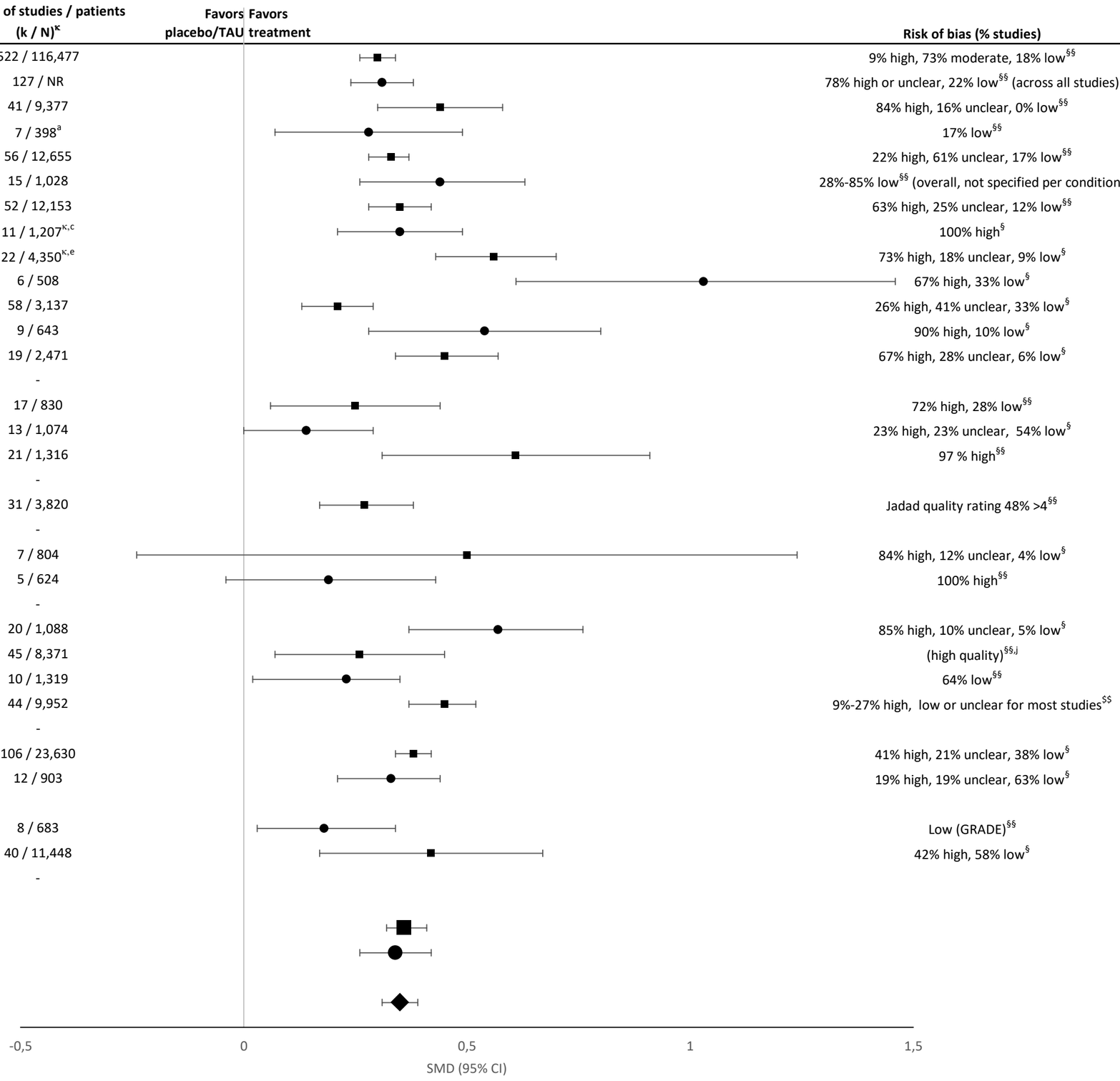
Supplement: Figures 2-4 with annotations

Figure 2 Effect sizes in the largest meta-analyses of pharmacotherapies (squares) and psychotherapies (circles) in comparison to placebo or treatment as usual (TAU) [∞]

