CQ 3-3

【評価シート　エビデンス総体】

診療ガイドライン	日本うつ病学会　うつ病診療ガイドライン 2025
対象	精神症性の特徴を伴わない中等度から重度のうつ病
介入	新規抗うつ薬
対照	プラセボ

*1 各ドメインは“高（−2）”“中/疑い（−1）”“低（0）”の 3 段階
*2 上昇要因は“高（+2）”“中（+1）”“低（0）”の 3 段階
*3 エビデンスの強さは“強（A）”“中（B）”“弱（C）”“非常に弱（D）”の 4 段階．エビデンスの強さについて，RCT は“強（A）”からスタート，観察研究は“弱（C）”からスタート
*4 重要性はアウトカムの重要性（1〜9）

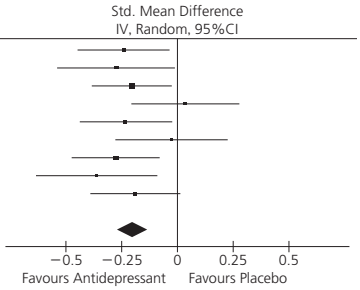
エビデンス総体

アウトカム	研究デザイン/研究数	バイアスリスク*1	非一貫性*1	不精確性*1	非直接性*1	その他（出版バイアスなど）*1	上昇要因（観察研究）*2	リスク人数（アウトカム率）						効果指標（種類）	効果指標統合値	95%信頼区間	エビデンスの強さ*3	重要性*4	コメント
								対照群分母	対照群分子	（%）	介入群分母	介入群分子	（%）						
抑うつ症状の軽減	RCT/9	0	0	0	0	0	0							SMD	−0.20	−0.27：−0.12	強（A）	9	
反応率	RCT/9	0	0	0	0	0	0			43.8			54.5	RR	1.23	1.13：1.32	強（A）	8	
寛解率	RCT/9	0	0	0	0	0	0			26.0			34.4	RR	1.30	1.16：1.45	強（A）	8	
すべての理由による治療中断率	RCT/9	0	0	0	0	0	0							RR	0.98	0.82：1.18	強（A）	7	
有害事象による治療中断率	RCT/9	0	0	0	0	0	0			3.9			5.7	RR	1.60	1.13：2.26	強（A）	7	
少なくとも 1 つの有害事象の出現頻度	RCT/9	0	0	0	0	0	0			66.5			76.4	RR	1.13	1.08：1.18	強（A）	6	

抑うつ症状の軽減

Study or Subgroup	Weight	Std. Mean Difference IV, Random, 95%CI
Higuchi 2009 DUL and PAR-IR	13.0%	−0.24[−0.44, −0.04]
Higuchi 2011 PAR-CR and PAR-IR	7.4%	−0.27[−0.53, −0.01]
Higuchi 2016 VEN	16.1%	−0.20[−0.38, −0.02]
Hirayasu 2011 ESC	8.9%	0.04[−0.20, 0.28]
Hirayasu 2011 ESC and PAR-IR	12.2%	−0.23[−0.43, −0.02]
Inoue 2018 VOR	8.2%	−0.02[−0.27, 0.23]
Inoue 2020 VOR	13.8%	−0.27[−0.47, −0.08]
Kinoshita 2009 MIR	7.1%	−0.36[−0.63, −0.09]
Nishimura 2018 VOR	13.3%	−0.18[−0.38, 0.01]
Total (95%CI)	100.0%	−0.20[−0.27, −0.12]

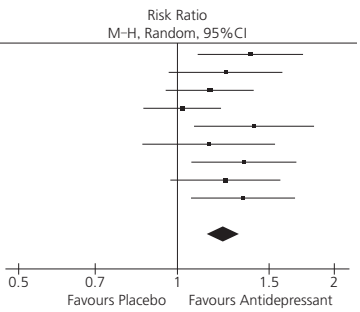
Heterogeneity: Tau²=0.00; Chi²=8.02, df=8 (P=0.43); I²=0%
Test for overall effect: Z=5.35 (P<0.00001)



反応率

Study or Subgroup	Weight	Risk Ratio M-H, Random, 95%CI
Higuchi 2009 DUL and PAR-IR	10.6%	1.39[1.10, 1.75]
Higuchi 2011 PAR-CR and PAR-IR	9.2%	1.24[0.97, 1.59]
Higuchi 2016 VEN	15.0%	1.16[0.96, 1.40]
Hirayasu 2011 ESC	18.6%	1.02[0.86, 1.21]
Hirayasu 2011 ESC and PAR-IR	8.4%	1.41[1.08, 1.83]
Inoue 2018 VOR	6.9%	1.15[0.86, 1.54]
Inoue 2020 VOR	10.6%	1.34[1.07, 1.69]
Kinoshita 2009 MIR	9.7%	1.24[0.97, 1.58]
Nishimura 2018 VOR	11.1%	1.34[1.07, 1.68]
Total (95%CI)	100.0%	1.23[1.13, 1.32]
Total events		

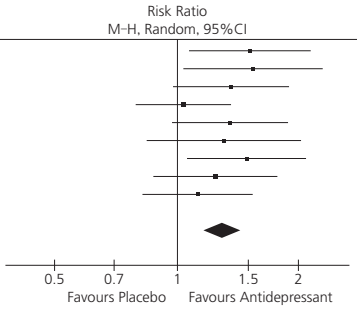
Heterogeneity: Tau²=0.00; Chi²=8.47, df=8 (P=0.39); I²=6%
Test for overall effect: Z=5.15 (P<0.00001)



寛解率

Study or Subgroup	Weight	Risk Ratio M-H, Random, 95%CI
Higuchi 2009 DUL and PAR-IR	10.7%	1.52[1.08, 2.15]
Higuchi 2011 PAR-CR and PAR-IR	8.2%	1.55[1.04, 2.29]
Higuchi 2016 VEN	11.8%	1.37[0.98, 1.90]
Hirayasu 2011 ESC	16.9%	1.04[0.79, 1.37]
Hirayasu 2011 ESC and PAR-IR	11.7%	1.36[0.98, 1.89]
Inoue 2018 VOR	6.7%	1.31[0.85, 2.02]
Inoue 2020 VOR	11.1%	1.49[1.06, 2.09]
Kinoshita 2009 MIR	10.4%	1.25[0.88, 1.77]
Nishimura 2018 VOR	12.6%	1.13[0.82, 1.55]
Total (95%CI)	100.0%	1.30[1.16, 1.45]
Total events		

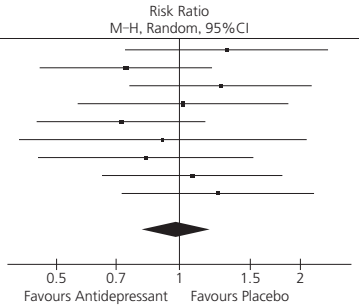
Heterogeneity: Tau²=0.00; Chi²=5.76, df=8 (P=0.67); I²=0%
Test for overall effect: Z=4.51 (P<0.00001)



