**APPENDIX**

**Data supplement 1. Additional information about treatment integrity.**

132 patients gave consent for video-taping of the sessions, videos were available for 129 patients. Treatment adherence (i.e. whether the therapy protocol was followed, measured by the CSPRS-6) was measured by independent expert raters (three psychotherapists and one psychologist) that assessed 51 randomly selected videos of different patients drawn between session 6 and 12. Drawn videos were stratified with respect to treatment completers (n=37 videos) versus treatment non-completers (n=14 videos) and depression severity among the completers (BDI-II score < 29 (n=14 videos) versus ≥ than 29 (n=23 videos). Stratification was done to the ratio of completers/non-completers and low versus high depression severity in the study at the moment of stratification (February 2018).

Videos for 126 patients were rated on treatment competence. Therapy competence(i.e. how well the therapy was performed) was rated by both experts and master’s level clinical graduate students specifically trained in each of the modalities (three students per modality). Clinical graduate students received a 4-hour training that focused on the rating scale (CTS or IPT-AQS) and included one practice rating after which independent scores were compared and discussed. After the training, students received weekly to biweekly booster sessions to discuss common video ratings. For the trained students, 2 sessions[[1]](#footnote-1) between sessions 6 and 12 were randomly drawn for each patient that was videotaped. Master level clinical graduate students rated a total of 116 unique CBT videos of 60 different patients and 131 unique IPT videos of 66 different patients. For each session that was rated by multiple raters mean scores for that specific session (for CBT, 28 sessions were double rated by students, 11 sessions by experts and 3 sessions by both students and experts; for IPT, 47 sessions were double rated by students, 4 sessions by experts and 9 sessions by both students and experts) were computed. Mean competence scores for each patient were computed by calculating the mean of the competence scores for the two rated sessions. Note that the ‘all sessions form’ of the IPT Adherence and Quality Scale was used (<https://iptinstitute.com/ipt-training-materials/ipt-quality-adherence-scale/>) and is referred to as the short version.

For the CSPRS scores (*adherence*), ICCs (two-way random, absolute agreement, single measures) between experts were .56 (*n* = 3 double ratings) and .43 (*n* = 2 double ratings) for CBT and IPT, respectively. For the CTS and IPT-AQS total scores (*competence),* ICCs (two-way random, absolute agreement, average measures) for the double student ratings were .69 (*n* = 30 double ratings) and .67 (*n* = 56 double ratings) for CBT and IPT, respectively. For the double student ratings on the CBT and IPT items (*competence*), ICCs were .51 (*n* = 30 double ratings) and .58 (*n* = 54 double ratings) for the double student ratings on the CBT and IPT items, respectively. Note that the ICCs on the adherence measure were computed using all item scores of the instrument, while the ICCs of the total scores on the competence measures were computed using the total scores on the CTS and IPT-AQS.

**Data supplement 2. Additional information about the data-analyses.**

The flowchart, pre-treatment variables, therapist characteristics, treatment- and study compliance, treatment integrity and descriptive statistics were described. Differences in days between sessions, treatment duration, number of sessions, days between the baseline assessment and start of treatment were tested between frequency conditions using independent t-tests, differences in attrition (treatment compliance) and study compliance between session frequencies were tested using a chi-squared test.

Multilevel regression analyses using maximum likelihood estimation were conducted to investigate the effect of condition on depression (BDI-II scores). Intervention was represented by two dichotomous variables: CBT (1) versus IPT (1) and once weekly (-1) versus twice weekly (1) sessions. Note that while in our original protocol [1] we planned to use the online assessments that were conducted on equal intervals across conditions as the outcome only (i.e., two weeks after the start of treatment and every month until month 6), because of the long waiting time (*M* = 35.20 days, *SD* = 27.05) between baseline and session 1 (see also Table 2), we decided to add the BDI-II score measured before the start of session 1 to the outcome (i.e. dependent variable = BDI-II score measured before session 1, two weeks and monthly after start of treatment up to month 6, leading to a total of eight time points). Because this method might be unusual compared to other trials that analysed the data from baseline to post-treatment and not included BDI-II session 1 scores, we checked whether the same analyses without adding session 1 in the outcome would lead to different results. This was not the case and therefore we chose to keep up with the model that includes session 1 BDI-II scores in the outcome. The model corrected for baseline values by adding BDI-II baseline values as a covariate in the model [2].

The initial basic model was a two-level model with repeated measurements (level 1) nested within patients (level 2) with two two-way interactions testing the difference in change of BDI-II scores over time in days for the different session frequencies (time\*frequency) and treatments (time\*treatment). Time in days was centred 90 days after the start of treatment (i.e., to reduce collinearity between variables representing time) and it was tested whether a quadratic or cubic function of time led to better model fit. Subsequently, the addition of random levels on therapist and treatment centre level were tested, followed by a test of random slopes for time on all fitted random levels. Fit of different (co)variance structures for both the random effects and residuals errors within the lowest level (repeated measurements) were compared. To test whether the effect of session frequency was different for CBT versus IPT, a model with a three-way interaction (frequency\*treatment\*time) was fitted.

