Residual (within-subjects) variance	41.681	31.808
	Intercept (between-subjects) variance	21.110	16.573
2	**Model Deviance**	1093.514	1024.050
	Change in Deviance, df	48.932, 1*	41.615, 1*
	Time-point	B = -3.876* (-4.859, -2.894)	B = -3.258* (-4.167, -2.350)
	Residual variance	26.750	21.537
	Intercept variance	25.513	19.554
3	**Model Deviance**	1083.879	1009.434
	Change in Deviance, df	9.635, 1*	14.616, 1*
	Time-point	B = -3.915* (-5.060, -2.770)	B = -3.398* (-4.521, -2.275)
	Residual variance	20.495	15.394
	Intercept variance	17.644	11.392
	Slope (Time) variance	7.673	8.385
4	**Model Deviance**	1081.075	1005.386
	Change in Deviance, df	2.804, 1	4.048, 1*
	Time-point	B = -3.925* (-5.050, -2.799)	B = -3.381* (-4.477, -2.284)
	Treatment (0 = TAU, 1 = ISTDP)	B = -2.364 (-5.115, 0.387)	B = -2.405* (-4.710, -0.100)
	Residual variance	20.984	15.956
	Intercept variance	16.676	10.137
	Slope (Time) variance	6.813	7.367
5	**Model Deviance**	1075.045	999.094
	Change in Deviance, df	6.030, 1*	6.292, 1*
	Time-point	B = -2.609* (-4.105, -1.113)	B = -1.982* (-3.451, -0.512)
	Treatment (0 = TAU, 1 = ISTDP)	B = -0.814 (-3.864, 2.236)	B = -1.155 (-3.709, 1.399)
	Time-point * Treatment	B = -2.704* (-4.854, -0.555)	B = -2.677* (-4.730, -0.624)
	Residual variance	20.387	15.836
	Intercept variance	17.664	10.979
	Slope (Time) variance	5.553	5.499

167 observations from 60 participants

* $p < 0.05$ (2-tailed tests for fixed effects and model improvement, Bonferroni corrected)

Table 4: Group Differences / post-hoc tests for Estimated Marginal Means at Times 1, 2 and 3 (taken from Model 5), and corresponding effect sizes

		Estimated Marginal Means		Estimated mean difference			Effect Size	
Measure	Time	ISTDP	TAU	Difference	SE	Observed SD	Cohen's D†	95% CI
HAM-D	T1	23.031	23.845	0.814	1.533	5.202		
	T2	17.718	21.236	3.518*	1.443	8.075	0.436	(0.085, 0.786)
	T3	12.405	18.627	6.223*	2.029	8.321	0.748	(0.270, 1.226)
PHQ-9	T1	18.855	20.012	1.157	1.276	4.737		
	T2	14.196	18.026	3.830*	1.256	6.881	0.557	(0.199, 0.915)
	T3	9.536	16.040	6.504*	1.911	7.793	0.835	(0.354, 1.315)

†Effect size computed using estimated marginal means, and the SD of the outcome variable from the observed data at the respective time-point.* $p < 0.05$ (2-tailed test, Bonferroni corrected)

Table 5: GEE models for HAM-D and PHQ-9 diagnostic coding - model fit and regression coefficients, with sample percentages by group

		Outcome			
Predictor Variable		HAM-D diagnostic: full remission	HAM-D diagnostic: partial remission	PHQ-9 diagnostic: full remission	PHQ-9 diagnostic: partial remission
Time-point	B	-0.006	0.055	0.591	0.347
	Odds Ratio (exp(B))	0.994	1.056	1.806	1.415
	95% CI for exp(B)	(0.961, 1.030)	(0.598, 1.867)	(0.453, 7.196)	(0.598, 3.349)
Treatment (0 = TAU, 1 = ISTDP)	B	-0.570	-1.492	-2.366	-2.533
	Odds Ratio (exp(B))	0.565	0.225	0.094	0.079,
	95% CI for exp(B)	(0.044, 7.247)	(0.038, 1.332)	(0.001, 12.883)	(0.005, 1.238)
Time-point * Treatment	B	1.216*	1.029*	1.427	1.663*
	Odds Ratio (exp(B))	3.373	2.799	4.167	5.278
	95% CI for exp(B)	(2.443, 4.657)	(1.377, 5.691)	(0.642, 27.058)	(1.701, 16.375)
Model Fit (QIC)		87.667	155.288	77.656	117.558
Sample %s showing remission by time 3	ISTDP group:	36.0	48.0	32.0	60.0
	TAU group:	3.7	18.5	4.3	8.7

167 observations from 60 participants
* $p < 0.05$ (2-tailed test, Bonferroni corrected)

Figure 2: The 17-item HAM-D scale sample mean sum score by treatment group and time-point.

