Supplement 1: Replication with the Beck Depression Inventory-II

A mixed-effects growth model including linear and quadratic slopes for time, treatment contrast (DIT *v.* LIT, DIT *v.* CBT), covariates (age, sex, higher education status, marital status, ethnicity, income bracket, co-occurring medical problems, and assessment variability), a random intercept for patient, and random linear slope for time, fit the data well (fixed portion: $\chi^2(17) = 105.65$, $p < .001$: random portion: $\chi^2(2) = 55.493$, $p < .001$). Patients varied in their baseline scores (random intercept = 30.50, 95% CI [17.68, 52.60]) and slopes (random slope = 18.23, 95% CI [9.31, 35.69]).

BDI-II scores showed a linear decline over time ($B = -5.35$, $z = -2.11$, $p = .034$, 95% CI [-10.32, -0.39]). Furthermore, the LIT group showed a marginally stronger quadratic pattern of change (i.e. U-shaped) compared to the DIT group ($B = 3.74$, $z = 1.95$, $p = .051$, 95% CI [-0.02, 7.49]; see Supplementary Figure 1a). A separate model which included all four assessment phases for the DIT group only (e.g., baseline, mid-treatment, post-treatment, and follow-up) showed that there was no significant difference between marginal means at post-treatment and follow-up (post-treatment = 16.7, follow-up = 16.0; $\chi^2(1) = 0.11$, $p = .740$, 95% CI [-4.40, 3.13]).

At post-treatment, the DIT group scored significantly lower on the BDI-II than the LIT group (16.7 vs. 25.5; $\chi^2(1) = 9.24$, $p = .002$, 95% CI [3.12, 14.44]; $d = .71$, 95% CI [0.25, 1.17]). There were no significant differences between the DIT and CBT groups in post-treatment marginal means (DIT = 16.7, CBT = 20.7, $\chi^2(1) = 1.29$, $p = .257$, 95% CI [-2.95, 11.04]; $d = .33$, 95% CI [-.24, .90]).

A logistic regression model with treatment contrast (DIT *v.* LIT, DIT *v.* CBT), baseline BDI-II scores and covariates (age, sex, higher education status, marital status, ethnicity, income bracket, co-occurring medical problems and end-point assessment variability) showed that compared to LIT patients, more DIT patients achieved clinically

significant change (marginal percentages: 50% vs. 0%; no chi-square test was available due to the lack of variability in the LIT group) and reliable improvement (68% vs. 31%; $\chi^2(1) = 8.81$, $p = .003$, 95% CI [-.62, -.13]; RR = 1.90, 95% CI [1.16, 3.10]). Moreover, the DIT group showed significantly fewer cases of no reliable change compared to the LIT group (30% vs. 57%; $\chi^2(1) = 3.61$, $p = .05$, 95% CI [.01, .54]; RR = 1.88, 95% CI [0.93, 3.77]). There were no differences between the marginal proportions of DIT and CBT patients who achieved clinically significant or reliable change (see Supplementary Table 4).

Supplement 2: Multiple Imputation

Imputation Models

We imputed the primary outcome (HRSD-17) and secondary outcomes (ECR-R, BSI, EQ-5D health status measure, IIP and SAS) separately. Fifty datasets were imputed by chained equations for HRSD-17 scores and assessment times at mid-treatment and post-treatment as well as baseline covariates with missing data, using baseline HRSD-17 and BDI-II scores, age, sex, and imputed data (e.g., imputed covariates and HRSD-17 scores at mid-treatment were used to impute HRSD-17 scores at end-of-treatment). We also included a dummy variable for randomization site to represent the multilevel structure of the data. Covariates were imputed first, followed by the HRSD-17. Multiple Imputation by Chained Equations (MICE) does not assume a joint multivariate distribution among imputed variables, making it suitable for imputing both continuous and categorical variables (White, Royston, & Wood, 2011).

Fifty datasets were imputed also by chained equations for each secondary outcome subscale at baseline and post-treatment. As a rule of thumb, the number of imputed datasets should match the rate of missingness (Graham, 2009). Hence, 50 datasets are suitable for reducing bias at ~50% missingness. Subscales within a measure were imputed together (and not with subscales from other measures) to avoid collinearity issues. Predictors included baseline HRSD-17 and BDI-II scores, age, sex, site, and imputed data (e.g., imputed scores for a given subscale at baseline were used to impute scores on the same subscale at post-treatment). Estimates from each imputed data set were combined using Rubin's rules (Rubin, 1987).