The same method was used to investigate effects on the secondary outcomes. Because we also had session ratings for current mood and both happiness items, in these models scores before session 1 were also added to the dependent variable. All models were corrected for baseline values of the outcome by adding their baseline scores as a covariate in the model [2].

All models were intention-to-treat, meaning that all 200 patients that were randomized into a condition were included in the analyses. Analyses were performed with SPSS [3] [3] and Stata[4] and significance levels were set to *p* < 0.05.

Effect sizes from session 1 to month 6 (Cohen’s *d*) [5] for the primary outcomes were computed (estimated mean at session 1– estimated mean at month 6) / mean estimated standard deviation at session 1). Note that we deliberately calculated effect sizes from session 1 to month 6, because the long waiting time between baseline and the start of therapy actually accounts for the observed difference between CBT and IPT on month 6 when the baseline measurement is used.

Three sensitivity analyses were conducted. First, it was investigated whether the total number of sessions, the presence of a comorbid anxiety disorder or the use of antidepressants were covariates and influenced the main results. Second, because a number of the time-points (week 2, month 1, month 2, month 3, month 4) are not comparing the same number of treatment sessions between frequency conditions (i.e., at month 2, weekly has had 8 sessions; twice weekly has had 16 sessions), and this could potentially skew the estimates in the analysis, we additionally tested the change in BDI-II scores from baseline to month 6, ignoring symptom levels across the other time points by using a multilevel that used time as dummy variables. The main effects of session frequency, treatment modality and time (using dummies for session 1, week 2, month 1-5 versus month 6 (reference category)), BDI baseline as a covariate and interactions between session frequency\*time and treatment modality\*time were included. By making time at month 6 the reference category the main effect of frequency would represent the change in BDI-II scores from baseline to month 6. Third, an completers-analyses was conducted by fitting the final model from the main analyses on completers-only (*n* = 152).

Based on the methodology of Jacobson & Truax [6] response (reliable change) was calculated as a decrease of at least 7 BDI-II points from session 1 to month 6 and remission (clinically significant change) as an absolute value of 17 or less on the BDI-II at month 6. However, because these criteria are not in line with previous definitions of response and remission in the field of depression and a BDI-II score of 17 still meets the criteria for mild depression, we decided to use more conservative cut-off scores (i.e. response = minimal decrease of 9 BDI-II points and remission an absolute value of 9 or less on the BDI-II at month 6) that have been used in previous randomized trials on the effects of psychotherapy for depression [7–9]. Differences in reliable change at month 6 and clinically significant change between session 1 and month 6 between session frequency conditions were tested using a chi-squared test. Cox regression was conducted to examine differences in the time in days to response and remission (censoring = 0, response/remission as = 1, response and remission were measured from session 1 to month 6) for weekly versus twice weekly sessions [[2]](#footnote-2) and including treatment modality and BDI-II scores at baseline as a covariate. The assumption of proportional hazards was tested by testing Schoenfeld residuals, testing time\*conditions interactions and inspecting the plot. Results showed that the proportional hazard assumption was supported.

Results are reported according to the CONSORT guidelines for reporting trials [10,11]. This study is registered: Netherlands Trial Register (registration number: NTR4856).

References data supplement 2.

1. Bruijniks SJE, Bosmans J, Peeters FPML, Hollon SD, van Oppen P, van den Boogaard M, et al. Frequency and change mechanisms of psychotherapy among depressed patients: Study protocol for a multicenter randomized trial comparing twice-weekly versus once-weekly sessions of CBT and IPT. BMC Psychiatry. 2015;15(1).

2. de Boer MR, Waterlander WE, Kuijper L, Steenhuis I, Twisk J. Testing for baseline differences in randomized controlled trials: an unhealthy research behavior that is hard to eradicate. Int J Behav Nutr Phys Act [Internet]. 2015;12(1):4. Available from: http://www.ijbnpa.org/content/12/1/4

3. IBM Corp. Released 2013. IBM SPSS Statistics for Windows. (Armonk, NY: IBM Corp.).

4. StataCorp, College Station TSL. StataCorp. Stata Statistical Software. 2015. p. Release 14.

5. Cohen J. Statistical power analysis for the behavioral sciences. 1988;

6. Jacobson NS, Truax P. Clinical Significance : A Statistical Approach to Denning Meaningful Change in Psychotherapy Research. 1991;59(1):12–9.

7. Lemmens LHJM, Arntz a, Peeters F, Hollon SD, Roefs a, Huibers MJH. Clinical effectiveness of cognitive therapy v. interpersonal psychotherapy for depression: results of a randomized controlled trial. Psychol Med [Internet]. 2015;(Ci):1–16. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25640151

8. Ekeblad A, Falkenström F, Andersson G, Vestberg R, Holmqvist R. Randomized Trial of Interpersonal Psychotherapy and Cognitive Behavioral Therapy for Major Depressive Disorder in a Community-Based Psychiatric Outpatient Clinic. Depress Anxiety. 2016;33(12):1090–8.