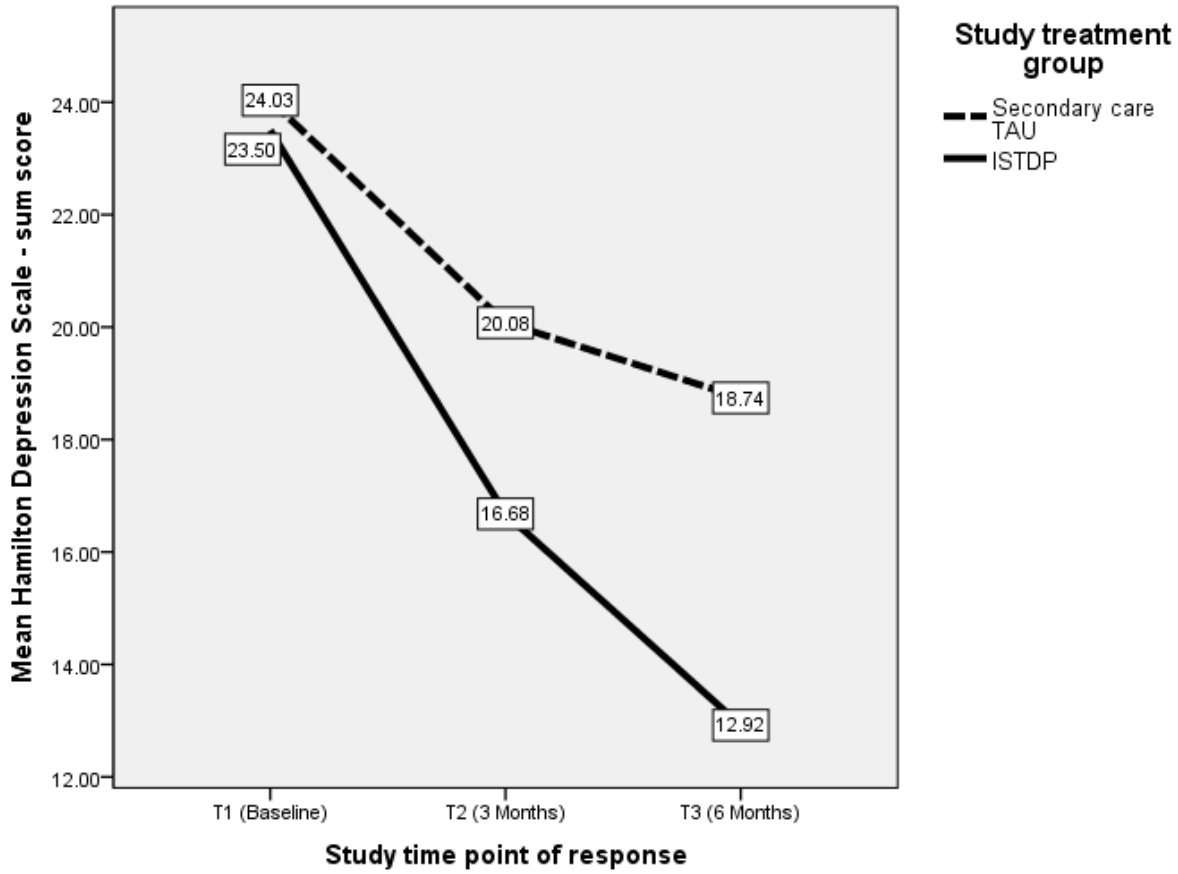