すべての理由による治療中断率

Study or Subgroup	Weight	Risk Ratio M-H, Random, 95%CI
Higuchi 2009 DUL and PAR-IR	10.1%	1.31[0.74, 2.34]
Higuchi 2011 PAR-CR and PAR-IR	14.3%	0.74[0.45, 1.20]
Higuchi 2016 VEN	12.6%	1.27[0.76, 2.13]
Hirayasu 2011 ESC	9.6%	1.03[0.57, 1.86]
Hirayasu 2011 ESC and PAR-IR	14.8%	0.72[0.45, 1.16]
Inoue 2018 VOR	5.1%	0.92[0.40, 2.08]
Inoue 2020 VOR	9.1%	0.83[0.45, 1.53]
Kinoshita 2009 MIR	13.0%	1.08[0.65, 1.81]
Nishimura 2018 VOR	11.4%	1.25[0.72, 2.16]
Total (95%CI)	100.0%	0.98[0.82, 1.18]
Total events		

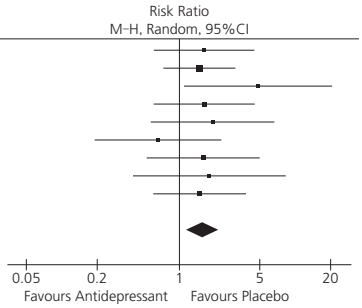
Heterogeneity: Tau²=0.00; Chi²=6.08, df=8 (P=0.64); I²=0%
Test for overall effect: Z=0.18 (P=0.86)



有害事象による治療中断率

Study or Subgroup	Weight	Risk Ratio M-H, Random, 95%CI
Higuchi 2009 DUL and PAR-IR	12.6%	1.66[0.63, 4.38]
Higuchi 2011 PAR-CR and PAR-IR	23.8%	1.50[0.74, 3.05]
Higuchi 2016 VEN	5.7%	4.68[1.10, 19.94]
Hirayasu 2011 ESC	12.5%	1.64[0.62, 4.35]
Hirayasu 2011 ESC and PAR-IR	8.1%	1.95[0.58, 6.53]
Inoue 2018 VOR	7.7%	0.67[0.19, 2.32]
Inoue 2020 VOR	9.7%	1.62[0.54, 4.89]
Kinoshita 2009 MIR	5.4%	1.82[0.41, 8.03]
Nishimura 2018 VOR	14.6%	1.50[0.61, 3.70]
Total (95%CI)	100.0%	1.60[1.13, 2.26]
Total events		

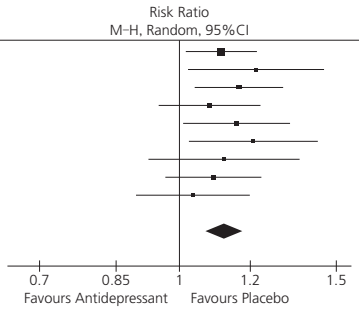
Heterogeneity: Tau²=0.00; Chi²=4.22, df=8 (P=0.84); I²=0%
Test for overall effect: Z=2.68 (P=0.007)



少なくとも 1 つの有害事象の出現頻度

Study or Subgroup	Weight	Risk Ratio M-H, Random, 95%CI
Higuchi 2009 DUL and PAR-IR	22.8%	1.12[1.02, 1.22]
Higuchi 2011 PAR-CR and PAR-IR	6.4%	1.22[1.02, 1.45]
Higuchi 2016 VEN	14.7%	1.17[1.04, 1.31]
Hirayasu 2011 ESC	11.5%	1.09[0.95, 1.24]
Hirayasu 2011 ESC and PAR-IR	10.1%	1.16[1.01, 1.33]
Inoue 2018 VOR	6.9%	1.21[1.02, 1.43]
Inoue 2020 VOR	5.1%	1.13[0.93, 1.37]
Kinoshita 2009 MIR	13.1%	1.09[0.97, 1.23]
Nishimura 2018 VOR	9.4%	1.04[0.90, 1.20]
Total (95%CI)	100.0%	1.13[1.08, 1.18]
Total events		

Heterogeneity: Tau²=0.00; Chi²=3.90, df=8 (P=0.87); I²=0%
Test for overall effect: Z=5.33 (P<0.00001)



	判断						
問題	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
望ましい効果	わずか	小	中	大		さまざま	わからない
望ましくない効果	大	中	小	わずか		さまざま	わからない
エビデンスの確実性	非常に低	低	中	高			わからない
価値観	重要な不確実性，ばらつきあり	重要な不確実性，ばらつきの可能性あり	重要な不確実性，ばらつきはおそらくなし	重要な不確実性，ばらつきなし			
効果のバランス	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	わからない
必要資源量	大きなコスト	中等度のコスト	無視できる程度のコスト	中等度の節減	大きな節減	さまざま	わからない
必要資源のエビデンスの確実性	非常に低	低	中	高			採用研究なし
費用対効果	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	採用研究なし
公平性	減る	おそらく減る	おそらく影響なし	おそらく増える	増える	さまざま	わからない
容認性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
実行可能性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない

CQ 3-4

【評価シート　エビデンス総体】

診療ガイドライン	日本うつ病学会　うつ病診療ガイドライン 2025
対象	中等度から重度のうつ病
介入	精神療法を実施
対照	無治療あるいはプラセボ

*1 各ドメインは“高（−2）”“中/疑い（−1）”“低（0）”の 3 段階
*2 上昇要因は“高（+2）”“中（+1）”“低（0）”の 3 段階
*3 エビデンスの強さは“強（A）”“中（B）”“弱（C）”“非常に弱（D）”の 4 段階．エビデンスの強さについて，RCT は“強（A）”からスタート，観察研究は“弱（C）”からスタート
*4 重要性はアウトカムの重要性（1〜9）

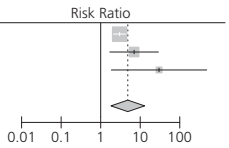
エビデンス総体

アウトカム	研究デザイン/研究数	バイアスリスク*1	非一貫性*1	不精確性*1	非直接性*1	その他（出版バイアスなど）*1	上昇要因（観察研究）*2	リスク人数（アウトカム率）						効果指標（種類）	効果指標統合値	95%信頼区間	エビデンスの強さ*3	重要性*4	コメント
								対照群分母	対照群分子	（%）	介入群分母	介入群分子	（%）						
反応率	RCT/3	−1	0	−1	0	−1	0	203	30	14.78	188	104	55.32	RR	4.85	1.82；12.94	弱（C）	9	
寛解率	RCT/9	−1	0	0	0	−1	0	697	209	29.99	640	391	61.09	RR	2.70	1.66；4.41	弱（C）	7	
抑うつ症状の軽減	RCT/19	−1	0	0	0	−1	0	874	NA	NA	842	NA	NA	SMD	−1.25	−1.65；−0.86	弱（C）	7	
すべての理由による治療中断率	RCT/17	−1	0	−1	0	−1	0	1067	145	13.59	1025	200	19.51	RR	1.31	0.83；2.08	弱（C）	7	