9. Elkin I, Shea MT, Watkins JT, Imber SD, Sotsky SM, Collins JF, et al. National Institute of Mental Health treatment of depression collaborative research program: General effectiveness of treatments. Arch Gen Psychiatry. 1989;46(11):971–82.

10. Boutron I, Altman DG, Moher D, Schulz KF, Ravaud P. CONSORT Statement for randomized trials of nonpharmacologic treatments: A 2017 update and a CONSORT extension for nonpharmacologic Trial Abstracts. Ann Intern Med. 2017;167(1):40–7.

11. Montgomery P, Grant S, Mayo-Wilson E, Macdonald G, Michie S, Hopewell S, et al. Reporting randomised trials of social and psychological interventions: The CONSORT-SPI 2018 Extension. Trials [Internet]. 2018;19:407. Available from: https://doi.org/10.1186/s13063-018-2733-1

**Data supplement 3. Treatment characteristics.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **CBT weekly (*n* = 49)** | **CBT twice-weekly (*n* = 49)** | **IPT weekly (*n* = 55)** | **IPT twice-weekly (*n* = 47)** | **Significant difference between session frequencies?** |
| ***Information about the waiting period*** |  |  |  |  |  |
| Days between baseline assessment and session 1, *M* (*SD*) | 34.98 (28.60) | 36.30 (25.16) | 31.78 (28.28) | 38.07 (26.40) | ns |
| Observed BDI-II score at first session | 30.64 (10.36) | 33.22 (9.44) | 31.85 (10.95) | 31.26 (11.47) | ns |
| ***Session frequency*** |  |  |  |  |  |
| Days between sessions, *M* (*SD*) | 9.03 (3.12) | 4.89 (2.89) | 8.94 (3.04) | 4.45 (1.35) | *p* <.001 |
| ***Treatment duration*** |  |  |  |  |  |
| Total treatment length, in days, *M* (*SD*) | 169.17 (63.88) | 111.38 (41.55) | 171.62 (70.87) | 115.09 (46.07) | *p* < .001 |
| Number of sessions, *M* (*SD*) | 15.89 (5.32) | 16.97 (4.38) | 15.74 (5.17) | 17.31 (4.38) | ns |
| Non-completer, *n* (%) | 13 (26.5) | 7 (14.3) | 19 (34.5) | 9 (19.1) | *p* < .05 |

*Note*. BDI-II, Beck Depression Inventory, second edition; *M* = mean; *SD* = standard deviation. Note that days between sessions include weekend days. Completers are defined as patients that received a minimum of 12 sessions and no deviations-other protocols or use of antidepressants before session 12. Note that, except for the data on the (non) completers, data in this Table does not include patients that did not start treatment (n=9). Additional data was missing for number of sessions: IPT twice weekly (n=1); session frequency: CBT twice-weekly (missing *n* = 1), IPT weekly (missing *n* = 1), IPT twice weekly (missing n=2); total treatment length: CBT twice-weekly (missing *n* = 1), IPT weekly (missing *n* = 1), IPT twice weekly (missing n=1); days between baseline and session 1: CBT twice-weekly (missing *n* = 1), IPT weekly (missing *n* = 1) IPT twice weekly (missing n=1); BDI-II score at the first session: CBT weekly (missing *n* = 2), IPT weekly (missing *n* = 2) and IPT twice-weekly (missing *n* = 4); finished or received at least 16 sessions of treatment by month 6: IPT twice weekly (n=1). Note that the somewhat lower number of sessions in the weekly conditions may be due to the number of drop-out. Number of sessions in completers only was: 17.97 (*SD* = 3.02; *n* = 36), 17.59 (*SD* = 3.66 *n* = 42), 17.97 (*SD* = 2.74 *n* = 36) and 18.47 (*SD* = 1.95 *n* = 38) for CBT weekly, CBT twice weekly, IPT weekly, IPT twice weekly respectively.