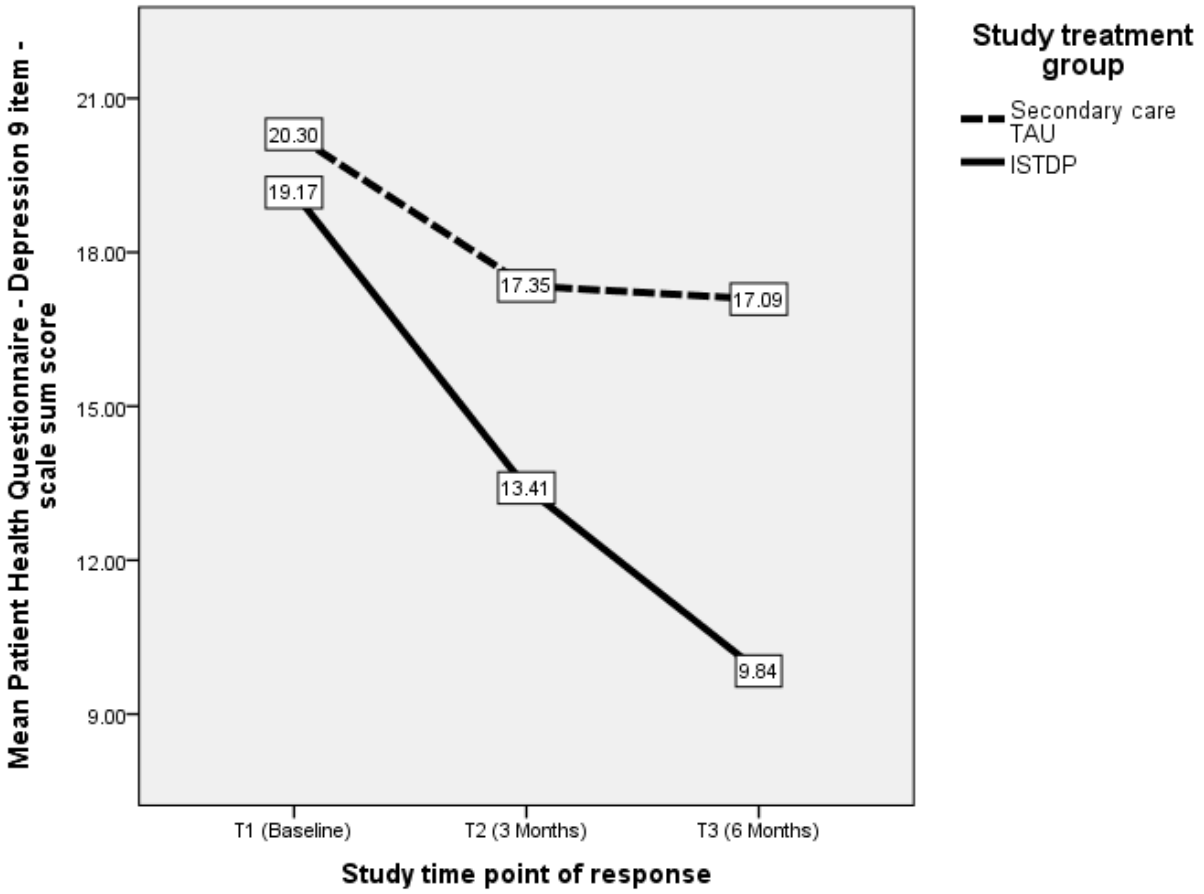