Disorder, treatment	SMD (95% CI)[Δ]	Number of studies / patients (k / N)[κ]	Risk of bias (% studies)
MDD, PHA[36]	0.30 (0.26-0.34)[P]	522 / 116,477	9% high, 73% moderate, 18% low[§§]
MDD, PSY[51]	0.31*° (0.24-0.38)[TAU]	127 / NR	78% high or unclear, 22% low[§§] (across all studies)
Panic disorder, PHA[53]	0.44 (0.30-0.58)[P]	41 / 9,377	84% high, 16% unclear, 0% low[§§]
Panic disorder, PSY[6]	0.28 (0.07-0.49)[P/TAU]	7 / 398[a]	17% low[§§]
GAD, PHA[56]	0.33° (0.28-0.37)[P]	56 / 12,655	22% high, 61% unclear, 17% low[§§]
GAD, PSY[55]	0.44[Δ,b] (0.26-0.63)[PP/TAU]	15 / 1,028	28%-85% low[§§] (overall, not specified per condition)
SAD, PHA[64]	0.35° (0.28-0.42)[P]	52 / 12,153	63% high, 25% unclear, 12% low[§§]
SAD, PSY[66]	0.35[Δ] (0.21-0.49)[P/PP]	11 / 1,207[κ,c]	100% high[§]
OCD, PHA[72]	0.56[Δ,d] (0.43-0.70)[P]	22 / 4,350[κ,e]	73% high, 18% unclear, 9% low[§]
OCD, PSY[74]	1.03[*,f] (0.61-1.46)[P/PP]	6 / 508	67% high, 33% low[§]
PTSD, PHA[76]	0.21 (0.13-0.29) [P]	58 / 3,137	26% high, 41% unclear, 33% low[§]
PTSD, PSY[54]	0.54 (0.28-0.80)[TAU]	9 / 643	90% high, 10% low[§]
Binge eating, PHA[82]	0.45 (0.34-0.57)[P]	19 / 2,471	67% high, 28% unclear, 6% low[§]
Binge eating, PSY	-[g]	-	
Anorexia nervosa, PHA[83]	0.25[h] (0.06-0.44)[P]	17 / 830	72% high, 28% low[§§]
Anorexia nervosa, PSY[85]	0.14 (0.00-0.29)[TAU]	13 / 1,074	23% high, 23% unclear, 54% low[§]
Bulimia nervosa, PHA[80]	0.61 (0.31-0.91)[P]	21 / 1,316	97 % high[§§]
Bulimia nervosa, PSY[80]	-[i]	-	
Insomnia, PHA[90]	0.27 (0.17-0.38)[P]	31 / 3,820	Jadad quality rating 48% >4[§§]
Insomnia, PSY	-	-	
Somatoform disorder, PHA[91]	0.50[Δ] (-0.24-1.24)[P]	7 / 804	84% high, 12% unclear, 4% low[§]
Somatoform disorder, PSY[92]	0.19 (-0.04-0.43)[TAU]	5 / 624	100% high[§§]
Borderline PD, PHA	-	-	
Borderline PD, PSY[93]	0.57 (0.37-0.76)[TAU]	20 / 1,088	85% high, 10% unclear, 5% low[§]
Substance use disorders, PHA[95]	0.26 (0.07-0.45)[P]	45 / 8,371	(high quality)[§§,j]
Substance use disorders, PSY[96]	0.23 (0.04-0.43)[TAU]	10 / 1,319	64% low[§§]
ADHD, PHA[99]	0.45 (0.37-0.52)[P]	44 / 9,952	9%-27% high, low or unclear for most studies[$§]
ADHD, PSY	-	-	
Schizophrenia, PHA[103]	0.38*°[,k] (0.34-0.42)[P]	106 / 23,630	41% high, 21% unclear, 38% low[§]
Schizophrenia, PSY[111]	0.33 (0.21-0.44)[TAU]	12 / 903	19% high, 19% unclear, 63% low[§]
Bipolar disorder, PHA	-[l]		
Bipolar disorder, PSY[121]	0.18[m] (0.03-0.34)[TAU]	8 / 683	Low (GRADE)[§§]
Bipolar depression, PHA[113]	0.42 (0.17-0.67)[P]	40 / 11,448	42% high, 58% low[§]
Bipolar depression, PSY	-	-	
Total PHA	0.36 (0.32-0.41)		
Total PSY	0.34 (0.26-0.42)		
Total overall	0.35 (0.31-0.39)		