**Data supplement 4. Mixed regression-based marginal means (95% CI) for all outcome measures in the intention-to-treat sample stratified to condition.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **CBT weekly (N=49)** | | **CBT twice weekly (N=49)** | | **IPT weekly (N=55)** | | **IPT twice weekly (N=47)** |
| *﻿****Primary outcome***  Beck Depression Inventory II | |  | |  | |  | |  |
| Session 1 | | 31.05 (29.29-32.81) | | 32.58 (30.84-34.32) | | 32.49 (30.77-34.21) | | 34.02 (32.25-35.79) |
| Week 2 | | 30.60 (29.02-32.17) | | 30.53 (28.98-32.08) | | 32.10 (30.58-33.63) | | 32.04 (30.46-33.62) |
| Month 1 | | 30.08 (28.40-31.77) | | 28.65 (26.98-30.31) | | 31.53 (29.88-33.17) | | 30.09 (28.44-31.75) |
| Month 2 | | 29.09 (27.07-31.11) | | 26.18 (24.20-28.17) | | 30.15 (28.17-32.14) | | 27.25 (25.26-29.24) |
| Month 3 | | 27.98 (25.67-30.29) | | 24.67 (22.40-26.94) | | 28.56 (26.29-30.83) | | 25.25 (22.95-27.54) |
| Month 4 | | 26.66 (23.99-29.33) | | 23.59 (20.94-26.25) | | 26.93 (24.28-29.57) | | 23.86 (21.23-26.49) |
| Month 5 | | 25.03 (21.93-28.14) | | 22.44 (19.36-25.52) | | 25.44 (22.36-28.53) | | 22.85 (19.80-25.91) |
| Month 6 | | 23.01 (19.45-26.57) | | 20.69 (17.18-24.21) | | 24.30 (20.77-27.83) | | 21.98 (18.44-25.52) |
| ***Secondary outcomes***  Mood | |  | |  | |  | |  |
| Session 1 | | 38.34 (34.51-42.18) | | 35.43 (31.64-39.22) | | 35.82 (32.09-39.54) | | 32.90 (29.06-36.75) |
| Week 2 | | 39.84 (36.38-43.30) | | 38.43 (35.03-41.84) | | 36.59 (33.23-39.94) | | 35.18 (31.70-38.66) |
| Month 1 | | 41.44 (38.10-44.78) | | 41.55 (38.27-44.84) | | 37.52 (34.27-40.76) | | 37.63 (34.29-40.97) |
| Month 2 | | 44.11 (40.46-47.76) | | 46.50 (42.89-50.12) | | 39.41 (35.84-42.98) | | 41.80 (38.22-45.38) |
| Month 3 | | 46.36 (42.26-50.47) | | 50.28 (46.21-54.35) | | 41.49 (37.45-45.52) | | 45.40 (41.39-49.42) |
| Month 4 | | 48.19 (43.72-52.66) | | 52.88 (48.47-57.29) | | 43.76 (39.34-48.17) | | 48.45 (44.03-52.86) |
| Month 5 | | 49.59 (44.74-54.44) | | 54.30 (49.52-59.08) | | 46.21 (41.41-51.02) | | 50.93 (46.11-55.75) |
| Month 6 | | 50.57 (45.02-56.12) | | 54.55 (49.04-60.06) | | 48.86 (43.34-54.38) | | 52.85 (47.38-58.31) |
| General life happiness | |  | |  | |  | |  |
| Session 1 | | 2.91 (2.69 - 3.12) | | 2.76 (2.56 - 2.97) | | 3.01 (2.81 - 3.22) | | 2.87 (2.66 - 3.08) |
| Week 2 | | 3.05 (2.87 - 3.23) | | 3.08 (2.91 - 3.26) | | 3.09 (2.92 - 3.27) | | 3.13 (2.95 - 3.31) |
| Month 1 | | 3.18 (2.99 - 3.36) | | 3.36 (3.18 - 3.54) | | 3.18 (3.01 - 3.36) | | 3.37 (3.19 - 3.55) |
| Month 2 | | 3.34 (3.14 - 3.54) | | 3.68 (3.49 - 3.88) | | 3.34 (3.14 - 3.53) | | 3.68 (3.49 - 3.88) |
| Month 3 | | 3.43 (3.23 - 3.64) | | 3.82 (3.61 - 4.02) | | 3.48 (3.28 - 3.68) | | 3.86 (3.66 - 4.07) |
| Month 4 | | 3.51 (3.28 - 3.73) | | 3.85 (3.62 - 4.08) | | 3.61 (3.39 - 3.84) | | 3.96 (3.74 - 4.18) |
| Month 5 | | 3.59 (3.34 - 3.85) | | 3.88 (3.62 - 4.13) | | 3.74 (3.48 - 3.99) | | 4.02 (3.77 - 4.27) |
| Month 6 | | 3.74 (3.45 - 4.03) | | 3.99 (3.70 - 4.27) | | 3.86 (3.57 - 4.15) | | 4.11 (3.82 - 4.40) |
| Happiness today | |  | |  | |  | |  |