Missing Data Mechanisms

Primary Outcome

We ran three sensitivity analyses to approximate the missing data mechanisms. The first analysis involved logistic regressions predicting the probability of missingness on the primary outcome (HRSD-17) or covariates (assessment variability, higher education status, marital status, ethnicity, income bracket, and co-occurring medical problems) at mid-treatment and post-treatment (we report logits as betas). Predictors included baseline HRSD-17 and BDI-II scores, age, sex, and treatment group (all of which had full data). No variables significantly predicted the probability of being missing on any covariate or on the HRSD-17 at mid- or post-treatment, but age marginally predicted missingness on the HRSD-17 at post-treatment (B = .03, z = 1.92, p = .054, 95% CI [.00, .06]).

The second sensitivity analysis involved Little's Missing Completely at Random (MCAR) test. There were no systematic missing data patterns on the HRSD-17 ($\chi^2(2) = 1.59$, p = .451), but the covariates showed a systematic missing data pattern ($\chi^2(25) = 21.78$, p = .026), which was largely monotone (e.g., the most frequent missing data pattern involved the same participants showing missing observations across covariates).

The final sensitivity analysis involved a comparison between the observed and imputed results. Supplementary Table 5 demonstrates that the significant regression coefficients and marginal means for the observed and imputed analyses are largely similar. The marginal proportions differ slightly between analyses, but the direction and significance of differences is similar, except for the contrast between DIT and CBT groups in the proportion of patients achieving clinically significant change, which is no longer significant.

The three sensitivity analyses suggest that the mechanisms underpinning missingness on the HRSD-17 are, at the very least, missing at random. While there may be unobserved covariates that predict the occurrence of missingness on the HRSD-17, the fact that the

HRSD-17 itself showed no association with missingness suggests that there is unlikely to be a systematic bias in the missing data mechanism (i.e. Not Missing at Random) that would preclude the use of multiple imputation, or indeed the use of all available cases in an intent-to-treat analysis. This is further bolstered by the results from Little's MCAR test, which did not suggest any systematic basis to missing data patterns on the HRSD-17. The covariates did show a systematic pattern of missingness, but this was not predicted by baseline depression severity. Furthermore, missing data rates on the covariates were minimal.

Secondary Outcomes

We applied the three sensitivity analyses described above to the secondary outcomes (ECR-R, BSI, EQ-5D health status measure, IIP and SAS). First, logistic regressions predicting the probability of missing on each secondary outcome subscale at mid-treatment and post-treatment showed that no subscales were predicted by baseline HRSD-17 and BDI-II scores, sex, or treatment group. However, age significantly predicted missingness on the extended family subscale (B = .04, z = 2.37, p = .018, 95% CI [0.01, .07]), social (friends) subscale (B = .03, z = 1.91, p = .057, 95% CI [-0.01, .06]), and total social problems scale (B = .03, z = 1.91, p = .057, 95% CI [-0.01, .06]) of the SAS.

Little's MCAR test confirmed the logistic regression results: aside from the social avoidance subscale ($\chi^2(2) = 6.54$, p = .038) and overly nurturing subscale ($\chi^2(2) = 6.08$, p = .048) of the IIP, the SAS was the only scale to show systematic missing data patterns, including Work (χ^2 Distance $(2) = 10.00$, p = .007), Household (χ^2 Distance $(2) = 7.02$, p = .029), and Total scores (χ^2 Distance $(2) = 11.60$, p = .003), the latter of which can be explained, in part, by variation in age. Lastly, the significant estimates were largely similar between imputed and observed analyses (see Supplementary Table 2).

The secondary outcomes had a substantial amount of missing data at baseline which precluded their use in predicting missingness at post-treatment. Therefore, while the

sensitivity analyses generally suggest that cases were missing at random, they should be treated with caution.

