**Supplementary Appendix**

Table S1. Mixed Model Repeated Measures predicting loneliness (intention-to-treat analysis)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  β | SE | *df* | *t*-statistic | *p* value |
| Condition |  |  |  |  |  |
| (G4H vs. CBT) | .158 | .148 | 65.76 | 1.06 | .291 |
| Time |  |  |  |  |  |
| (1 v. 0)  | -.302 | .105 | 1078.98 | -2.87 | .004 |
| (2 v. 0) | -.439 | .109 | 1084.86 | -4.01 | <.0001 |
| (3 v. 0) | -.539 | .112 | 1087.40 | -4.82 | <.0001 |
| (4 v. 0) | -.680 | .113 | 1088.03 | -6.02 | <.0001 |
| (5 v. 0) | -.873 | .110 | 1086.05 | -7.96 | <.0001 |
| (6 v. 0) | -.854 | .109 | 1084.54 | -7.85 | <.0001 |
| (7 v. 0) | -.994 | .109 | 1083.59 | -9.11 | <.0001 |
| Condition X Time |  |  |  |  |  |
| (1 v. 0)  | .078 | .146 | 1077.46 | .54 | .592 |
| (2 v. 0) | -.237 | .150 | 1081.25 | -1.58 | .114 |
| (3 v. 0) | -.097 | .152 | 1082.86 | -.64 | .524 |
| (4 v. 0) | -.184 | .153 | 1083.35 | -1.20 | .229 |
| (5 v. 0) | -.010 | .150 | 1081.52 | -.07 | .946 |
| (6 v. 0) | -.68 | .148 | 1080.38 | -.46 | .646 |
| (7 v. 0) | .119 | .149 | 1079.51 | .80 | .421 |

*Note.*

This analysis is based on 1262 observations across 8 timepoints, 174 participants, and 26 therapy groups. Level 1 (time) was modelled as a categorical fixed effect, Level 2 (participant) was modelled as a random intercept, and Level 3 (therapy group) was a random intercept. Hypothesis testing and Figure 2 in the manuscript utilise estimated marginal means and planned comparisons based on this model.

Table S2. Mixed Model Repeated Measures predicting depression (intention-to-treat analysis).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  β | SE | *df* | *t*-statistic | *p* value |
| Condition |  |  |  |  |  |
| (G4H vs. CBT) | 0.100 | .162 | 60.84 | 0.621 | .537 |
| Time |  |  |  |  |  |
| (1 v. 0)  | -.236 | .109 | 1077.63 | -2.16 | .003 |
| (2 v. 0) | -.501 | .113 | 1084.62 | -4.43 | <.0001 |
| (3 v. 0) | -.626 | .116 | 1087.63 | -5.40 | <.0001 |
| (4 v. 0) | -.942 | .117 | 1088.37 | -8.05 | <.0001 |
| (5 v. 0) | -1.033 | .114 | 1085.82 | -9.08 | <.0001 |
| (6 v. 0) | -1.073 | .113 | 1084.12 | -9.52 | <.0001 |
| (7 v. 0) | -.741 | .113 | 1083.01 | -6.55 | <.0001 |
| Condition X Time |  |  |  |  |  |
| (1 v. 0)  | -.184 | .152 | 1075.97 | -1.22 | .224 |
| (2 v. 0) | -.202 | .155 | 1080.58 | -1.30 | .193 |
| (3 v. 0) | -.207 | .158 | 1082.53 | -1.31 | .190 |
| (4 v. 0) | .075 | .158 | 1083.04 | .474 | .636 |
| (5 v. 0) | -.017 | .155 | 1080.81 | -.108 | .914 |
| (6 v. 0) | .098 | .154 | 1079.47 | .638 | .523 |
| (7 v. 0) | -.226 | .154 | 1078.48 | -1.466 | .143 |

*Note.*

This analysis is based on 1262 observations across 8 timepoints, 174 participants, and 26 therapy groups. Level 1 (time) was modelled as a categorical fixed effect, Level 2 (participant) was modelled as a random intercept, and Level 3 (therapy group) was a random intercept. Hypothesis testing utilises estimated marginal means and planned comparisons based on this model.

Per-protocol analyses

Table S3. Mixed Model Repeated Measures predicting loneliness (per-protocol analysis).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  β | SE | *df* | *t*-statistic | *p* value |
| Condition |  |  |  |  |  |
| (G4H vs. CBT) | .117 | .165 | 59.62 | .71 | .480 |
| Time |  |  |  |  |  |
| (1 v. 0)  | -.328 | .115 | 998.42 | -2.85 | .004 |
| (2 v. 0) | -.464 | .116 | 998.95 | -4.00 | <.0001 |
| (3 v. 0) | -.567 | .117 | 1000.06 | -4.84 | <.0001 |
| (4 v. 0) | -.720 | .118 | 1000.30 | -6.12 | <.0001 |
| (5 v. 0) | -.916 | .115 | 998.48 | -7.96 | <.0001 |
| (6 v. 0) | -.906 | .117 | 999.62 | -7.77 | <.0001 |
| (7 v. 0) | -1.11 | .118 | 1000.38 | -9.35 | <.0001 |
| Condition X Time |  |  |  |  |  |
| (1 v. 0)  | .068 | .157 | 998.53 | 0.44 | .663 |
| (2 v. 0) | -.283 | .157 | 998.69 | -1.80 | .072 |
| (3 v. 0) | -.131 | .158 | 999.32 | -.83 | .409 |
| (4 v. 0) | -.198 | .159 | 999.56 | -1.25 | .212 |
| (5 v. 0) | -.63 | .157 | 998.45 | -.40 | .688 |
| (6 v. 0) | -.090 | .158 | 999.01 | -.57 | .568 |
| (7 v. 0) | .165 | .159 | 999.44 | 1.04 | .298 |

*Note.*

This analysis is based on 1164 observations across 8 timepoints, 152 participants, and 26 therapy groups. Level 1 (time) was modelled as a categorical fixed effect, Level 2 (participant) was modelled as a random intercept, and Level 3 (therapy group) was a random intercept. Hypothesis testing utilises estimated marginal means and planned comparisons based on this model.