反応率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Markkula 2019	78	124	28	134	3.01	[2.11; 4.30]
Rickhi 2011	14	39	2	39	7.00	[1.70; 28.77]
Choi 2012	12	25	0	30	29.90	[1.86; 480.71]
Random effects model		188		203	4.85	[1.82; 12.94]

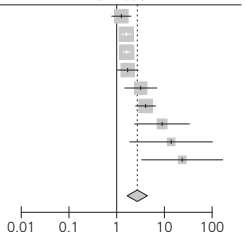
Heterogeneity: I^2=46.8%, τ^2=0.3740, p=0.1526



寛解率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Mohr 2011	21	41	18	44	1.25	[0.79; 1.99]
Arjadi 2018	78	120	63	154	1.59	[1.26; 2.00]
Patel 2017/Weobong 2017	147	247	91	248	1.62	[1.34; 1.97]
Hatcher 2018	16	22	12	28	1.70	[1.03; 2.79]
Choi 2012	16	25	6	30	3.20	[1.48; 6.94]
Naeem 2014	64	94	15	89	4.04	[2.50; 6.54]
Arean 1993	17	19	2	20	8.95	[2.38; 33.62]
Rickhi 2011	12	39	1	45	13.85	[1.88; 101.74]
Jamison 1995	20	33	1	39	23.64	[3.35; 166.82]
Random effects model		640		697	2.70	[1.66; 4.41]

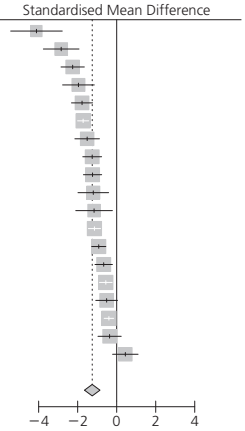
Heterogeneity: I^2=76.4%, τ^2=0.3907, p<0.0001



抑うつ症状の軽減

Study	Experimental Total	Control Total	SMD	95% CI
Hojatifar 2017	15	15	−4.11	[−5.43; −2.79]
Arean 1993	19	20	−2.84	[−3.76; −1.93]
Jamison 1995	33	39	−2.26	[−2.85; −1.66]
Lindegaard 2019	20	16	−1.96	[−2.77; −1.15]
Rickhi 2011	35	39	−1.78	[−2.32; −1.23]
Markkula 2019	124	134	−1.72	[−2.01; −1.43]
Choi 2012	23	28	−1.51	[−2.14; −0.88]
Hamdan-Mansour 2009	44	36	−1.26	[−1.74; −0.77]
Alavi 2016	43	41	−1.24	[−1.71; −0.77]
Beach 1992	15	15	−1.20	[−1.99; −0.41]
Hamamci 2006	10	11	−1.16	[−2.10; −0.22]
Naeem 2014	94	89	−1.14	[−1.45; −0.83]
Naeem 2015	66	63	−0.92	[−1.28; −0.56]
Mohr 2011	40	41	−0.67	[−1.11; −0.22]
Kramer 2014	74	77	−0.56	[−0.88; −0.23]
Krejtz 2018	26	25	−0.51	[−1.07; 0.04]
Arjadi 2018	120	145	−0.40	[−0.64; −0.15]
Hatcher 2018	22	22	−0.36	[−0.96; 0.23]
Costa 2016	19	18	0.44	[−0.21; 1.10]
Random effects model	842	874	−1.25	[−1.65; −0.86]

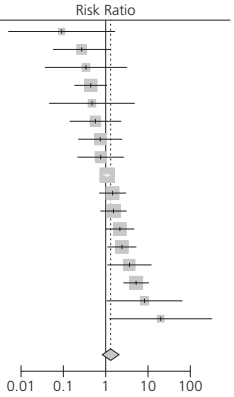
Heterogeneity: I^2=88.4%, τ^2=0.6818, p<0.0001



すべての理由による治療中断率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Hamdan-Mansour 2009	0	44	4	36	0.09	[0.01; 1.64]
Naeem 2014	2	94	7	89	0.27	[0.06; 1.27]
Mohr 2011	1	40	3	41	0.34	[0.04; 3.15]
Lindegaard 2019	5	20	9	16	0.44	[0.19; 1.06]
Costa 2016	1	19	2	18	0.47	[0.05; 4.78]
Naeem 2015	3	66	5	63	0.57	[0.14; 2.30]
Rickhi 2011	4	35	6	39	0.74	[0.23; 2.42]
Alavi 2016	4	43	5	41	0.76	[0.22; 2.64]
Kramer 2014	57	74	55	77	1.08	[0.89; 1.30]
Patel 2017/Weobong 2017	17	230	12	236	1.45	[0.71; 2.98]
Markkula 2019	17	124	12	134	1.53	[0.76; 3.08]
Hatcher 2018	13	22	6	22	2.17	[1.01; 4.66]
Usaf 1990	12	19	6	23	2.42	[1.12; 5.22]
Choi 2012	9	23	3	28	3.65	[1.12; 11.94]
Arjadi 2018	39	120	9	145	5.24	[2.64; 10.37]
Jamison 1995	7	33	1	39	8.27	[1.07; 63.84]
Arean 1993	9	19	0	20	19.97	[1.24; 320.49]
Random effects model		1025		1067	1.31	[0.83; 2.08]

Heterogeneity: I^2=69.0%, τ^2=0.5332, p<0.0001



	判断						
問題	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
望ましい効果	わずか	小	中	大		さまざま	わからない
望ましくない効果	大	中	小	わずか		さまざま	わからない
エビデンスの確実性	非常に低	低	中	高			わからない
価値観	重要な不確実性，ばらつきあり	重要な不確実性，ばらつきの可能性あり	重要な不確実性，ばらつきはおそらくなし	重要な不確実性，ばらつきなし			
効果のバランス	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	わからない
必要資源量	大きなコスト	中等度のコスト	無視できる程度のコスト	中等度の節減	大きな節減	さまざま	わからない
必要資源のエビデンスの確実性	非常に低	低	中	高			採用研究なし
費用対効果	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	採用研究なし
公平性	減る	おそらく減る	おそらく影響なし	おそらく増える	増える	さまざま	わからない
容認性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
実行可能性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない

CQ 6-4

【評価シート　エビデンス総体】

診療ガイドライン	日本うつ病学会　うつ病診療ガイドライン 2025
対象	中等度から重度の老年期うつ病
介入	抗うつ薬
対照	プラセボ

*1 各ドメインは“高（−2）”“中/疑い（−1）”“低（0）”の 3 段階
*2 上昇要因は“高（+2）”“中（+1）”“低（0）”の 3 段階
*3 エビデンスの強さは“強（A）”“中（B）”“弱（C）”“非常に弱（D）”の 4 段階．エビデンスの強さについて，RCT は“強（A）”からスタート，観察研究は“弱（C）”からスタート
*4 重要性はアウトカムの重要性（1〜9）