Replication of the Main Analysis with Imputed Datasets

We re-ran the mixed-effects models reported in the main analysis using an imputed dataset to ensure that the analysis of all available cases was replicable. Note that the --mimrgns-- Stata module for computing marginal estimates from imputed datasets does not compute chi-square values due to the uncertainty in combining estimates. We thus determine significance by confidence intervals. Furthermore, we do not report marginal risk ratios as they were computed from chi-square values.

As was found in the intent-to-treat analysis, HRSD-17 scores showed a linear decline over time (B = -3.65, t = -2.76, p = .006, 95% CI [-6.24, -1.05]). Moreover, the LIT group showed a stronger quadratic pattern of change (i.e. U-shaped) compared to the DIT group (B = 2.41, z = 2.24, p = .026, 95% CI [0.30, 4.53]). A separate model which included all four assessment phases for the DIT group only (e.g., baseline, mid-treatment, post-treatment, and follow-up) showed that there was no significant difference marginal means at post-treatment and follow-up (post-treatment = 10.7, follow-up = 11.7; contrast = 1.04, 95% CI [-2.88, 5.35]). At post-treatment, the DIT group showed significantly lower marginal means than the LIT group (10.7 vs. 14.4; contrast = 3.72, 95% CI [0.52, 6.91]; d = 54, 95% CI [0.09, 1.00]). There were no significant differences in post-treatment marginal means between the DIT and CBT groups (10.7 vs. 11.5, contrast = 0.79, 95% CI [-2.82, 4.40], d = .13, 95% CI [-.44, .69]).

Logistic regression models showed that more DIT patients achieved clinically significant change than LIT patients (marginal percentages: 42% vs. 13%; contrast = -.29, 95% CI [-.47, -.10]). Nonetheless, the difference in the proportion of DIT and CBT patients achieving clinically significant change was no longer significant (42% vs. 32%; contrast = -

.09, 95% CI [-.34, .15]). All treatment groups showed moderate-to-high levels of reliable improvement, moderate levels of no reliable change, and low levels of deterioration (see Supplementary Table 5).

Supplement 3: Collapsing the DIT Sites

We ran a series of sensitivity analyses to ensure that collapsing the DIT sites over the pilot trial and feasibility study was statistically plausible. First, we tested whether there was any significant variation in HRSD-17 scores associated with the randomization site and whether this interacted with the treatment group. Adding a level-3 random intercept for randomization site significantly improved the two-level model with linear and quadratic fixed slopes for time, treatment contrasts, and a random intercept and slope for patient and time at level-2, respectively ($\chi^2 = 4.38$, $p = .036$). However, the amount of variance in HRSD-17 scores explained by randomization site was minimal (random intercept = 1.14, 95% CI [0.15, 8.90]), and was reduced to near-zero when covariates were added to the model (hence why we used two-level models in the main analysis that included covariates).

Adding a random slope for treatment contrast at level-3 did not improve model fit (both random slopes were equal to zero; $\chi^2(1) = 0$, $p = 1.00$). Therefore, initial differences between DIT and CBT, or DIT and LIT, were not more or less pronounced at different sites. In a separate three-level model, we tested a random slope for time at level 3, but this too did not improve the model (random slope = .08, 95% CI [0.00, 178.27]; $\chi^2(1) = .09$, $p = .770$). Therefore, different sites were not significantly associated with weaker or stronger changes over time across treatment groups. Finally, adding both level-3 random slopes for treatment group and time to a single model did not improve its fit compared to including only one random slope ($\chi^2(2) = .09$, $p = .960$), and the model was not further improved by estimating a random interaction slope between treatment group and time (random slopes for each interaction contrast were zero; $\chi^2(6) = 0$, $p = 1.00$). In other words, treatment differences between DIT and CBT, or DIT and LIT, over time were not differentially associated with randomization site.