Note

ADHD: attention deficit hyperactivity disorder; GAD: generalized anxiety disorder; MDD: major depressive disorder; NR: not reported; OCD: obsessive compulsive disorder; P: pill placebo; PD: personality disorder; PHA: pharmacotherapy; PP: psychological placebo; PSY: psychotherapy; PTSD: posttraumatic stress disorder; SAD: social anxiety disorder; SMD: standardized mean difference

[∞] if an available network meta-analysis contained few direct comparisons for a specific condition, we present results of a standard meta-analysis if the latter contained more relevant studies (results of the network meta-analyses concerned can be found in table S1)

[κ] if network meta-analyses did not report an overall effect size across treatments, we present k and N of the available direct comparisons of relevance

[Δ] if a (network) meta-analysis reported multiple results for different treatments, we meta-analytically calculated an aggregated mean effect across the relevant treatments (all other data can be found in table S1)

* adjusted for risk of bias by the authors of the respective meta-analysis (e.g., evaluation of high-quality studies only)

° adjusted for small-study effects by the authors of the respective meta-analysis

[§] risk of bias rating: overall risk of bias of a meta-analysis was rated as high if at least one domain was rated as high by the authors of the meta-analysis; overall risk of bias was rated as low if most domains were rated as low (e.g. 3 of 4 or 4 of 6) and none was rated as high and overall risk of bias was rated as unclear if most domains were rated as unclear and none was rated as high, following Higgins et al. (Br Med J 2011; 343: d5928)

[§§] risk of bias / quality judgment as reported by the authors of the respective meta-analysis

[a] N not reported, estimated from included trials[6]

[b] the large effect size of 1.44 in comparison to pill placebo was not regarded as reliable by Carl et al. (p. 12).[55] Correspondingly, we only present the effects in comparison to psychological placebo and treatment as usual

[c] the total number of studies / patients included in this network meta-analysis is 101/13,164[66]

[d] we calculated SMDs by dividing the mean differences reported by Skapinakis et al. excluding waiting list controlled trials by the pooled standard deviation of the YBOCS mean post-therapy across all treatments (SD_{pooled}=7.11)[72]

[e] the total number of studies / patients included in this network meta-analysis is 54/6,652[72]

[f] as CIs were not reported for adjusted SMDs, the CI of the unadjusted SMDs are presented here[74]

[g] for psychotherapy of binge eating, the included meta-analyses reported only data for mixed controls.[81,82] Thus, no effect for psychotherapy was available

[h] since de Vos et al. identified one study as an outlier, we present data with the outlier removed.[83] As CIs were not reported for the SMD with the outlier removed, the CI of the SMD across all studies is used here

[i] for psychotherapy the Svaldi et al. meta-analysis included only comparisons with waiting list[80]

[j] information on quality of included trials was only available for studies on naltrexone (i.e., 100% high quality), not for studies on acamprosate[95]

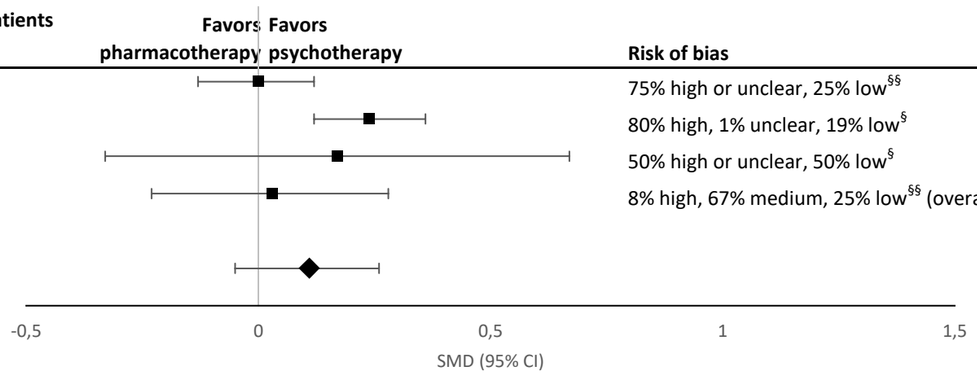
[k] personal communication Maximilian Huhn, 22 August 2019[103]

[l] for drug treatment of bipolar disorder no valid meta-analysis reporting SMDs was available

[m] the reported effect size refers to individual PT (additional data on group PT can be found in table S1)[121]

Figure 3 Effect sizes in the largest meta-analyses for head-to-head comparisons of psychotherapies (PSY) vs. pharmacotherapies (PHA)[∞]