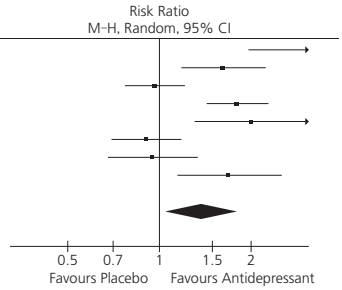
エビデンス総体

アウトカム	研究デザイン/研究数	バイアスリスク*1	非一貫性*1	不精確性*1	非直接性*1	その他（出版バイアスなど）*1	上昇要因（観察研究）*2	リスク人数（アウトカム率）						効果指標（種類）	効果指標統合値	95%信頼区間	エビデンスの強さ*3	重要性*4	コメント
								対照群分母	対照群分子	（%）	介入群分母	介入群分子	（%）						
反応率	RCT/8	−1	−1	0	0	−1	0							RR	1.38	1.04：1.83	中（B）	8	
抑うつ症状の軽減	RCT/9	−1	−1	0	0	0	0							SMD	−0.62	−0.92：−0.33	中（B）	9	
寛解率	RCT/5	0	−1	−1	0	0	0							RR	1.27	0.98：1.65	弱（C）	8	
すべての理由による治療中断率	RCT/9	−1	0	0	0	0	0							RR	1.06	0.88：1.27	中（B）	7	
有害事象による治療中断率	RCT/9	−1	0	0	0	0	0							RR	1.94	1.30：2.88	中（B）	7	
少なくとも 1 つの有害事象の出現率	RCT/9	−1	0	0	0	0	0							RR	1.11	1.02：1.21	中（B）	6	

反応率

Study or Subgroup	Weight	Risk Ratio M-H, Random, 95% CI
Chen 2011 ESC	1.0%	31.25[1.98, 493.21]
Emsley 2018 ESC	14.1%	1.62[1.18, 2.23]
Kasper 2005 ESC	15.6%	0.97[0.78, 1.22]
Katona 2012 DUL/VOR	15.5%	1.82[1.44, 2.29]
Raskin 2007 DUL	12.1%	2.03[1.30, 3.16]
Robinson 2014 DUL	15.1%	0.91[0.70, 1.18]
Schatzberg 2006 VEN	13.7%	0.95[0.68, 1.35]
Schweizer 1998 IMI	12.9%	1.70[1.15, 2.53]
Total (95% CI)	100.0%	1.38[1.04, 1.83]
Total events		

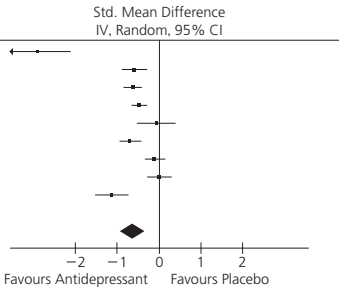
Heterogeneity: Tau²=0.12; Chi²=38.71, df=7 (P<0.00001); I²=82%
Test for overall effect: Z=2.25 (P=0.02)



抑うつ症状の軽減

Study or Subgroup	Weight	Std. Mean Difference IV, Random, 95% CI
Chen 2011 ESC	6.6%	−2.92[−3.73, −2.11]
Emsley 2018 ESC	11.8%	−0.59[−0.87, −0.31]
Kasper 2005 ESC	12.4%	−0.62[−0.84, −0.41]
Katona 2012 DUL/VOR	12.5%	−0.47[−0.67, −0.27]
Lin 2022 SER	10.0%	−0.06[−0.53, 0.40]
Raskin 2007 DUL	12.1%	−0.68[−0.92, −0.43]
Robinson 2014 DUL	12.1%	−0.10[−0.35, 0.14]
Schatzberg 2006 VEN	11.8%	0.00[−0.29, 0.29]
Schweizer 1998 IMI	10.7%	−1.14[−1.53, −0.75]
Total (95% CI)	100.0%	−0.62[−0.92, −0.33]

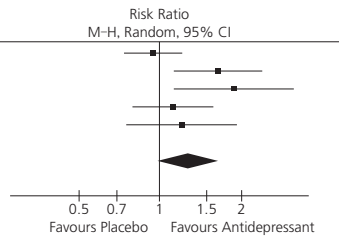
Heterogeneity: Tau²=0.17; Chi²=74.01, df=8 (P=0.00001); I²=89%
Test for overall effect: Z=4.15 (P<0.0001)



寛解率

Study or Subgroup	Weight	Risk Ratio M-H, Random, 95% CI
Kasper 2005 ESC	26.6%	0.95[0.74, 1.22]
Katona 2012 DUL/VOR	20.5%	1.65[1.14, 2.39]
Raskin 2007 DUL	14.7%	1.89[1.13, 3.18]
Robinson 2014 DUL	22.0%	1.13[0.80, 1.58]
Schatzberg 2006 VEN	16.1%	1.21[0.75, 1.95]
Total (95% CI)	100.0%	1.27[0.98, 1.65]
Total events		

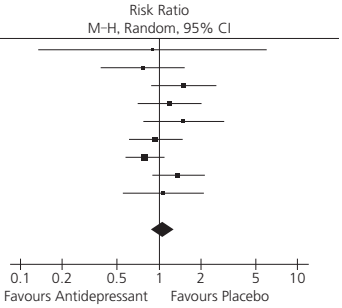
Heterogeneity: Tau²=0.05; Chi²=9.67, df=4 (P=0.05); I²=59%
Test for overall effect: Z=1.81 (P=0.07)



すべての理由による治療中断

Study or Subgroup	Weight	Risk Ratio M-H, Random, 95% CI
Chen 2011 ESC	0.9%	0.90[0.14, 5.92]
Emsley 2018 ESC	6.7%	0.76[0.38, 1.52]
Kasper 2005 ESC	10.8%	1.51[0.89, 2.56]
Katona 2012 DUL/VOR	10.9%	1.19[0.71, 2.02]
Lin 2022 SER	7.1%	1.50[0.77, 2.92]
Raskin 2007 DUL	15.2%	0.94[0.61, 1.46]
Robinson 2014 DUL	26.0%	0.79[0.58, 1.08]
Schatzberg 2006 VEN	15.2%	1.38[0.90, 2.14]
Schweizer 1998 IMI	7.1%	1.08[0.55, 2.09]
Total (95% CI)	100.0%	1.06[0.88, 1.27]
Total events		

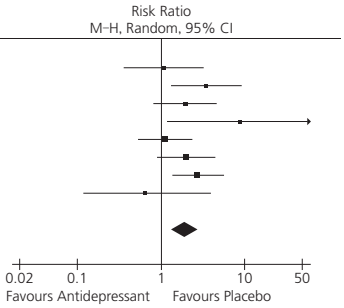
Heterogeneity: Tau²=0.01; Chi²=8.98, df=8 (P=0.34); I²=11%
Test for overall effect: Z=0.62 (P=0.54)