We also compared the post-treatment marginal means and slopes for DIT patients in

the pilot trial and feasibility study. A mixed-effects model including the fixed effects of time (linear and quadratic slopes), randomization group (pilot trial *v.* feasibility study), time-by-randomization group interactions, and random effects of patient and time, showed that there were no differences in the linear slope (B = -.80, z = -0.34, p = .735, 95% CI [-5.44, 3.84]) or quadratic slope (b = .14, z = 0.13, p = .898, 95% CI [-2.07, 2.36]) between randomization groups. Furthermore, marginal end-point means were similar across randomization groups (pilot trial = 9.6, feasibility study = 10.5; $\chi^2(1) = 0.22$, p = .641, 95% CI [-2.92, 4.75]). DIT patients from each randomization group also showed similarities in age (pilot trial = 40, feasibility study = 38; $t(70) = -0.33$, p = .741, 95% CI [-7.98, 5.70]), sex (pilot trial = 55%, female, feasibility study = 72%; $x(1) = 1.83$, p = .176), marital status (pilot trial = 60% single, feasibility study = 71%; $x(1) = 0.85$, p = .355), higher education status (pilot trial = 65% attended higher education, feasibility study = 65%; $x(1) = 0.00$, p = .974), income bracket (pilot trial = 32% low household income, 42% medium household income; feasibility study = 46% low, 36% medium; $x(1) = 1.22$, p = .544), co-occurring medical problems (pilot trial = 47%, feasibility study = 44%; $x(1) = 0.07$, p = .788), but not ethnicity (pilot trial = 100% Caucasian, feasibility study = 71% Caucasian; $x(1) = 7.35$, p = .007).

Collectively, these sensitivity analyses demonstrate that the DIT patients from different randomization sites and groups were comparable in treatment effects and demographics, supporting their collapse.

Supplement 4: Excluding Mild Cases

Our sample included mild cases because 52 patients (35%) scored in the mild range on the HRSD-17, but in the moderate-to-severe range on the BDI-II. We did not exclude these patients because both clinician- and patient-reported depression outcomes provide valuable information (Uher et al., 2012). Instead, we re-ran the main analysis whilst excluding these cases to determine whether they had an impact on the results. Supplementary Table 6 demonstrates that estimates between the standard analysis (N = 147) and analysis with mild cases excluded (N = 95) were similar. Post-treatment marginal estimates increased slightly in the analysis excluding mild cases (as would be expected), but group differences remained significant and of a medium strength.

References

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual review of psychology*, *60*, 549-576.

Uher, R., Perlis, R. H., Placentino, A., Dernovšek, M. Z., Henigsberg, N., Mors, O., ... & Farmer, A. (2012). Self-report and clinician-rated measures of depression severity: can one replace the other?. *Depression and anxiety*, *29*(12), 1043-1049.

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: issues and guidance for practice. *Statistics in medicine*, *30*(4), 377-399.

Supplementary Table 1. *Regression coefficients for covariates in the main analysis (mixed-effects model with HRSD-17 scores as the outcome) using either available or imputed data.*

Covariate	*B*	*z*	*p*	95% CI
Intention-to-Treat				
Age	-.04	-1.23	.218	[-0.11, 0.03]
Assessment time	-.04	-1.83	.067	[-0.09, 0.00]
Sex	1.66	2.05	**.040**	**[0.07, 3.24]**
Ethnicity	0.88	0.96	.339	[-0.93, 2.69]
Marital Status	1.36	1.64	.101	[-0.27, 2.99]
Income Bracket				
£20,000-50,000	-0.98	-1.15	.249	[-2.63, 0.68]
£50,000-100,000+	-2.45	-2.22	**.027**	**[-4.62, -0.28]**
Higher Education Status	1.35	1.45	0.146	[-0.47, 3.1]
Medical Problems	0.25	0.32	0.751	[-1.28, 1.77]
Multiple Imputation				
Age	-.03	-1.00	.316	[-0.10, 0.03]
Assessment time	-.03	-1.21	.231	[-0.09, 0.02]
Sex	1.62	2.03	**.042**	**[0.06, 3.18]**
Ethnicity	1.78	2.05	**.040**	**[0.08, 3.48]**
Marital Status	1.61	2.02	**.044**	**[0.04, 3.17]**
Income Bracket				
£20,000-50,000	-0.91	-1.05	.295	[-2.63, 0.80]
£50,000-100,000+	-1.88	-1.68	.093	[-4.08, 0.32]
Higher Education Status	1.85	2.16	**.031**	**[0.17, 3.52]**
Medical Problems	0.83	1.12	.264	[-0.62, 2.28]

Note. *B*, unstandardized beta, CI, confidence interval**.** Contrasts include: sex (0 = Male, 1 = Female), ethnicity (0 = Caucasian, 1 = all other categories, including black, mixed race, Asian, other), Marital Status (0 = Married or Cohabiting, 1 = all other categories, including single, divorced, widowed, or separated), income bracket (0 = <£10,000-20,000, 1 = £20,000-50,000, 2 = £50,000-100,000+), higher education status (0 = attended higher education, 1 = no higher education reported), medical problems (0 = no co-occurring health issues, 1 co-occurring health issues).