Disorder	SMD (95% CI)[Δ]	Number of studies / patients (k / N)[κ]	Risk of bias
MDD[37]	0.00 (-0.13-0.12)	50 / NR[a]	75% high or unclear, 25% low[§§]
SAD[66]	0.24[Δ] (0.12-0.36)	8 / 864[κ,b]	80% high, 1% unclear, 19% low[§]
OCD[74]	0.17° (-0.33-0.67)[c]	4 / 409	50% high or unclear, 50% low[§]
PTSD[126]	0.03[d] (-0.23-0.28)	4 / NR[κ,e]	8% high, 67% medium, 25% low[§§] (overall, not specified per condition)
Total PSY vs. PHA	0.11 (-0.05-0.26)		



Note

MDD: major depressive disorder; NR: not reported; OCD: obsessive compulsive disorder; PHA: pharmacotherapies; PSY: psychotherapies; PTSD: posttraumatic stress disorder; SAD: social anxiety disorder; SMD: standardized mean difference

[∞] if an available network meta-analysis contained few direct comparisons for a specific condition, we present results of a standard meta-analysis if the latter contained more relevant studies (results of the network meta-analyses concerned can be found in table S1)

[k] if network meta-analyses did not report an overall effect size across treatments, we present k and N of the available direct comparisons of relevance

[Δ] if a (network) meta-analysis reported multiple results for different treatments, we meta-analytically calculated an aggregated mean effect across the relevant treatments (all other data can be found in table S1)

° adjusted for small-study effects by the authors of the respective meta-analysis

[§] risk of bias rating: overall risk of bias of a meta-analysis was rated as high if at least one domain was rated as high by the authors of the meta-analysis; overall risk of bias was rated as low if most domains were rated as low (e.g. 3 of 4 or 4 of 6) and none was rated as high and overall risk of bias was rated as unclear if most domains were rated as unclear and none was rated as high, following Higgins et al. (Br Med J 2011; 343: d5928)

[§§] risk of bias / quality judgment of the authors of the respective meta-analysis

[a] the total number of studies / patients included in this meta-analysis is 101/11,910[37]

[b] the total number of studies / patients included in this network meta-analysis is 101/13,164[66]

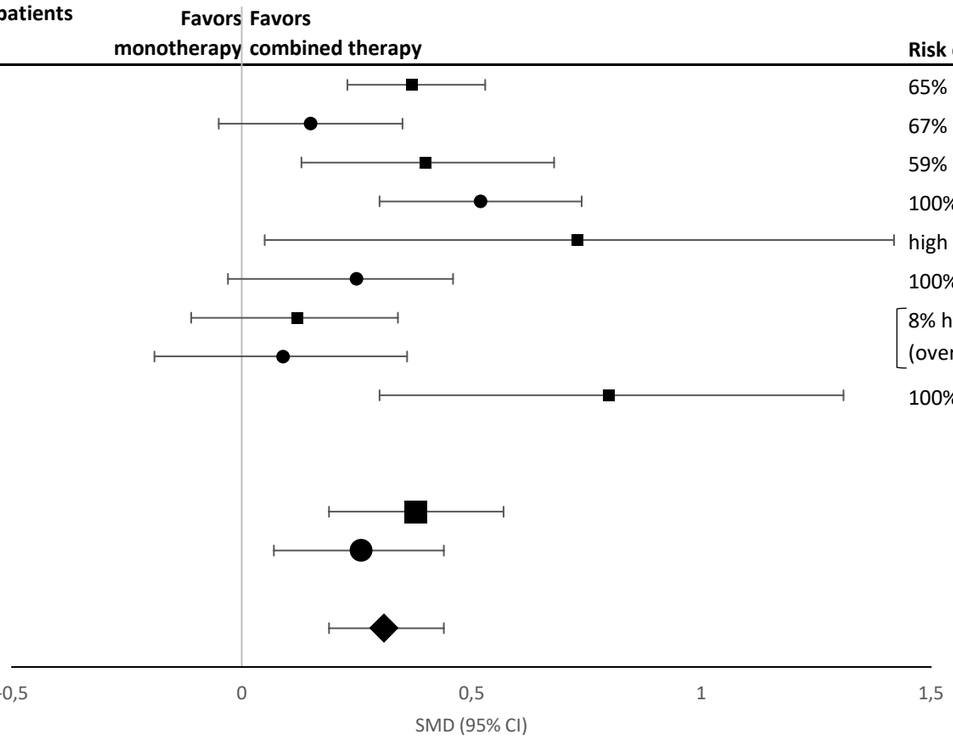
[c] as CIs were not reported for adjusted SMDs, the CI of the unadjusted SMDs are presented here[74]