Table S4. Mixed Model Repeated Measures predicting depression (per-protocol analysis).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  β | SE | *df* | *t*-statistic | *p* value |
| Condition |  |  |  |  |  |
| (G4H vs. CBT) | .042 | .127 | 60.14 | 0.24 | .238 |
| Time |  |  |  |  |  |
| (1 v. 0)  | -.262 | .119 | 998.39 | -2.21 | .028 |
| (2 v. 0) | -.569 | .119 | 999.03 | -4.76 | <.0001 |
| (3 v. 0) | -.679 | .121 | 1000.36 | -5.63 | <.0001 |
| (4 v. 0) | -1.02 | .121 | 1000.70 | -8.40 | <.0001 |
| (5 v. 0) | -1.10 | .119 | 998.48 | -9.25 | <.0001 |
| (6 v. 0) | -1.15 | .120 | 999.84 | -9.59 | <.0001 |
| (7 v. 0) | -0.832 | .122 | 1000.61 | -6.83 | <.0001 |
| Condition X Time |  |  |  |  |  |
| (1 v. 0)  | -.156 | .162 | 998.53 | -.97 | 334 |
| (2 v. 0) | -.180 | .162 | 998.72 | -1.11 | .266 |
| (3 v. 0) | -.182 | .163 | 999.48 | -1.11 | .265 |
| (4 v. 0) | .121 | .164 | 999.79 | .74 | .461 |
| (5 v. 0) | -.004 | .162 | 998.43 | -.02 | .982 |
| (6 v. 0) | .125 | .162 | 999.12 | .77 | .441 |
| (7 v. 0) | -.130 | .164 | 999.56 | -.79 | .429 |

*Note.*

This analysis is based on 1164 observations across 8 timepoints, 152 participants, and 26 therapy groups. Level 1 (time) was modelled as a categorical fixed effect, Level 2 (participant) was modelled as a random intercept, and Level 3 (therapy group) was a random intercept. Hypothesis testing and Figure 2 in the manuscript utilise estimated marginal means and planned comparisons based on this model.

Per-protocol test of H1

H1a predicted significant positive change in loneliness in the G4H condition. Contrasts were specified to evaluate this within the G4H condition from baseline to program completion, *t*(999)=-7.96, *p*<.0001, to six month follow-up, *t*(1000)=-7.77, *p*<.0001, and to 12-month follow up, *t*(1001)=-9.35, *p*<.0001. Each of these analyses were significant, supporting H1a.

H1b predicted that improvements in loneliness would be greater in the G4H condition than the CBT condition. To assess this, the interaction terms between time and condition were considered. Although the linear interaction was not significant, *t*(1001)=0.66, *p*=.512, the *quadratic* interaction was significant, *t*(1000)=2.61, *p*=.009. This suggests that the non-linear shape of the slope of recovery differs between the two groups. In this case, we can see that while the treatments had a comparable benefit for loneliness up until T3, a gap widened to the 12 month timepoint. In the second half of the program, loneliness scores continued to improve in the G4H condition (T3-T7 contrast: *t*(1001)=4.46, *p*<.001, but stabilised in the CBT condition (T3-T7 contrast: *t*(999)=2.26, *p=*.315. Therefore, H1b was supported.

Per-protocol test of H2

H2 predictedthat G4H would be non-inferior to CBT in improving depression symptom severity, conceptualised as a change score difference that was both non-significant and less than 2.20 at all follow-up timepoints. To assess this, planned contrasts were conducted which assessed the differences between conditions in the degree of change between T0 and each timepoint. All seven of these contrasts were non-significant (*p*s>.264), and crucially for H2, indicated that at no timepoint did the difference exceed 2.2 (the largest was -1.71 for the T0-T2 comparison). Therefore, H2 was supported.

Furthermore, the *cubic interaction* of time and condition on depression was significant, *t*(1000)=-2.37, *p*=.018. This means that, although the overall degree of depression symptom improvement did not differ between conditions, the trajectory of depression symptom change differed somewhat between the groups. On inspection of the estimated marginal means, this was attributable to a steeper slope of improvement in the G4H condition, such that the majority of symptom change occurred between T2 and T6: *t*(1002)=4.77, *p*<.0001. By contrast, the rate of change in the CBT condition was more gradual but of similar overall magnitude (e.g., the T2-T6 contrast was non-significant, *t*(999)=2.60, *p*=.190.

Test of H3

H3 predicted that the benefits of G4H for loneliness and depression would be sustained over time. To assess this, planned contrasts were specified comparing T5 and T6, and then T6 and T7. In the G4H condition, there was no significant change in loneliness between program completion and 6 month follow-up, *t*(1001)=-.09, *p*>.999, or between 6 month and 12 month follow-up, *t*(1000)=1.66, *p*=.713. Similarly, for depression, there was no significant change in the G4H condition between program completion and 6 month follow-up, *t*(1001)=.45, *p*>.999, and limited evidence of any change between 6 and 12 month follow ups: *t*(1009)=-2.58, *p*=.164.