Supplementary Table 2. *Comparison of significant findings for the secondary outcomes between intention-to-treat (IIT) and multiple imputation (MI) analyses*

Subscale	Marginal Mean (*SE*)			Contrast [95% CI]	
	LIT	DIT	CBT	DIT *v.* LIT	DIT *v.* CBT
EQ-5D					
IIT	61.15 (2.74)	71.51 (2.16)	67.31 (3.76)	**-9.86 [-16.90, -2.12]****	-4.20 [-12.86, 4.45]
MI	60.14 (2.44)	69.66 (2.03)	64.92 (4.05)	**-9.52 [-15.81, -3.24]****	-4.75 [–13.83, 4.34]
BSI Depression					
IIT	1.84 (0.19)	1.01 (0.16)	1.20 (0.29)	**0.82 [0.30, 1.34]**	0.18 [-0.50, 0.87]
MI	1.87 (0.18)	1.03 (0.15)	1.20 (0.25)	**0.85 [0.37, 1.32]*****	0.18 [–0.40, 0.75]
BSI Anxiety	1.51 (0.17)	0.93 (0.13)	0.83 (0.24)	**0.57 [0.15, 1.00]**	-0.10 [-0.66, 0.47]
IIT	1.52 (0.19)	1.00 (0.13)	0.98 (0.23)	**0.52 [0.06, 0.98]***	-0.02 [–0.55, 0.51]
MI					
BSI Psychoticism					
IIT	1.14 (0.14)	0.67 (0.11)	0.80 (0.20)	**0.47 [0.12, 0.82]****	0.13 [-0.33, 0.59]
MI	1.17 (0.13)	0.77 (0.11)	0.81 (0.17)	**0.41 [0.07, 0.75]***	0.05 [–0.36, 0.46]
BSI OC					
IIT	2.07 (0.20)	1.26 (0.15)	1.57 (0.27)	**0.80 [0.31, 1.29]****	0.21 [-0.45, 0.87]
MI	1.92 (0.19)	1.24 (0.15)	1.46 (0.28)	**0.68 [0.19, 1.18]****	0.31 [–0.32, 0.93]
BSI GSI					
IIT	1.31 (0.14)	0.86 (0.11)	0.96 (0.20)	**0.45 [0.10, 0.80]****	0.11 [–0.35, 0.57]
MI	1.35 (0.14)	0.90 (0.10)	1.01 (0.18)	**0.46 [0.11, 0.81]****	0.11 [–0.30, 0.52]
BSI PSDI					
IIT	2.11 (0.14)	1.61 (0.10)	1.80 (0.20)	**0.50 [0.15, 0.85]****	0.19 [–0.27, 0.65]
MI	2.13 (0.14)	1.61 (0.10)	1.72 (0.17)	**0.52 [0.17, 0.88]****	0.12 [–0.27, 0.51]

SAS Work					
IIT	2.29 (0.12)	1.89 (0.10)	2.05 (0.17)	**0.40 [0.08, 0.72]***	0.16 [-0.25, 0.57]
MI	2.36 (0.11)	1.99 (0.09)	2.33 (0.17)	**0.37 [0.09, 0.65]****	0.35 [-0.04, 0.73]
SAS Social					
IIT	2.78 (0.12)	2.30 (0.10)	2.40 (0.17)	**0.49 [0.18, 0.80]****	0.10 [–0.31, 0.51]
MI	2.72 (0.11)	2.39 (0.09)	2.50 (0.14)	**0.33 [0.07, 0.60]****	0.11 [–0.23, 0.46]
SAS Total					
IIT	2.50 (0.08)	2.24 (0.07)	2.50 (0.12)	**0.25 [0.05, 0.47]***	0.26 [–0.01, 0.53]
MI	2.45 (0.09)	2.19 (0.07)	2.44 (0.13)	**0.26 [0.03, 0.49]***	0.25 [–0.06, 0.56]

