Supplementary appendix (S1)

Transition probabilities, i.e. remission, dropout, and relapse rates between Markov states were based on input parameters from multiple studies. In order to include as much of the available evidence as possible, event rates were pooled using random-effects meta-regression as proposed by DerSimonian and Laird [70]. Fixed-effects modelling was not applied because it presumes that all studies are functionally equivalent. This would be the case if sample characteristics that could possibly influence event rates (such as population age) were basically identical in all studies and consequently, the difference in event rates across those studies were only due to sampling error. However, the studies included in this analysis reveal considerable differences with respect to their design, baseline populations, and measurement methods. Hence, random effects were applied to account for interstudy variability. In addition to Cochran’s Q, \( I^2 \) is reported, which is the total amount of true heterogeneity (variance) on the scale of the original effect measure. \( I^2 \) statistics were calculated to quantify the share of dispersion across the effects that is due to true heterogeneity rather than due to sampling error. Afterwards, the estimated transition probabilities are interpolated to the used cycle length of one week. All calculations were done in R using the packages heemod [56] and metafor [57].

All figures follow a similar structure. The first column contains the authors and publication year of the relevant publications. The second reports the number of participants or patients, which have been observed. Column three and four provides the absolute frequency of patients in remission (REM) and in a depressed state (DEP) or the frequency of dropouts and compliant patients, respectively. Column five reports the observed time periods named as the treatment length (TL) for remission and dropout rates or the follow-up (FU) period for relapse rates and rates for spontaneous remission. The last column shows the mean value and the corresponding 95% confidence intervals. The size of the square of the forest plot in column six reflects the respective weighting of the individual studies depending on the number of patients observed. The estimated results are located at the bottom right. The horizontal expansion of the diamond reflects the uncertainty of the point estimator in terms of its confidence interval.
Figure 1 reports the included studies to estimate the pooled remission rate after FCBT. Pooling the reported remission rates of 11 groups from 10 publications results in an estimated remission rate of 0.609 for a treatment length of 12 weeks. The confidence interval ranges from 0.531 to 0.687.

**Figure 1** Remission rates after FCBT

| Author(s) and Year | N  | REM | DEP | TL | 12-week remission rate [95% CI] |
|-------------------|----|-----|-----|----|---------------------------------|
| [61] DeRubeis et al. 2005 | 51 | 24  | 27  | 16 | 0.379 [0.246, 0.513] |
| [62] Elkin et al. 1990  | 37 | 24  | 13  | 16 | 0.544 [0.383, 0.704] |
| [44] Evans et al. 2009  | 16 | 10  | 6   | 12 | 0.625 [0.388, 0.862] |
| [45] Gorner et al. 1998 | 44 | 25  | 19  | 12 | 0.568 [0.422, 0.715] |
| [42] Hautzinger et al. 2004 | 55 | 34  | 21  | 12 | 0.618 [0.490, 0.747] |
| [39] Jarrett et al. 2001 | 130| 87  | 43  | 13 | 0.640 [0.557, 0.722] |
| [43] Jarrett et al. 1998 | 49 | 37  | 12  | 10 | 0.815 [0.706, 0.924] |
| [43] Jarrett et al. 1998 | 30 | 21  | 9   | 10 | 0.764 [0.612, 0.916] |
| [60] Murphy et al. 1984 | 19 | 10  | 9   | 12 | 0.526 [0.302, 0.751] |
| [40] Shaia et al. 1992  | 40 | 23  | 17  | 16 | 0.474 [0.319, 0.628] |
| [41] Thase et al. 1992  | 64 | 50  | 14  | 16 | 0.680 [0.566, 0.794] |

RE Model ($Q = 35.07, T^2 = 0.012; I^2 = 71.5\%, df = 10, p < 0.000$) 0.609 [0.531, 0.687]

**Notes:** CI confidence interval, DEP number of depressed patients, df degrees of freedom, $I^2$ percentage of variation across studies caused by heterogeneity, N number patients, $p$ probability of error, Q Cochran’s Q-test, RE random effects, REM number of remitted patients, $T^2$ tau-squared (between study variance or true heterogeneity between studies), TL treatment lengths (in weeks)
Figure 2 reports the included studies to estimate the pooled remission rate after ICBT. Pooling the reported remission rates of 7 groups from 7 publications results in an estimated remission rate of 0.517 for a treatment length of 12 weeks. The confidence interval ranges from 0.388 to 0.647.