Figure 3: The PHQ-9 scale sample mean sum score by treatment group and time-point.



References

Abbass, A., 2015. Reaching through Resistance. Seven Leaves Press, Kansas City.
Abbass, A., Town, J., Driessen, E., 2011. The efficacy of short-term psychodynamic psychotherapy for depressive disorders with comorbid personality disorder. Psychiatry 74, 58-71.
Abbass, A.A., 2006. Intensive Short-Term Dynamic Psychotherapy of treatment-resistant depression: a pilot study. Depress Anxiety 23, 449-452.
Baldwin, S., Imel, Z.E., 2013. Therapist Effects, In: Lambert, M.J. (Ed.), Bergin and Gafield's handbook of psychotherapy and behaviour change, 6th ed. Wiley, New York.
Bech, P., Engelhardt, N., Evans, K., Kalali, A., Kobak, K., Lipsitz, J., Olin, J., Pearson, J., Rothman, M., Williams, J.B.W., 2003. GRID-HAM-17 GRID-HAM-21 Structured Interview Guide International Society for CNS Drug Development, San Diego, CA, USA.
Choi-Kain, L.W., Rodriguez-Villa, A.M., 2015. Borderline Personality Disorder, Atypical Depression, and Cyclothymia: Diagnostic Distinctions Crossing Mood and Personality Disorders Borders, Borderline Personality and Mood Disorders. Springer, pp. 39-63.
Connolly-Gibbons, M.B., Gallop, R., Thompson, D., Luther, D., Crits-Christoph, K., Jacobs, J., Yin, S., Crits-Christoph, P., 2016. Comparative effectiveness of cognitive therapy and dynamic psychotherapy for major depressive disorder in a community mental health setting: a randomized clinical noninferiority trial. JAMA psychiatry 73, 904-911.
Davanloo, H., 2000. Intensive Short-term Dynamic Psychotherapy. Wiley, Chichester.
Driessen, E., Cuijpers, P., de Maat, S.C., Abbass, A.A., de Jonghe, F., Dekker, J.J., 2010. The efficacy of short-term psychodynamic psychotherapy for depression: a meta-analysis. Clin Psychol Rev 30, 25-36.
Driessen, E., Hegelmaier, L.M., Abbass, A.A., Barber, J.P., Dekker, J.J., Van, H.L., Jansma, E.P., Cuijpers, P., 2015. The efficacy of short-term psychodynamic psychotherapy for depression: A meta-analysis update. Clin Psychol Rev 42, 1-15.
Driessen, E., Van, H.L., Don, F.J., Peen, J., Kool, S., Westra, D., Hendriksen, M., Schoevers, R.A., Cuijpers, P., Twisk, J.W., Dekker, J.J., 2013. The efficacy of cognitive-behavioral therapy and psychodynamic therapy in the outpatient treatment of major depression: a randomized clinical trial. Am J Psychiatry 170, 1041-1050.
Fava, M., 2003. Diagnosis and definition of treatment-resistant depression. Biol Psychiatry 53, 649-659.
First, M.B., Gibbon, M., Spitzer, R.L., Williams, J.B.W., Benjamin, L.S., 1997. Structured clinical interview for DSM-IV axis II personality disorders, (SCID-II) American Psychiatric Press, Inc, Washington, DC.
Fonagy, P., Rost, F., Carlyle, J.A., McPherson, S., Thomas, R., Pasco Fearon, R.M., Goldberg, D., Taylor, D., 2015. Pragmatic randomized controlled trial of long-term psychoanalytic psychotherapy for treatment-resistant depression: the Tavistock Adult Depression Study (TADS). World Psychiatry 14, 312-321.
Hamilton, M., 1967. Development of a rating scale for primary depressive illness The British Journal of Social and Clinical Psychology 6, 278-296.
Hilsenroth, M.J., Blagys, M.D., Ackerman, S.J., Bonge, D.R., Blais, M.A., 2005. Measuring Psychodynamic-Interpersonal and Cognitive-Behavioral Techniques: Development of the Comparative Psychotherapy Process Scale. Psychotherapy: Theory, Research, Practice, Training 42, 340.
Kocsis, J.H., Gelenberg, A.J., Rothbaum, B.O., Klein, D.N., Trivedi, M.H., Manber, R., Keller, M.B., Leon, A.C., Wisniewski, S.R., Arnow, B.A., Markowitz, J.C., Thase, M.E., 2009. Cognitive behavioral analysis