| Session 1 | | 3.35 (3.17 – 3.53) | | 3.29 (3.11 – 3.47) | | 3.26 (3.08 – 3.44) | | 3.20 (3.01 – 3.38) |
| Week 2 | | 3.39 (3.21 - 3.56) | | 3.36 (3.19 - 3.53) | | 3.30 (3.13 - 3.47) | | 3.28 (3.10 - 3.45) |
| Month 1 | | 3.43 (3.26 - 3.60) | | 3.45 (3.28 - 3.61) | | 3.35 (3.19 - 3.52) | | 3.37 (3.20 - 3.54) |
| Month 2 | | 3.51 (3.34 - 3.68) | | 3.61 (3.44 - 3.77) | | 3.44 (3.28 - 3.61) | | 3.54 (3.37 - 3.71) |
| Month 3 | | 3.59 (3.41 - 3.77) | | 3.77 (3.59 - 3.95) | | 3.54 (3.36 - 3.72) | | 3.72 (3.53 - 3.90) |
| Month 4 | | 3.67 (3.46 - 3.88) | | 3.93 (3.73 - 4.14) | | 3.63 (3.42 - 3.84) | | 3.89 (3.68 - 4.10) |
| Month 5 | | 3.75 (3.51 - 3.99) | | 4.09 (3.85 - 4.33) | | 3.72 (3.48 - 3.96) | | 4.06 (3.82 - 4.30) |
| Month 6 | | 3.83 (3.55 - 4.11) | | 4.25 (3.98 - 4.53) | | 3.81 (3.53 - 4.09) | | 4.24 (3.96 - 4.52) |
| Remission of Depression Questionnaire | | |  | |  | |  | |
| *RDQ - Depressive symptoms* | | |  | |  | |  | |
| Month 3 | 13.00 (11.53-14.47) | | 12.61 (11.17-14.05) | | 13.46 (11.99-14.93) | | 13.06 (11.61-14.51) | |
| Month 6 | 11.64 (10.09-13.18) | | 10.47 (8.91-12.02) | | 11.97 (10.43-13.53) | | 10.80 (9.27-12.34) | |
| *RDQ - Other symptoms often present in depressed patients such as anxiety and irritability* | | | | | | |  | |
| Month 3 | 5.38 (4.59-6.16) | | 5.05 (4.25-5.85) | | 5.78 (4.96-6.59) | | 5.44 (4.62-6.27) | |
| Month 6 | 4.75 (3.93-5.58) | | 4.39 (3.55-5.24) | | 5.18 (4.32-6.04) | | 4.82 (3.96-5.68) | |
| *RDQ - Coping ability* |  | |  | |  | |  | |
| Month 3 | 3.45 (2.90-3.99) | | 3.26 (2.72-3.79) | | 3.39 (2.84-3.94) | | 3.20 (2.67-3.74) | |
| Month 6 | 3.58 (3.09-4.06) | | 2.93 (2.44-3.43) | | 3.20 (2.71-3.68) | | 2.55 (2.08-3.03) | |
| *RDQ - Positive health* |  | |  | |  | |  | |
| Month 3 | 14.74 (13.07-16.41) | | 13.95 (12.24-15.65) | | 13.24 (11.51-14.97) | | 12.45 (10.71-14.18) | |
| Month 6 | 13.47 (11.84-15.09) | | 11.97 (10.31-13.62) | | 12.22 (10.53-13.91) | | 10.72 (9.03-12.41) | |
| *RDQ -Functioning* |  | |  | |  | |  | |
| Month 3 | 3.07 (2.56-3.59) | | 2.85 (2.35-3.36) | | 2.95 (2.44-3.47) | | 2.73 (2.22-3.24) | |
| Month 6 | 3.00 (2.53-3.47) | | 2.54 (2.07-3.01) | | 2.93 (2.46-3.40) | | 2.47 (2.01-2.93) | |
| *RDQ - Life satisfaction* |  | |  | |  | |  | |
| Month 3 | 3.82 (3.29-4.35) | | 3.71 (3.19-4.23) | | 3.54 (3.00-4.08) | | 3.43 (2.91-3.96) | |
| Month 6 | 3.37 (2.89-3.85) | | 2.91 (2.42-3.39) | | 3.32 (2.85-3.80) | | 2.86 (2.40-3.32) | |
| *RDQ - General sense of wellbeing* | | |  | |  | |  | |
| Month 3 | 4.08 (3.50-4.65) | | 3.75 (3.19-4.31) | | 3.72 (3.15-4.30) | | 3.40 (2.83-3.96) | |
| Month 6 | 3.71 (3.20-4.23) | | 3.07 (2.55-3.59) | | 3.55 (3.04-4.06) | | 2.90 (2.40-3.40) | |
| RAND-36 – General health perception | | |  | |  | |  | |
| Week 2 | 46.88 (43.74-50.29) | | 43.95 (40.72-47.18) | | 44.51 (41.19-47.84) | | 41.58 (38.23-44.93) | |
| Month 3 | 48.45 (45.58-51.32) | | 47.98 (45.20-50.76) | | 46.09 (43.27-48.91) | | 45.62 (42.80-48.44) | |
| Month 6 | 50.31 (46.23-54.39) | | 52.76 (48.75-56.77) | | 47.95 (43.89-52.02) | | 50.40 (46.40-54.41) | |