[d] the reported effect size refers to short-term effects (additional data on long-term effects can be found in table S1)[126]

[e] the total number of studies / patients included in this network meta-analysis is 12/922[126]

Figure 4 Effect sizes in the largest meta-analyses for combined therapy vs. pharmacological (squares) or psychological (circles) monotherapy[∞]

Disorder, treatment	SMD (95% CI)[Δ]	Number of studies / patients (k / N)[κ]	Risk of bias
MDD, COM vs. PHA[37]	0.37[a] (0.23-0.53)	41 / NR[b]	65% high or unclear, 35% low[§§]
MDD, COM vs. PSY[37]	0.15[a] (-0.05-0.35)	19 / NR[b]	67% high or unclear, 33% low[§§]
SAD, COM vs. PHA[66]	0.40[Δ] (0.13-0.68)	5 / 629[κ,c]	59% high, 2% uncertain, 39% low[§]
SAD, COM vs. PSY[66]	0.52[Δ] (0.30-0.74)	2 / 461[κ,c]	100% high[§]
OCD, COM vs. PHA[72]	0.73 (0.05-1.42)[d]	1 / 12[κ,e]	high for 50% of domains[§§]
OCD, COM vs. PSY[74]	0.25° (-0.03-0.46)[f]	6 / 447	100% high[§]
PTSD, COM vs. PHA[126]	0.12[g] (-0.11-0.34)	5 / NR[h]	8% high, 67% medium, 25% low[§§] (overall, not specified per condition)
PTSD, COM vs. PSY[126]	0.09[g] (-0.19-0.36)	2 / NR[h]	
ADHD, COM vs. PHA[134]	0.80 (0.30-1.31)	2 / 65	100% high[§]
ADHD, COM vs. PSY	-	-	
Total, COM vs. PHA	0.38 (0.19-0.57)		
Total, COM vs. PSY	0.26 (0.07-0.44)		
Total, COM vs. MONO	0.31 (0.19-0.44)		



Note

ADHD: attention deficit hyperactivity disorder; COM: combined therapy; MDD: major depressive disorder; MONO: monotherapy; NR: not reported; OCD: obsessive compulsive disorder; PHA: pharmacotherapy; PSY: psychotherapy; PTSD: posttraumatic stress disorder; SAD: social anxiety disorder; SMD: standardized mean difference

[∞] if an available network meta-analysis contained few direct comparisons for a specific condition, we present results of a standard meta-analysis if the latter contained more relevant studies (results of the network meta-analyses concerned can be found in table S1)

[κ] if network meta-analyses did not report an overall effect size across treatments, we present k and N of the available direct comparisons of relevance

[Δ] if a (network) meta-analysis reported multiple results for different treatments, we meta-analytically calculated an aggregated mean effect across the relevant treatments (all other data can be found in table S1)

° adjusted for small-study effects by the authors of the respective meta-analysis

[§] risk of bias rating: overall risk of bias of a meta-analysis was rated as high if at least one domain was rated as high by the authors of the meta-analysis; overall risk of bias was rated as low if most domains were rated as low (e.g. 3 of 4 or 4 of 6) and none was rated as high and overall risk of bias was rated as unclear if most domains were rated as unclear and none was rated as high, following Higgins et al. (Br Med J 2011; 343: d5928)

[§§] risk of bias / quality judgment of the authors of the respective meta-analysis

[a] the reported effect size refers to short-term effects (additional data on long-term effects can be found in table S1)[132]

[b] the total number of studies / patients included in this meta-analysis is 101/11,910[37]

[c] the total number of studies / patients included in this network meta-analysis is 101/13,164[66]

[d] we calculated SMDs by dividing the mean differences reported by Skapinakis et al. excluding waiting list controlled trials by the pooled standard deviation of the YBOCS mean post-therapy across all treatments (SD_{pooled}=7.11)[72]

[e] the total number of studies / patients included in this network meta-analysis is 54/6,652[72]

[f] as CIs were not reported for adjusted SMDs, the CI of the unadjusted SMDs are presented here[74]

[g] the reported effect size refers to short-term effects (additional data on long-term effects can be found in table S1)[126]

[h] the total number of studies / patients included in this network meta-analysis is 12/922[126]