有害事象による治療中断

Study or Subgroup	Weight	Risk Ratio M-H, Random, 95% CI
Chen 2011 ESC		Not estimable
Emsley 2018 ESC	10.3%	1.08[0.36, 3.24]
Kasper 2005 ESC	12.5%	3.54[1.33, 9.38]
Katona 2012 DUL/VOR	14.7%	1.97[0.83, 4.69]
Lin 2022 SER	3.6%	9.00[1.20, 67.69]
Raskin 2007 DUL	18.0%	1.12[0.53, 2.37]
Robinson 2014 DUL	16.6%	2.01[0.91, 4.46]
Schatzberg 2006 VEN	19.6%	2.77[1.37, 5.59]
Schweizer 1998 IMI	4.6%	0.67[0.12, 3.85]
Total (95% CI)	100.0%	1.94[1.30, 2.88]
Total events		

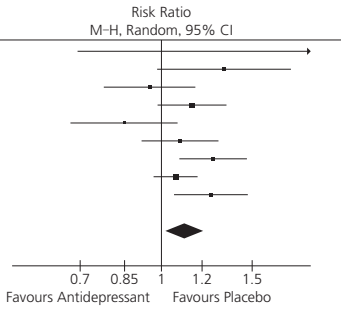
Heterogeneity: $Tau^2=0.08$; $Chi^2=9.35$, df=7 (P=0.23); $I^2=25\%$
Test for overall effect: Z=3.27 (P=0.001)



少なくとも 1 つの有害事象の出現

Study or Subgroup	Weight	Risk Ratio M-H, Random, 95% CI
Chen 2011 ESC	0.2%	5.33[0.69, 41.20]
Emsley 2018 ESC	6.6%	1.32[0.98, 1.76]
Kasper 2005 ESC	10.7%	0.95[0.78, 1.17]
Katona 2012 DUL/VOR	14.4%	1.14[0.98, 1.32]
Lin 2022 SER	8.7%	0.85[0.67, 1.08]
Raskin 2007 DUL	12.8%	1.09[0.92, 1.29]
Robinson 2014 DUL	14.4%	1.26[1.08, 1.46]
Schatzberg 2006 VEN	18.8%	1.07[0.97, 1.17]
Schweizer 1998 IMI	13.4%	1.24[1.06, 1.46]
Total (95% CI)	100.0%	1.11[1.02, 1.21]
Total events		

Heterogeneity: $Tau^2=0.01$; $Chi^2=16.32$, df=8 (P=0.04); $I^2=51\%$
Test for overall effect: Z=2.36 (P=0.02)



	判断						
問題	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
望ましい効果	わずか	小	中	大		さまざま	わからない
望ましくない効果	大	中	小	わずか		さまざま	わからない
エビデンスの確実性	非常に低	低	中	高			わからない
価値観	重要な不確実性，ばらつきあり	重要な不確実性，ばらつきの可能性あり	重要な不確実性，ばらつきはおそらくなし	重要な不確実性，ばらつきなし			
効果のバランス	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	わからない
必要資源量	大きなコスト	中等度のコスト	無視できる程度のコスト	中等度の節減	大きな節減	さまざま	わからない
必要資源のエビデンスの確実性	非常に低	低	中	高			採用研究なし
費用対効果	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	採用研究なし
公平性	減る	おそらく減る	おそらく影響なし	おそらく増える	増える	さまざま	わからない
容認性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
実行可能性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない

CQ 8-3

【評価シート　エビデンス総体】

診療ガイドライン	日本うつ病学会　うつ病診療ガイドライン 2025
対象	不眠症状を伴ううつ病
介入	抗うつ薬＋睡眠薬（$GABA_A$ 受容体作動薬，メラトニン受容体作動薬，オレキシン受容体拮抗薬）
対照	抗うつ薬単剤

*1 各ドメインは“高（−2）”“中/疑い（−1）”“低（0）”の 3 段階
*2 上昇要因は“高（＋2）”“中（＋1）”“低（0）”の 3 段階
*3 エビデンスの強さは“強（A）”“中（B）”“弱（C）”“非常に弱（D）”の 4 段階．エビデンスの強さについて，RCT は“強（A）”からスタート，観察研究は“弱（C）”からスタート
*4 重要性はアウトカムの重要性（1〜9）

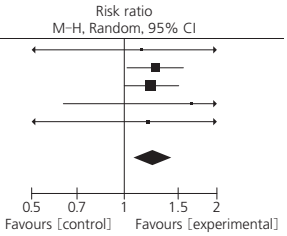
エビデンス総体

アウトカム	研究デザイン/研究数	バイアスリスク*1	非一貫性*1	不精確性*1	非直接性*1	その他（出版バイアスなど）*1	上昇要因（観察研究）*2	リスク人数（アウトカム率）						効果指標（種類）	効果指標統合値	95% 信頼区間	エビデンスの強さ*3	重要性*4	コメント
								対照群分母	対照群分子	（%）	介入群分母	介入群分子	（%）						
長期の抑うつ症状の寛解率	RCT/0																	9	
短期の抑うつ症状の寛解率	RCT/5	−1	0	0	0	0	0	564	192	34.0	557	235	42.2	RR	1.24	1.07：1.43	強（A）	8	
短期の抑うつ症状の改善度	RCT/6	−1	0	0	0	0	0	943	NA	NA	957	NA	NA	SMD	0.20	0.05：0.36	強（A）	7	【非一貫性】NCT00826111 の結果が外れ値だが n が非常に少なく，除いて感度分析を行っても結果にほとんど影響がなかった（0.21, 0.08-0.35）ため「0」とした．
短期の不眠症状の改善度	RCT/4	−1	0	0	0	0	0	439	NA	NA	442	NA	NA	SMD	0.60	0.22：0.97	強（A）	7	
重篤な有害事象の発現率	RCT/7	−1	0	−1	0	0	0	903	9	1.0	912	8	0.9	RR	0.87	0.33：2.27	中（B）	9	【不精確性】95% 信頼区間が 0.33-2.27
有害事象による脱落率	RCT/4	−1	0	−1	0	0	0	865	65	7.51	866	61	7.04	RR	0.93	0.59：1.48	中（B）	7	【不精確性】95% 信頼区間が 0.59-1.48
眠気の発現率	RCT/6	−1	0	−1	0	0	0	850	47	5.53	860	62	7.21	RR	1.28	0.88：1.84	中（B）	6	【不精確性】95% 信頼区間が 0.88-1.84

抑うつ症状の寛解率

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk ratio M-H, Random, 95% CI
Dominguez 1984	8	63	7	63	2.3%	1.14[0.44, 2.96]
Fava 2006	113	269	91	274	44.0%	1.26[1.02, 1.57]
Fava 2011	104	190	85	190	50.5%	1.22[1.00, 1.50]
McCall 2010	8	25	5	26	2.2%	1.66[0.63, 4.40]
NCT00642694	2	10	4	11	1.0%	0.55[0.13, 2.38]
Total		557		564	100.0%	1.24[1.07, 1.43]
Total events:	235		192			