Note. B, unstandardized regression coefficient; *SE*, standard error, *CI*, confidence intervals, *CBT*, Cognitive-behavioural therapy; *DIT*, Dynamic Interpersonal Therapy; Low-intensity Treatment; BSI, Brief Symptom Inventory; OC, Obsessive-compulsive; GSI, Global Severity Index; PSDI, Positive Symptom Distress Index; SAS, Social Adjustment Scale.
Significant results are in bold: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Supplementary Table 3. *Descriptive statistics for Dynamic Interpersonal Therapy (DIT) therapists and treatments*

	Mean	*SD*	Range
Competency (DIT Adherence Scale)			
Overall (*n* = 52)	53.4	10.6	19–65
Initial (*n* = 66)	54.7	9.9	12–69
Mid (*n* = 60)	52.9	12.8	19–69
Late (*n* = 53)	52.1	13.2	20–70
Sessions			
Sessions offered (*n* = 72)	13.4	5.1	1–18
Sessions attended (*n* = 72)	11.8	5.6	0–17
Therapists in the trial (*n* = 17)			
Number of patients seen	4.6	2.9	2–12
Number of sessions undertaken	11.8	5.61	1–17
Number of sessions rated for competence	2.5	1.00	1–3

SD, Standard deviation. Initial phase included first four sessions; mid phase included sessions 5–12 (approximately), and late phase included sessions 13–16 (approximately).

Supplementary Table 4. *Marginal means and percentages for the BDI-II for each treatment group at each assessment point*

	Marginal Means (*SE*)			Contrast [95% CI]	
Outcome measure	LIT	DIT	CBT	DIT *v.* LIT	DIT *v.* CBT
Mean Score	32.61 (1.29)	32.72 (1.08)	33.30 (1.99)	-0.11 [–3.42, 3.21]	0.58 [-3.87, 5.02]
Baseline	26.63 (1.60)	26.03 (1.36)	27.11 (2.39)	0.60 [–3.54, 4.73]	1.08 [–4.31, 6.46]
3 months	25.46 (2.28)	16.68 (1.76)	20.73 (3.10)	**8.78 [3.12, 14.44]****	4.05 [–2.95, 1.04]
6 months		16.05 (1.99)			
12 months					
RCI					
CSC[a]	0% (0)	50% (.06)	41% (.13)	N/A	-9% [-0.39, 0.20]
Improvement	31% (.10)	68% (.07)	64% (.12)	**-38% [-0.62, -0.13]****	-4% [-0.32, 0.24]
No Change	57% (.12)	30% (.07)	38% (.13)	26% [-0.01%, 0.54]	8% [-0.21, 0.36]
Deterioration[a]	10% (.05)	2% (.02)	0% (0%)	8% [-0.02, 0.19]	N/A

Note. SE, Standard error; CI, confidence interval; LIT, low-intensity therapy; DIT, Dynamic Interpersonal Therapy; CBT, cognitive-behavioural therapy; RCI, Reliable Change Indices; CSC, clinically significantly change.
Outcomes at 12 months were not collected for the LIT and CBT groups.
[a]Chi-square difference testing could not be performed due to lack of variation in at least one of the groups estimates.
Significant results are in bold: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Supplementary Table 5. *Comparison of the significant regression coefficients and marginal estimates between intention-to-treat (IIT) and multiple imputation (MI) analyses of the HRSD-17*

	B		*z*		*p*		95% CI	
Predictor	IIT	MI	IIT	MI	IIT	MI	IIT	MI
Time (linear)	-3.81	-3.65	-3.11	-2.76	.002	.006	[-6.21, -1.41]	[-6.24, -1.05]
Time (Quadratic) x Group (DIT *v.* LIT)	2. 57	2.41	2.89	2.24	.004	.026	[0.89, 4.62]	[0.30, 4.53]

Outcome	Marginal criteria (*SE*)			Contrast [95% CI]	
	LIT	DIT	CBT	DIT *v.* LIT	DIT *v.* CBT
End-point Means					
Observed	14.84 (1.39)	9.96 (.93)	13.22 (1.70)	**4.88 [1.59, 8.16]****	3.25 [-0.56, 7.01]
Imputed	14.43 (1.41)	10.70 (.90)	11.49 (1.59)	**3.71 [0.52, 6.91]**	0.78 [-2.82, 4.40]
CSC					
Observed	9% (.08)	51% (.07)	20% (.20)	**-.42 [-.64, -.21]*****	**-.31 [-.56, -.06]***
Imputed	13% (.07)	42% (.06)	32% (.11)	**-.29 [-.47, -.10]**	-.09 [-.15, .34]