system of psychotherapy and brief supportive psychotherapy for augmentation of antidepressant nonresponse in chronic depression: the REVAMP Trial. Arch Gen Psychiatry 66, 1178-1188.
Kroenke, K., Spitzer, R.L., 2002. The PHQ-9: A new depression diagnostic and severity measure Psychiatric Annals 32, 509-521.
Kubitz, N., Mehra, M., Potluri, R.C., Garg, N., Cossrow, N., 2013. Characterization of treatment resistant depression episodes in a cohort of patients from a US commercial claims database. PLoS One 8, e76882.
Lam, R.W., Kennedy, S.H., Grigoriadis, S., McIntyre, R.S., Milev, R., Ramasubbu, R., Parikh, S.V., Patten, S.B., Ravindran, A.V., 2009. Canadian Network for Mood and Anxiety Treatments (CANMAT) clinical guidelines for the management of major depressive disorder in adults. III. Pharmacotherapy. J Affect Disord 117 Suppl 1, S26-43.
Luborsky, L., Diguer, L., Seligman, D.A., Rosenthal, R., Krause, E.D., Johnson, S., Halperin, G., Bishop, M., Berman, J.S., Schweizer, E., 1999. The researcher's own therapy allegiances: A “wild card” in comparisons of treatment efficacy. Clinical Psychology: Science and Practice 6, 95-106.
MacQueen, G., Santaguida, P., Keshavarz, H., Jaworska, N., Levine, M., Beyene, J., Raina, P., 2017. Systematic Review of Clinical Practice Guidelines for Failed Antidepressant Treatment Response in Major Depressive Disorder, Dysthymia, and Subthreshold Depression in Adults. Canadian journal of psychiatry. Revue canadienne de psychiatrie 62, 11-23.
Malhi, G.S., Adams, D., Porter, R., Wignall, A., Lampe, L., O'Connor, N., Paton, M., Newton, L.A., Walter, G., Taylor, A., Berk, M., Mulder, R.T., 2009. Clinical practice recommendations for depression Acta Psychiatrica Scandinavica 119, 8-26.
Markowitz, J.C., 2008. When should psychotherapy be the treatment of choice for major depressive disorder? Curr Psychiatry Rep 10, 452-457.
Moher, D., Hopewell, S., Schulz, K.F., Montori, V., Gotzsche, P.C., Devereaux, P.J., Elbourne, D., Egger, M., Altman, D.G., 2010. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 340, c869.
Murray, C.J.L., Lopez, A.D., 1997. The Global Burden of Disease. A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries and Risk Factors in 1990 and Projected to 2020. Harvard University Press, Cambridge, Massachussetts.
Murray, G., Michalek, E.E., Axler, A., Yaxley, D., Hayashi, B., Westrin, A., Ogrodniczuk, J.S., Tam, E.M., Yatham, L.N., Lam, R.W., 2010. 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 123, 243-248.
Rush, A.J., Trivedi, M.H., Wisniewski, S.R., Nierenberg, A.A., Stewart, J.W., Warden, D., Niederehe, G., Thase, M.E., Lavori, P.W., Lebowitz, B.D., McGrath, P.J., Rosenbaum, J.F., Sackeim, H.A., Kupfer, D.J., Luther, J., Fava, M., 2006. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. Am J Psychiatry 163, 1905-1917.
Sheehan, D., Janavs, J., Baker, R., Harnett-Sheehan, K., Knapp, E., Sheehan, M., Lecrubier, Y., Weiller, E., Hergueta, T., Amorim, P., Bonora, L.I., Lépine, J.P., 1992. MINI International Neuropsychiatric Interview English version DSM-IV Sheehan, D. V., & Lecrubier, Y.
Shrout, P.E., Fleiss, J.L., 1979. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 86, 420-428.
Singer, J., Willet, J.B., 2003. Applied Longitudinal Data Analysis: Modelling Change and Event Occurrence. Oxford Univeristy Press, NY.

Souza, L.H., Salum, G.A., Mosqueiro, B.P., Caldieraro, M.A., Guerra, T.A., Fleck, M.P., 2016. Interpersonal psychotherapy as add-on for treatment-resistant depression: A pragmatic randomized controlled trial. J Affect Disord 193, 373-380.
Thase, M.E., Friedman, E.S., Biggs, M.M., Wisniewski, S.R., Trivedi, M.H., Luther, J.F., Fava, M., Nierenberg, A.A., McGrath, P.J., Warden, D., Niederehe, G., Hollon, S.D., Rush, A.J., 2007. Cognitive therapy versus medication in augmentation and switch strategies as second-step treatments: a STAR*D report. Am J Psychiatry 164, 739-752.
Thase, M.E., Rush, A.J., 1997. When at first you don't succeed: sequential strategies for antidepressant nonresponders. J Clin Psychiatry 58 Suppl 13, 23-29.
Trivedi, R.B., Nieuwsma, J.A., Williams, J.W., Jr., 2011. Examination of the utility of psychotherapy for patients with treatment resistant depression: a systematic review. J Gen Intern Med 26, 643-650.
Twisk, J.W., 2004. Longitudinal data analysis. A comparison between generalized estimating equations and random coefficient analysis. Eur J Epidemiol 19, 769-776.
WHO, 2013. Mental health action plan 2013-2020. WHO, Geneva, Switzerland.
Wiles, N., Thomas, L., Abel, A., Ridgway, N., Turner, N., Campbell, J., Garland, A., Hollinghurst, S., Jerrom, B., Kessler, D., Kuyken, W., Morrison, J., Turner, K., Williams, C., Peters, T., Lewis, G., 2013. Cognitive behavioural therapy as an adjunct to pharmacotherapy for primary care based patients with treatment resistant depression: results of the CoBalT randomised controlled trial. Lancet 381, 375-384.
Zeger, S.L., Liang, K.Y., 1986. Longitudinal data analysis for discrete and continuous outcomes. Biometrics 42, 121-130.