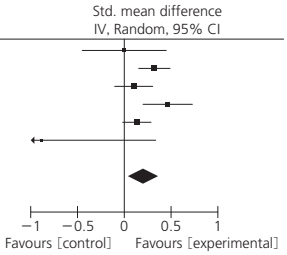
Test for overall effect: Z=2.89 (P=0.004)
Test for subgroup dofferences: Not applicable
Heterogeneity: Tau²=0.00; Chi²=1.61, df=4 (P=0.81); I²=0%



抑うつ症状の改善度

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Std. mean difference IV, Random, 95% CI
Dominguez 1984	13.15	8.26	42	13.15	5.91	36	8.9%	0.00[−0.45, 0.45]
Fava 2006	12.9	6.1	269	10.9	6.3	274	24.6%	0.32[0.15, 0.49]
Fava 2011	11.4	6.1	190	10.8	6.3	190	21.9%	0.10[−0.10, 0.30]
Ji 2007	13.9	5.4	112	11.3	5.8	109	17.0%	0.46[0.20, 0.73]
NCT00435279	15.2	6.36	334	14.3	6.51	330	26.1%	0.14[−0.01, 0.29]
NCT00826111	13.4	8.26	10	20.75	5.91	4	1.5%	−0.89[−2.11, 0.33]
Total			957			943	100.0%	0.20[0.05, 0.36]

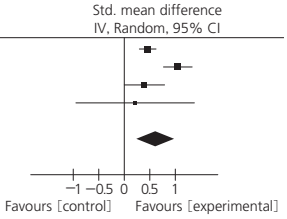
Test for overall effect: Z=2.57 (P=0.01)
Test for subgroup differences: Not applicable
Heterogeneity: Tau²=0.02; Chi²=11.11, df=5 (P=0.05); I²=55%



不眠症状の改善度

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Std. mean difference IV, Random, 95% CI
Fava 2006	3.4	2.04	269	2.5	2.01	274	34.6%	0.44[0.27, 0.61]
Ji 2007	9.7	3.6	112	6	3.5	109	30.8%	1.04[0.76, 1.32]
McCall 2019	11.51	8.3	51	8.71	5.45	52	26.5%	0.40[0.01, 0.79]
NCT00826111	11.2	8.3	10	9.5	5.45	4	8.1%	0.21[−0.96, 1.37]
Total			442			439	100.0%	0.60[0.22, 0.97]

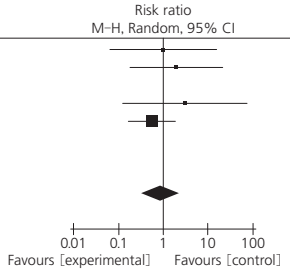
Test for overall effect: Z=3.11 (P=0.002)
Test for subgroup differences: Not applicable
Heterogeneity: Tau²=0.10; Chi²=13.90, df=3 (P=0.003); I²=78%



重篤な有害事象の発現率

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk ratio M-H, Random, 95% CI
Fava 2006	1	269	1	274	12.1%	1.02[0.06, 16.20]
Fava 2011	2	193	1	192	16.2%	1.99[0.18, 21.76]
McCall 2010	0	30	0	30		Not estimable
McCall 2019	1	51	0	52	9.2%	3.06[0.13, 73.36]
NCT00435279	4	341	7	336	62.5%	0.56[0.17, 1.91]
NCT00642694	0	14	0	15		Not estimable
NCT00826111	0	14	0	4		Not estimable
Total		912		903	100.0%	0.87[0.33, 2.27]
Total events:	8		9			

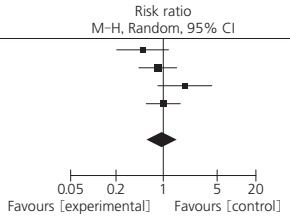
Test for overall effect: Z=0.29 (P=0.77)
Test for subgroup differences: Not applicable
Heterogeneity: Tau²=0.00; Chi²=1.56, df=3 (P=0.67); I²=0%



有害事象による脱落率

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk ratio M-H, Random, 95% CI
Dominguez 1984	7	63	14	63	20.2%	0.50[0.22, 1.16]
Fava 2006	17	269	21	274	29.0%	0.82[0.44, 1.53]
Fava 2011	14	193	7	192	18.7%	1.99[0.82, 4.82]
NCT00435279	23	341	23	336	32.1%	0.99[0.56, 1.72]
Total		866		865	100.0%	0.93[0.59, 1.48]
Total events:	61		65			

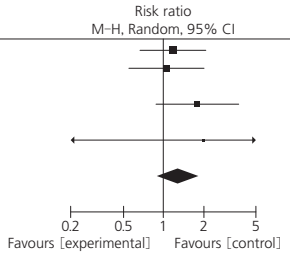
Test for overall effect: Z=0.31 (P=0.76)
Test for subgroup differences: Chi^2=2.48, df=1 (P=0.12); I^2=59.6%
Heterogeneity: Tau^2=0.09; Chi^2=5.14, df=3 (P=0.16); I^2=42%



眠気の発現率

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk ratio M-H, Random, 95% CI
Fava 2006	23	269	20	274	40.7%	1.17[0.66, 2.08]
Fava 2011	17	192	16	190	31.6%	1.05[0.55, 2.02]
McCall 2010	0	30	0	30		Not estimable
NCT00435279	20	341	11	336	26.0%	1.79[0.87, 3.68]
NCT00642694	0	14	0	15		Not estimable
NCT00826111	2	14	0	5	1.6%	2.00[0.11, 35.81]
Total		860		850	100.0%	1.28[0.88, 1.84]
Total events:	62		47			

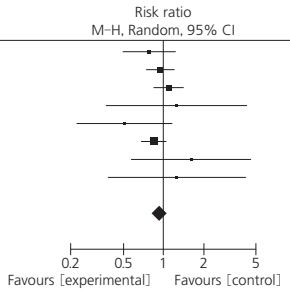
Test for overall effect: Z=1.30 (P=0.19)
Test for subgroup differences: Not applicable
Heterogeneity: Tau^2=0.00; Chi^2=1.37, df=3 (P=0.71); I^2=0%



すべての理由の脱落率

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk ratio M-H, Random, 95% CI
Dominguez 1984	21	63	27	63	8.1%	0.78[0.50, 1.22]
Fava 2006	83	269	89	274	27.1%	0.95[0.74, 1.22]
Fava 2011	74	193	67	192	23.9%	1.10[0.84, 1.43]
McCall 2010	5	30	4	30	1.1%	1.25[0.37, 4.21]
McCall 2019	7	51	14	52	2.5%	0.51[0.22, 1.16]
NCT00435279	101	341	118	336	34.6%	0.84[0.68, 1.05]
NCT00642694	6	14	4	15	1.5%	1.61[0.57, 4.52]
NCT00826111	7	14	2	5	1.2%	1.25[0.38, 4.13]
Total		975		967	100.0%	0.93[0.82, 1.06]
Total events:	304		325			