Figure S2 Remission rates for ICBT

| Author(s) and Year | N  | REM | DEP | TL | 12-week remission rate [95% CI] |
|-------------------|----|-----|-----|----|---------------------------------|
| [69] Andersson et al. 2013 | 31 | 13  | 18  | 9  | 0.516 [0.340, 0.692] |
| [58] Hedman et al. 2014   | 905 | 579 | 326 | 12 | 0.640 [0.609, 0.671] |
| [84] Kessler et al. 2009  | 113 | 43  | 70  | 16 | 0.302 [0.217, 0.386] |
| [67] Ruwaard et al. 2009  | 33  | 16  | 17  | 11 | 0.515 [0.344, 0.686] |
| [55] Thio et al. 2010     | 41  | 23  | 18  | 8  | 0.709 [0.570, 0.848] |
| [66] Vernmark et al. 2010 | 27  | 10  | 17  | 8  | 0.500 [0.312, 0.689] |
| [63] Warmerdam et al. 2008| 65  | 21  | 44  | 8  | 0.443 [0.322, 0.564] |

RE Model (Q = 65.10, T² = 0.026; I² = 90.8%, df = 6, p < 0.000) 0.517 [0.388, 0.647]

Notes: CI confidence interval, DEP number of depressed patients, df degrees of freedom, I² percentage of variation across studies caused by heterogeneity, N number patients, p probability of error, Q Cochran’s Q-test, RE random effects, REM number of remitted patients, T² tau-squared (between study variance or true heterogeneity between studies), TL treatment lengths (in weeks)
Figure 3 reports the included studies for estimation of the pooled dropout rate after FCBT. Pooling the reported dropout rates of 11 groups from 10 publications results in an estimated dropout rate of 0.170 for a treatment length of 12 weeks. The confidence interval ranges from 0.135 to 0.206.

**Figure S3 Dropout rates during FCBT**

| Author(s) and Year | N  | dropout | compliant | TL | 12-week dropout rate [95% CI] |
|--------------------|----|---------|-----------|----|-------------------------------|
| [61] DeRubeis et al. 2005 | 60 | 9       | 51        | 16 | 0.115 [0.034, 0.195]          |
| [62] Elkin et al. 1990   | 56 | 19      | 37        | 16 | 0.267 [0.151, 0.383]          |
| [44] Evans et al. 2009  | 25 | 9       | 16        | 12 | 0.360 [0.172, 0.548]          |
| [45] Gortner et al. 1998 | 50 | 6       | 44        | 12 | 0.120 [0.030, 0.210]          |
| [42] Hautzinger et al. 2004 | 65 | 10      | 55        | 12 | 0.154 [0.066, 0.242]          |
| [39] Jarrett et al. 2001 | 156 | 26      | 130       | 13 | 0.155 [0.098, 0.212]          |
| [43] Jarrett et al. 1998 | 60 | 11      | 49        | 10 | 0.216 [0.112, 0.320]          |
| [43] Jarrett et al. 1998 | 34 | 4       | 30        | 10 | 0.139 [0.023, 0.256]          |
| [60] Murphy et al. 1984 | 24 | 5       | 19        | 12 | 0.209 [0.046, 0.371]          |
| [40] Shea et al. 1992   | 59 | 19      | 40        | 16 | 0.253 [0.142, 0.364]          |
| [41] Thase et al. 1992  | 76 | 12      | 64        | 16 | 0.121 [0.048, 0.194]          |

RE Model (Q = 14.91, T² = 0.001; I² = 32.9%, df = 10, p < 0.135)  

---

**Notes:** CI confidence interval, df degrees of freedom, I² percentage of variation across studies caused by heterogeneity, N number patients, p probability of error, Q Cochran’s Q-test, RE random effects, T² tau-squared (between study variance or true heterogeneity between studies), TL treatment lengths (in weeks)
Figure 4 reports the included studies for estimation of the pooled dropout rate after ICBT. Pooling the reported dropout rates of 7 groups from 7 publications results in an estimated dropout rate of 0.170 for a treatment length of 12 weeks. The confidence interval ranges from 0.135 to 0.206.

**Figure S4 Dropout rates during ICBT**

| Author(s) and Year      | N  | REM | DEP | TL | 12-week dropout rate [95% CI] |
|-------------------------|----|-----|-----|----|-------------------------------|
| [69] Andersson et al. 2013 | 33 | 2   | 31  | 9  | 0.080 [0.000, 0.173]          |
| [68] Hedman et al. 2014  | 1203 | 298 | 905 | 12 | 0.248 [0.223, 0.272]          |
| [64] Kessler et al. 2009 | 149 | 36  | 113 | 16 | 0.187 [0.125, 0.250]          |
| [67] Ruwaard et al. 2009 | 36 | 3   | 33  | 11 | 0.091 [0.000, 0.184]          |
| [65] Titov et al. 2010   | 47 | 6   | 41  | 8  | 0.185 [0.074, 0.296]          |
| [66] Vernmark et al. 2010| 29 | 2   | 27  | 8  | 0.102 [0.000, 0.212]          |
| [63] Warmerdam et al. 2008| 80 | 15  | 65  | 8  | 0.268 [0.171, 0.365]          |