Note. B, unstandardized regression coefficient; *CI*, confidence interval; CBT, Cognitive-behavioural therapy; *CSC*, clinically significant change; *DIT*, Dynamic Interpersonal Therapy; IIT, intention-to-treat analysis; *LIT*, Low-intensity Treatment; *MI*, multiple imputation analysis. Significant results are in bold: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$. Significance values were not available for imputed mixed-effect analyses.

Supplementary Table 6. *Comparison of the main regression coefficients and marginal estimates of the HRSD-17 between the full sample estimates (full, N = 147) and moderate-to-severe (MTS, N = 95) sample excluding mild cases.*

	B		*z*		*p*		95% CI	
Predictor	Full	MTS	Full	MTS	Full	MTS	Full	MTS
Time (linear)	-3.81	-3.81	-3.11	-2.50	.002	.012	[-6.21, -1.41]	[-6.79, -0.83]
Time (Quadratic) x Group (DIT *v.* LIT)	2. 57	2.48	2.89	2.26	.004	.024	[0.89, 4.62]	[0.33, 4.64]

Outcome	Marginal criteria (*SE*)			Contrast [95% CI]	
	LIT	DIT	CBT	DIT *v.* LIT	DIT *v.* CBT
End-point Means					
Full	14.84 (1.39)	9.96 (.93)	13.22 (1.70)	**4.88 [1.59, 8.16]****	3.25 [-0.56, 7.01]
MTS	16.12 (1.53)	12.01 (1.23)	15.50 (2.01)	**4.11 [0.24, 7.98]***	3.49 [-1.12, 8.11]
CSC					
Full	9% (.08)	51% (.07)	20% (.20)	**-.42 [-.64, -.21]*****	**-.31 [-.56, -.06]***
MTS	11% (.05)	47% (.05)	3% (.03)	**-.37 [-.50, -.23]*****	**-.45 [-.56, -.34]*****

Note. B, unstandardized regression coefficient; *CI*, confidence interval; CBT, Cognitive-behavioural therapy; *CSC*, clinically significant change; *DIT*, Dynamic Interpersonal Therapy; IIT, intention-to-treat analysis; *MTS*, moderate-to-severe, *SE*, standard error.
Significant results are in bold: $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Supplementary Table 7. Treatment differences between prorated and imputed means on the secondary outcome measures at baseline.

	Marginal Mean (*SE*)			Contrast [95% CI]	
Outcome measure	LIT	DIT	CBT	DIT *v.* LIT	DIT *v.* CBT
ECR					
Avoidance	4.31 (0.19)	4.24 (0.16)	4.66 (0.31)	0.07 [–0.42, 0.56]	0.42 [–0.26, 1.11]
Anxiety	4.52 (0.20)	4.54 (0.16)	4.76 (0.34)	-0.02 [–0.52, 0.49]	0.22 [–0.53, 0.96]
EQ-5D					
Index	0.77 (0.01)	0.78 (0.01)	0.78 (0.02)	-0.01 [–0.04, 0.03]	0 [–0.05, 0.05]
Continuous	54.59 (2.68)	60.14 (2.41)	67.24 (4.39)	-5.55 [-12.67, -1.58]	7.01 [–3.08, 17.28]
BSI					
Somatic	1.18 (0.15)	1.12 (0.13)	1.07 (0.24)	0.06 [–0.32, 0.44]	-0.05 [–0.61, 0.52]
OC	2.32 (0.14)	2.24 (0.12)	2.26 (0.23)	0.07 [-0.28, 0.43]	0.02 [–0.50, 0.54]
IS	2.09 (0.14)	2.08 (0.12)	2.06 (0.23)	0.01 [–0.36, 0.38]	-0.02 [–0.54, 0.50]
Depression	2.37 (0.13)	2.18 (0.11)	2.38 (0.21)	0.18 [-0.14, 0.51]	0.20 [–0.27, 0.66]
Anxiety	1.75 (0.16)	1.81 (0.14)	1.51 (0.26)	-0.06 [-0.47, 0.35]	-0.30 [–0.90, 0.30]
Hostility	0.95 (0.10)	1.06 (0.09)	0.60 (0.18)	-0.11 [–0.38, 0.16]	**-0.46 [–0.85, -0.07]***
Phobic	1.20 (0.17)	1.14 (0.16)	1.43 (0.29)	0.06 [–0.40, 0.52]	0.29 [–0.37, 0.95]
PI	1.39 (0.14)	1.41 (0.13)	1.17 (0.24)	-0.02 [–0.40, 0.36]	-0.24 [–0.78, 0.31]
Psychoticism	1.51 (0.11)	1.49 (0.09)	1.44 (0.17)	0.03 [-0.25, 0.31]	-0.04 [–0.43, 0.34]
GSI	1.65 (0.10)	1.61 (0.09)	1.58 (0.16)	0.05 [-0.21, 0.30]	-0.02 [–0.40, 0.33]
PSDI	2.41 (0.09)	2.25 (0.07)	2.31 (0.15)	0.16 [-0.08, 0.40]	0.06 [–0.28, 0.51]
IIP					
Total distress	1.75 (0.07)	1.75 (0.06)	1.77 (0.11)	0.00 [–0.18, 0.18]	0.02 [–0.22, 0.27]