*Note*. BDI-II, Beck Depression Inventory, second edition; RAND36= RAND-36 item Health Survey; RDQ = Remission of Depression Questionnaire. Note that baseline scores are not corrected because baseline scores cannot correct for themselves. Note that lower scores on the BDI-II and RDQ, and higher scores on current mood, the happiness measures and the RAND-36 indicate better functioning.

**Data supplement 5. ﻿ ﻿Results of multilevel analyses on the BDI-II scores.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **B** | **Std. Err.** | **95% CI** | **z** | **p** |
|  |  |  |  |  |  |
| Model 1 Effect of session frequency over time | | | |  |  |
| Intercept | -3.89 | 1.74 | -7.30/-.48 | -2.24 | .02 |
| **BDI baseline** | **.88** | **.04** | **.79/.97** | **19.11** | **<.001** |
| **Time in days** | **-.04** | **.07e-1** | **-.06/-.03** | **-6.42** | **<.001** |
| Time in days ^2 | -.11e-3 | .56e-4 | 1.60e-6/.22e-3 | 1.99 | .04 |
| Time in days ^3 | -1.01e-6 | 6.35e--7 | -2.26e-6/2.31e-7 | -1.60 | .11 |
| **Session frequency** | **-1.65** | **.67** | **-2.97/-.33** | **-2.46** | **.01** |
| Treatment | .28 | .67 | -1.03/1.60 | .43 | .67 |
| Session frequency \* time in days | -.13e-3 | .07e-1 | -.01/.01 | -.02 | .98 |
| **Session frequency \* time in days^2** | **.18e-3** | **.56e-4** | **.70e-4/.28e-3** | **3.21** | **.001** |
| **Session frequency \* time in days^3** | **-1.30e-6** | **6.48e-7** | **-2.57e-6/-3.39e-8** | **-2.01** | **.04** |
| Treatment \* time in days | -.74e-2 | .76e-2 | -.02/.75e-2 | -.97 | .33 |
| Treatment \* time in days^2 | .48e-4 | .56e-4 | -.61e-4/.15e-3 | .87 | .38 |
| **Treatment \* time in days^3** | **8.64e-7** | **6.83e-7** | **-4.74e-07/2.20e-6** | **1.27** | **.20** |

*Note.* BDI-II, Beck Depression Inventory, second edition. Note that best fit on the primary outcomes was reached with a model that included a random intercept and slope on patient level, time modelled as a cubic function, using an unstructured covariance structure (i.e. meaning that each (co)variance for the random effects are estimated separately) and including a (co)variance matrix for the residuals errors that followed an MA (1) structure (i.e. meaning that the residual (co)variance is estimated depending on the residual of the preceding error). For the three-way interaction the same model structure showed best fit. Note that condition was coded in the following way: weekly (-1) versus twice weekly (1) and CBT (-1) versus IPT (1).

**Data supplement 6. Results of multilevel models on the secondary outcomes.**   
Current mood

Best fit on mood was reached with a model that included a random intercept and slope on patient level, time modelled as a quadratic function, using an unstructured covariance structure and including a (co)variance matrix for the residuals errors that followed an AR (1) structure (i.e. meaning that the residual (co)variance is estimated depending on the preceding error). There was a significant difference between session frequencies over time in favour of the twice-weekly sessions (time^2\*frequency condition: B = -.20e-3, S.E = .10e-3, z = -1.99, 95% *CI* -.41e-3/-3.42e-6, *p* = .046), but there were no significant differences between treatments over time (time^2\*treatment condition: B = .17e-3, S.E = .10e-3, z = 1.62, 95% *CI* -.36e-4/.37e-3, *p* = .10).

Happiness in general/Happiness today

Best fit on happiness in general was reached with a model that included a random intercept and slope on patient level, time modelled as a cubic function, using an unstructured covariance structure. There was a significant difference between session frequencies over time in favour of the twice-weekly sessions (time^3\*frequency condition: B = 1.52e-07, S.E = 7.53e-08, z = 2.02, 95% *CI* 4.27e-09/3e-07, *p* = .044), but there were no significant differences between treatments over time (time^3\*treatment condition: B = -1.19e-07, S.E = 7.68e-08, z = -1.56, 95% *CI* -2.7e-07/3.1e-08, *p* = .12). Best fit on happiness today was reached with a model that included a random intercept and slope on patient level, time modelled as a linear function, using an unstructured covariance structure and including a (co)variance matrix for the residuals errors that followed an AR (1) structure (i.e. meaning that the residual (co)variance is estimated depending on the preceding error). There was a significant difference between session frequencies over time in favour of the twice-weekly sessions (time\*frequency condition: B =.13e-2, S.E = .49e-3, z = 2.79, 95% *CI* .40e-3/.23e-2, *p* = .005), but there were no significant differences between treatments over time (time\*treatment condition: B = .19e-3, S.E = .49e-3, z = .41, 95% *CI* -.76e-3/.11e-2, *p* = .68).