Test for overall effect: Z=1.14 (P=0.25)
Test for subgroup differences: Not applicable
Heterogeneity: Tau^2=0.00; Chi^2=6.54, df=7 (P=0.48); I^2=0%



	判断						
問題	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
望ましい効果	わずか	小	中	大		さまざま	わからない
望ましくない効果	大	中	小	わずか		さまざま	わからない
エビデンスの確実性	非常に低	低	中	高			わからない
価値観	重要な不確実性，ばらつきあり	重要な不確実性，ばらつきの可能性あり	重要な不確実性，ばらつきはおそらくなし	重要な不確実性，ばらつきなし			
効果のバランス	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	わからない
必要資源量	大きなコスト	中等度のコスト	無視できる程度のコスト	中等度の節減	大きな節減	さまざま	わからない
必要資源のエビデンスの確実性	非常に低	低	中	高			採用研究なし
費用対効果	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	採用研究なし
公平性	減る	おそらく減る	おそらく影響なし	おそらく増える	増える	さまざま	わからない
容認性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
実行可能性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない

CQ 9-3

【評価シート　エビデンス総体】

診療ガイドライン	日本うつ病学会　うつ病診療ガイドライン 2025
対象	初期治療で十分な効果が得られなかったうつ病
介入	現在の抗うつ薬増量
対照	現在の抗うつ薬同量

*1 各ドメインは“高（−2）”“中/疑い（−1）”“低（0）”の 3 段階
*2 上昇要因は“高（+2）”“中（+1）”“低（0）”の 3 段階
*3 エビデンスの強さは“強（A）”“中（B）”“弱（C）”“非常に弱（D）”の 4 段階．エビデンスの強さについて，RCT は“強（A）”からスタート，観察研究は“弱（C）”からスタート
*4 重要性はアウトカムの重要性（1〜9）

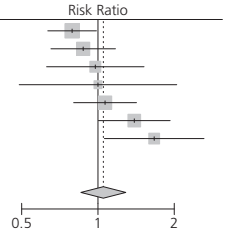
エビデンス総体

アウトカム	研究デザイン/研究数	バイアスリスク*1	非一貫性*1	不精確性*1	非直接性*1	その他（出版バイアスなど）*1	上昇要因（観察研究）*2	リスク人数（アウトカム率）						効果指標（種類）	効果指標統合値	95%信頼区間	エビデンスの強さ*3	重要性*4	コメント
								対照群分母	対照群分子	（%）	介入群分母	介入群分子	（%）						
反応率	RCT/7	−1	−1	−1	−2	−1	0	553	253	45.75	532	243	45.68	RR	1.05	0.86：1.29	弱（C）	9	【非直接性】日本で承認されていない用量を使用した研究が含まれている．
寛解率	RCT/6	−1	0	−1	−2	−1	0	512	151	29.49	496	152	30.65	RR	1.05	0.86：1.29	弱（C）	7	【非直接性】日本で承認されていない用量を使用した研究が含まれている．
抑うつ症状の軽減	RCT/3	−1	0	−1	−2	−1	0	341	NA	NA	323	NA	NA	SMD	−0.18	−0.43：0.08	弱（C）	7	【非直接性】日本で承認されていない用量を使用した研究が含まれている．
すべての理由による治療中断率	RCT/6	0	−1	−1	−2	0	0	512	103	20.12	496	100	20.16	RR	1.07	0.81：1.41	弱（C）	7	【非直接性】日本で承認されていない用量を使用した研究が含まれている．
有害事象による治療中断率	RCT/4	0	−1	−1	−2	0	0	392	22	5.612	371	34	9.164	RR	1.47	0.24：8.98	弱（C）	7	【非直接性】日本で承認されていない用量を使用した研究が含まれている．

反応率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Licht 2002	54	98	69	99	0.79	[0.63; 0.99]
Kornstein 2008	48	124	58	131	0.87	[0.65; 1.17]
Schweizer 1990	18	36	21	41	0.98	[0.63; 1.52]
Ruhe 2009	10	30	10	30	1.00	[0.49; 2.05]
Dornseif 1989	63	181	62	190	1.07	[0.80; 1.42]
Schweizer 2001	30	38	21	37	1.39	[1.00; 1.93]
Kim 2019	20	25	12	25	1.67	[1.06; 2.62]
Random effects model		532		553	1.05	[0.86; 1.29]

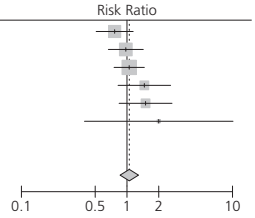
Heterogeneity: I^2=57.7%, τ^2=0.0410, p=0.0275



寛解率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Licht 2002	28	98	37	99	0.76	[0.51; 1.14]
Kornstein 2008	36	124	39	131	0.98	[0.67; 1.43]
Dornseif 1989	51	181	51	190	1.05	[0.75; 1.46]
Schweizer 2001	18	38	12	37	1.46	[0.82; 2.59]
Kim 2019	15	25	10	25	1.50	[0.84; 2.67]
Ruhe 2009	4	30	2	30	2.00	[0.40; 10.11]
Random effects model		496		512	1.05	[0.86; 1.29]

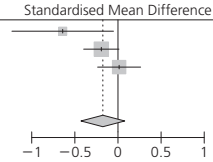
Heterogeneity: I^2=14.7%, τ^2=0.0063, p=0.3197



抑うつ症状の軽減

Study	Experimental Total	Control Total	SMD	95% CI
Kim 2019	25	22	−0.64	[−1.23; −0.05]
Dornseif 1989	180	189	−0.19	[−0.40; 0.01]
Kornstein 2008	118	130	0.01	[−0.24; 0.26]
Random effects model	323	341	−0.18	[−0.43; 0.08]

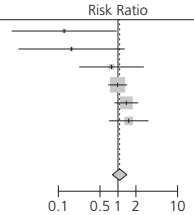
Heterogeneity: I^2=55.4%, τ^2=0.0259, p=0.1060



すべての理由による治療中断率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Ruhe 2009	1	30	8	30	0.13	[0.02; 0.94]
Kim 2019	1	25	6	25	0.17	[0.02; 1.29]
Schweizer 2001	4	38	5	37	0.78	[0.23; 2.68]
Dornseif 1989	45	181	48	190	0.98	[0.69; 1.40]
Kornstein 2008	34	124	26	131	1.38	[0.88; 2.16]
Licht 2002	15	98	10	99	1.52	[0.72; 3.21]
Random effects model		496		512	1.07	[0.81; 1.41]

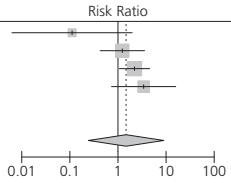
Heterogeneity: I^2=50.4%, τ^2=0.0106, p=0.0728