Domineering	–0.69 (0.09)	–0.71 (0.08)	–0.72 (0.14)	0.02 [–0.21, 0.26]	-0.01 [–0.32, 0.31]
Vindictive	–0.42 (0.08)	–0.46 (0.07)	–0.53 (0.13)	0.04 [–0.17, 0.25]	0.06 [–0.34, 0.22]
Cold	–0.08 (0.09)	–0.09 (0.08)	–0.02 (0.14)	0.01 [–0.22, 0.24]	0.06 [–0.25, 0.38]
Socially Inhibited	0.34 (0.10)	0.46 (0.08)	0.50 (0.16)	-0.12 [–0.39, 0.14]	0.04 [-0.32, 0.40]
Non-assertive	0.61 (0.09)	0.66 (0.08)	0.80 (0.14)	-0.05 [–0.28, 0.18]	0.13 [–0.18, 0.45]
OA	0.29 (0.08)	0.26 (0.08)	0.29 (0.12)	0.03 [–0.17, 0.24]	0.03 [–0.25, 0.31]
Self-sacrificing	0.32 (0.07)	0.38 (0.06)	0.24 (0.12)	-0.06 [–0.26, 0.13]	-0.14 [–0.40, 0.12]
Intrusive	–0.38 (0.10)	-0.50 (0.09)	–0.56 (0.16)	0.12 [–0.14, 0.39]	-0.06 [–0.42, 0.30]
SAS					
Total	2.74 (0.06)	2.71 (0.05)	2.63 (0.11)	0.03 [-0.13, 0.20]	-0.07 [–0.32, 0.17]
Work	2.64 (0.11)	2.63 (0.10)	2.37 (0.19)	0 [-0.30, 0.31]	0.26 [-0.70, 0.18]
Social (friends)	3.02 (0.08)	3.02 (0.07)	2.82 (0.14)	-0.01 [0.23, 0.22]	-0.20 [–0.52, 0.13]
Extended family	2.63 (0.09)	2.55 (0.08)	2.72 (0.15)	0.07 [–0.16, 0.31]	0.17 [–0.17, 0.51]
Household	2.48 (0.12)	2.47 (0.10)	2.33 (0.20)	0.02 [–0.30, 0.33]	-0.14 [–0.58, 0.30]

Note. SE, Standard error; CI, confidence interval; LIT, low-intensity therapy; DIT, Dynamic Interpersonal Therapy; CBT, cognitive-behavioural therapy; BSI, Brief Symptom Inventory; OC, Obsessive-compulsive; IS, Interpersonal sensitivity; PI, Paranoid ideation; GSI, Global Severity Index; PSDI, Positive Symptom Distress Index; IIP, Inventory of Interpersonal Problems; OA; Overly Accommodating; SAS, Social Adjustment Scale.
Significant results are in bold: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.