RAND-36

Best fit on the RAND-36 was reached with a model that included a random intercept and slope on patient level, time modelled as a linear function, using an unstructured covariance structure. There were no significant differences between session frequencies over time (time\*frequency condition: B = .01, S.E = .83e-2, z = 1.94, 95% *CI* -.19e-3/.03, *p* = .05), or between treatments over time (time\*treatment condition: B = .39e-4, S.E =.83e-2, z = .00, 95% *CI* -.01/.01, *p* = .99).

RDQ-depressive symptoms

Best fit on the RDQ-depressive symptoms was reached with a model that included a random intercept and slope on patient level, time modelled as a linear function, using the default independent structure for both the covariance and residual covariance. There were no significant differences between session frequencies over time (time\*frequency condition: B = -.43e-2, S.E = .42e-2, z = -1.01, 95% *CI* -.01/.40e-2, *p* = .31), or between treatments over time (time\*treatment condition: B = -.06e-2, S.E =.42e-2, z = -.15, 95% *CI* -.90e-2/.77e-2, *p* = .88).

RDQ-other symptoms

Best fit on the RDQ-other symptoms was reached with a model that included a random intercept and slope on patient level, time modelled as a cubic function, using the default independent structure for both the covariance and residual covariance. There were no significant differences between session frequencies over time (time^3\*frequency condition: B=-1.63e-6, S.E = 1.71e-6, z=-.95, 95% CI -4.98e-6/1.73e-6, p=.34), or between treatments over time (time^3\*treatment condition: B=3.64e-7, S.E =1.83e-6, z=.20, 95% CI -3.22e-6/3.95e-6, p=.84).

RDQ-coping ability

Best fit on the RDQ-coping ability was reached with a model that included a random intercept on patient level, time modelled as a linear function, using the default independent structure for both the covariance and residual covariance. There were no significant differences between session frequencies over time (time\*frequency condition: B=-.25e-2, S.E = .15e-2, z=-1.58, 95% CI -.56e-2/.60e-3, p=.11), or between treatments over time (time\*treatment condition: B=-.18e-2, S.E =.15e-2, z=-1.13, 95% CI -.49e-2/.13e-2, p=.26).

RDQ-positive health

Best fit on the RDQ-positive health was reached with a model that included a random intercept on patient level, time modelled as a cubic function, using the default independent structure for both the covariance and residual covariance. There were no significant differences between session frequencies over time (time^3\*frequency condition: B = -.5.95e-6, S.E = 3.32e-6 z = -1.79, 95% *CI* -.12e-4/5.6e-7, *p* = .07), or between treatments over time (time^3\*treatment condition: B = 6.34e-6, S.E =3.55e-6, z = 1.79, 95% *CI* -6.1e-7/.13e-4, *p* = .07).

RDQ-functioning

Best fit on the RDQ-functioning was reached with a model that included a random intercept on patient level, time modelled as a linear function, using the default independent structure for both the covariance and residual covariance. There were no significant differences between session frequencies over time (time\*frequency condition: B = -.13e-2, S.E = .13e-2 z = -.95, 95% *CI* -.40e-2/.14e-2, *p* = .34), or between treatments over time (time\*treatment condition: B = .28e-3, S.E =.13e-2, z = .20, 95% *CI* -.24e-2/.30e-2, *p* = .84).

RDQ-life satisfaction

Best fit on the RDQ-life satisfaction was reached with a model that included a random intercept on patient level, time modelled as a linear function, using the default independent structure for both the covariance and residual covariance. There were no significant differences between session frequencies over time (time\*frequency condition: B = -.19e-2, S.E = .15e-2 z = -1.28, 95% *CI* -.50e-2/.10e-2, *p* = .20), or between treatments over time (time\*treatment condition: B = .12e-2, S.E =.15e-2, z = .83, 95% *CI* -.17e-2/.43e-2, *p* = .41).

RDQ-general wellbeing

Best fit on the RDQ-wellbeing was reached with a model that included a random intercept on patient level, time modelled as a linear function, using the default independent structure for both the covariance and residual covariance. There were no significant differences between session frequencies over time (time\*frequency condition: B = -.17e-2, S.E = .16e-2 z = -1.05, 95% *CI* -.50e-2/.15e-2, *p* = .29), or between treatments over time (time\*treatment condition: B = .10e-2, S.E = .16e-2, z = .61, 95% *CI* -.22e-2/.43e-2, *p* = .54).