RE Model (Q = 27.87, T² = 0.005; I² = 78.5%, df = 6, p < 0.000) 0.172 [0.112, 0.231]

**Notes:** CI confidence interval, df degrees of freedom, I² percentage of variation across studies caused by heterogeneity, N number patients, p probability of error, Q Cochran’s Q-test, RE random effects, T² tau-squared (between study variance or true heterogeneity between studies), TL treatment lengths (in weeks)
Figure 5 reports the included studies for estimation of the pooled relapse rate of remitted patients after CBT. Pooling the reported relapse rates of 9 groups from 9 publications results in an estimated relapse rate of 0.296 for a follow-up period of one year. The confidence interval ranges from 0.232 to 0.360.

**Figure S5** Relapse rates of remitted patients after CBT

| Author(s) and Year | N | REM | DEP | FU | 1-year relapse rate [95% CI] |
|--------------------|---|-----|-----|----|-------------------------------|
| [44] Evans et al. 1992 | 10 | 7   | 3   | 104 | 0.163 [0.000, 0.392] |
| [45] Gortner et al. 1998 | 26 | 14  | 12  | 104 | 0.286 [0.096, 0.436] |
| [42] Hautzinger et al. 2004 | 36 | 25  | 11  | 52  | 0.306 [0.155, 0.456] |
| [46] Hollon et al. 2005 | 35 | 24  | 11  | 52  | 0.314 [0.160, 0.468] |
| [47] Jacobson 1993 | 13 | 11  | 2   | 52  | 0.154 [0.000, 0.350] |
| [39] Jarrett et al. 2001 | 37 | 19  | 18  | 104 | 0.283 [0.138, 0.429] |
| [43] Jarrett et al. 1998 | 18 | 3   | 15  | 104 | 0.592 [0.365, 0.819] |
| [40] Shea et al. 1992 | 23 | 14  | 9   | 78  | 0.282 [0.098, 0.466] |
| [41] Thase et al. 1992 | 50 | 34  | 16  | 52  | 0.320 [0.191, 0.449] |

RE Model (Q = 10.20, T² = 0.002; I² = 21.5%, df = 8, p < 0.252) 0.296 [0.232, 0.360]

Notes: CI confidence interval, DEP number of depressed patients, df degrees of freedom, FU follow-up (in weeks), I² percentage of variation across studies caused by heterogeneity, N number patients, p probability of error, Q Cochran’s Q-test, RE random effects, REM number of remitted patients, T² tau-squared (between study variance or true heterogeneity between studies)
Figure 6 reports the included studies for estimation of the pooled spontaneous remission rate of depressed patients without treatment. Pooling the reported remission rates of 9 groups from 8 publications results in an estimated remission rate of 0.166 for a follow-up period of 12 weeks. The confidence interval ranges from 0.108 to 0.224.

**Figure S6** Spontaneous remission rates

| Author(s) and Year | N | REM | DEP | FU | 12-week spontaneous remission rate [95% CI] |
|--------------------|---|-----|-----|----|------------------------------------------|
| [71] de Graaf et al. 2009 | 92 | 29  | 63  | 26 | 0.160 [0.085, 0.235] |
| [72] Farrer et al. 2011 | 27 | 9   | 18  | 26 | 0.171 [0.029, 0.313] |
| [72] Farrer et al. 2011 | 22 | 2   | 20  | 26 | 0.043 [0.000, 0.128] |
| [42] Hautzinger et al. 2004 | 30 | 4   | 26  | 12 | 0.133 [0.012, 0.255] |
| [64] Kessler et al. 2009 | 101 | 26 | 75  | 35 | 0.097 [0.039, 0.155] |
| [67] Ruwaard et al. 2009 | 16 | 5   | 11  | 11 | 0.336 [0.104, 0.567] |
| [65] Titov et al. 2010 | 39 | 8   | 31  | 11 | 0.222 [0.091, 0.352] |
| [66] Varnmark et al. 2010 | 29 | 4   | 25  | 8  | 0.200 [0.054, 0.345] |
| [63] Warmerdam et al. 2008 | 71 | 15  | 56  | 8  | 0.300 [0.193, 0.406] |

RE Model (Q = 21.08, I² = 62.0%, df = 8, p < 0.007) 0.166 [0.108, 0.224]

**Notes:** CI confidence interval, DEP number of depressed patients, df degrees of freedom, FU follow-up (in weeks), I² percentage of variation across studies caused by heterogeneity, N number patients, p probability of error, Q Cochran’s Q-test, RE random effects, REM number of remitted patients, T² tau-squared (between study variance or true heterogeneity between studies)