**Data supplement 7. Reliable change rates (%) on the Beck Depression Inventory-II (observed values).**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Reliable change rates (%)** | | |  |
|  | **CBT weekly** | **CBT twice weekly** | **IPT weekly** | **IPT twice weekly** |
| Session 1 - week 2 | 11.1 | 9.4 | 3.6 | 4.2 |
| Session 1 - month 1 | 17.5 | 28.6 | 7.7 | 21.6 |
| Session 1 - month 2 | 23.7 | 44.7 | 20.6 | 38.9 |
| Session 1 - month 3 | 37.1 | 47.2 | 25 | 42.9 |
| Session 1 - month 4 | 48.6 | 51.4 | 35.5 | 62.1 |
| Session 1 - month 5 | 43.2 | 56.7 | 31 | 59.4 |
| Session 1 - month 6 | 44.4 | 60 | 32.3 | 52.9 |

*Note.* Note that the n is slightly different per time point and group (session 1-week 2: 27, 32, 28, 24; session 1-month 1: 40, 35, 39, 37; session 1-month 2: 38, 38, 34, 36; session 1-month 3: 35, 36, 32, 35; session 1-month 4: 35, 37, 31, 29; session 1-month 5: 37, 30, 29, 32; session 1-month 6: 36, 35, 31, 34, for CBT weekly, CBT twice-weekly, IPT weekly, IPT twice-weekly, respectively).

**Data supplement 8. Clinically significant change rates (%) on the Beck Depression Inventory-II (observed values).**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Clinically significant change rates (%)** | | |  |
|  | **CBT weekly** | **CBT twice weekly** | **IPT weekly** | **IPT twice weekly** |
| Session 1 | 4.4 | 0 | 4.1 | 0 |
| Week 2 | 7.2 | 0 | 0 | 0 |
| Month 1 | 5 | 2.9 | 0 | 7.3 |
| Month 2 | 10.3 | 10.5 | 5.6 | 15 |
| Month 3 | 11.4 | 8.1 | 3 | 23.7 |
| Month 4 | 17.1 | 16.2 | 12.5 | 28.1 |
| Month 5 | 17.8 | 12.5 | 10 | 29.5 |
| Month 6 | 21.6 | 25.7 | 17.6 | 33.3 |

*Note.* Note that the n is slightly different per time point and group (session 1: 45, 48, 49, 42; week 2: 27, 33, 29, 27; month 1: 40, 35, 41, 41; month 2: 39, 38, 36, 40; month 3: 35, 37, 33, 38; month 4: 35, 37, 32, 32; month 5: 45, 48, 50, 44; month 6: 37, 35, 34, 39, for CBT weekly, CBT twice-weekly, IPT weekly, IPT twice-weekly, respectively).

**Data supplement 9. Within- and between-group effect sizes from before session 1 to post-treatment (month 6).**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Within group effect size** | | | | | | | **Between group effect size** | | | | | |  |
|  | **CBT weekly** | | **CBT twice weekly** | | **IPT weekly** | | **IPT twice weekly** | | **Weekly vs. twice weekly** |  | | **CBT vs. IPT** | |  |
|  | **Mean change** | ***d*** | **Mean change** | ***d*** | **Mean change** | ***d*** | **Mean change** | ***d*** | **Mean change difference** | | ***d*** | **Mean change difference** | ***d*** | |
| Session 1 – week 2 | -.45 | -.06 | -2.05 | -.29 | -.39 | -.05 | -1.98 | -.28 | -1.59 | | .22 | .06 | .00 | |
| Session 1 – month 1 | -.97 | -.13 | -3.93 | -.56 | -.96 | -.13 | -3.93 | -.56 | -2.96 | | .42 | .00 | .00 | |
| Session 1 – month 2 | -1.96 | -.28 | -6.4 | -.91 | -2.34 | -.33 | -6.77 | -.97 | -4.43 | | .63 | -.37 | -.05 | |
| Session 1 – month 3 | -3.07 | -.44 | -7.91 | -1.13 | -3.93 | -.56 | -8.77 | -1.25 | -4.84 | | .69 | -.86 | -.12 | |
| Session 1 – month 4 | -4.39 | -.63 | -8.99 | -1.29 | -5.56 | -.79 | -10.16 | -1.45 | -4.60 | | .66 | -1.17 | -.16 | |
| Session 1 – month 5 | -6.02 | -.86 | -10.14 | -1.45 | -7.05 | -1.01 | -11.17 | -1.60 | -4.12 | | .59 | -1.03 | -.14 | |
| Session 1 – month 6 | -8.04 | -1.15 | -11.89 | -1.70 | -8.19 | -1.17 | -12.04 | -1.72 | -3.85 | | .55 | -.15 | -.02 | |

*Note.* Effect size *d* = (estimated mean session 1 – estimated mean at time i) / mean estimated standard deviation at session 1. Between group effect size are the differences in effect size *d* between frequency and treatment conditions. The Pearson correlations between BDI-II scores before session 1 and subsequent measurements were: .81 (week 2), .70 (month 1), .69 (month 2), .58 (month 3), .51 (month 4), .55 (month 5), .48 (month 6).

1. For three patients receiving CBT and four patients receiving IPT only one session was rated. [↑](#footnote-ref-1)
2. In these analyses, time points > 200 days were considered as outliers and days and status at previous time points were used for these cases. [↑](#footnote-ref-2)