有害事象による治療中断率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Ruhe 2009	0	30	4	30	0.11	[0.01; 1.98]
Kornstein 2008	7	124	6	131	1.23	[0.43; 3.57]
Dornseif 1989	21	181	10	190	2.20	[1.07; 4.55]
Schweizer 2001	6	36	2	41	3.42	[0.74; 15.88]
Random effects model		371		392	1.47	[0.24; 8.98]

Heterogeneity: I^2=40.6%, τ^2=0.7720, p=0.1680



	判断						
問題	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
望ましい効果	わずか	小	中	大		さまざま	わからない
望ましくない効果	大	中	小	わずか		さまざま	わからない
エビデンスの確実性	非常に低	低	中	高			わからない
価値観	重要な不確実性，ばらつきあり	重要な不確実性，ばらつきの可能性あり	重要な不確実性，ばらつきはおそらくなし	重要な不確実性，ばらつきなし			
効果のバランス	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	わからない
必要資源量	大きなコスト	中等度のコスト	無視できる程度のコスト	中等度の節減	大きな節減	さまざま	わからない
必要資源のエビデンスの確実性	非常に低	低	中	高			採用研究なし
費用対効果	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	採用研究なし
公平性	減る	おそらく減る	おそらく影響なし	おそらく増える	増える	さまざま	わからない
容認性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
実行可能性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない

CQ 9-4

【評価シート　エビデンス総体】

診療ガイドライン	日本うつ病学会　うつ病診療ガイドライン 2025
対象	初期治療で十分な効果が得られなかったうつ病
介入	別の抗うつ薬へ切り替え
対照	現在の抗うつ薬継続

*1 各ドメインは“高（−2）”“中/疑い（−1）”“低（0）”の 3 段階
*2 上昇要因は“高（+2）”“中（+1）”“低（0）”の 3 段階
*3 エビデンスの強さは“強（A）”“中（B）”“弱（C）”“非常に弱（D）”の 4 段階．エビデンスの強さについて，RCT は“強（A）”からスタート，観察研究は“弱（C）”からスタート
*4 重要性はアウトカムの重要性（1〜9）

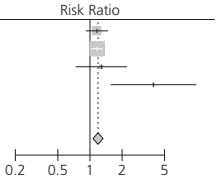
エビデンス総体

アウトカム	研究デザイン/研究数	バイアスリスク*1	非一貫性*1	不精確性*1	非直接性*1	その他（出版バイアスなど）*1	上昇要因（観察研究）*2	リスク人数（アウトカム率）						効果指標（種類）	効果指標統合値	95%信頼区間	エビデンスの強さ*3	重要性*4	コメント
								対照群分母	対照群分子	（%）	介入群分母	介入群分子	（%）						
反応率	RCT/4	−1	0	−1	0	−1	0	678	294	43.36	680	356	52.35	RR	1.19	1.07：1.33	弱（C）	9	
寛解率	RCT/3	−1	0	0	0	−1	0	610	142	23.28	612	198	32.35	RR	1.82	0.99：3.33	弱（C）	7	
抑うつ症状の軽減	RCT/2	−1	0	−1	0	0	0	106	NA	NA	101	NA	NA	SMD	−0.21	−0.48：0.07	弱（C）	7	
すべての理由による治療中断率	RCT/4	0	−1	−1	0	0	0	678	38	5.605	680	34	5	RR	0.93	0.50：1.71	弱（C）	7	
有害事象による治療中断率	RCT/2	0	−1	−1	0	0	0	106	1	0.943	102	11	10.784	RR	6.14	0.00：560641.88	弱（C）	7	

反応率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Xiao 2020	50	68	43	68	1.16	[0.92; 1.46]
Kato 2018	275	558	233	551	1.17	[1.02; 1.33]
Ferreri 2001	16	34	14	38	1.28	[0.74; 2.21]
Nakajima 2011	15	20	4	21	3.94	[1.57; 9.85]
Random effects model		680		678	1.19	[1.07; 1.33]

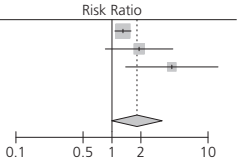
Heterogeneity: I^2=55.5%, τ^2<0.0001, p=0.0807



寛解率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Kato 2018	174	558	132	551	1.30	[1.07; 1.58]
Ferreri 2001	12	34	7	38	1.92	[0.85; 4.30]
Nakajima 2011	12	20	3	21	4.20	[1.39; 12.71]
Random effects model		612		610	1.82	[0.99; 3.33]

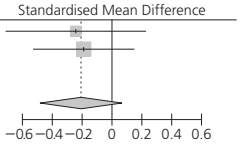
Heterogeneity: I^2=58.9%, τ^2=0.1712, p=0.0879



抑うつ症状の軽減

Study	Experimental Total	Control Total	SMD	95% CI
Ferreri 2001	33	38	−0.24	[−0.71; 0.23]
Xiao 2020	68	68	−0.19	[−0.53; 0.15]
Random effects model	101	106	−0.21	[−0.48; 0.07]

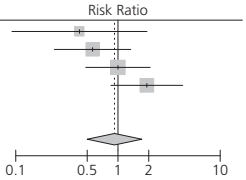
Heterogeneity: I^2=0.0%, τ^2=0, p=0.8573



すべての理由による治療中断率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Nakajima 2011	2	20	5	21	0.42	[0.09; 1.92]
Kato 2018	8	558	14	551	0.56	[0.24; 1.33]
Xiao 2020	12	68	12	68	1.00	[0.48; 2.07]
Ferreri 2001	12	34	7	38	1.92	[0.85; 4.30]
Random effects model		680		678	0.93	[0.50; 1.71]

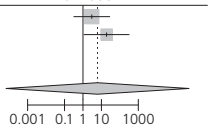
Heterogeneity: I^2=44.5%, τ^2=0.1740, p=0.1441



有害事象による治療中断率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Xiao 2020	3	68	1	68	3.00	[0.32; 28.13]
Ferreri 2001	8	34	0	38	18.97	[1.14; 316.66]
Random effects model		102		106	6.14	[0.00; 560641.88]

Heterogeneity: I^2=1.0%, τ^2=0.0174, p=0.3148



	判断						
問題	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
望ましい効果	わずか	小	中	大		さまざま	わからない
望ましくない効果	大	中	小	わずか		さまざま	わからない
エビデンスの確実性	非常に低	低	中	高			わからない
価値観	重要な不確実性，ばらつきあり	重要な不確実性，ばらつきの可能性あり	重要な不確実性，ばらつきはおそらくなし	重要な不確実性，ばらつきなし			
効果のバランス	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	わからない
必要資源量	大きなコスト	中等度のコスト	無視できる程度のコスト	中等度の節減	大きな節減	さまざま	わからない
必要資源のエビデンスの確実性	非常に低	低	中	高			採用研究なし
費用対効果	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	採用研究なし
公平性	減る	おそらく減る	おそらく影響なし	おそらく増える	増える	さまざま	わからない
容認性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
実行可能性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない

CQ 9-5

【評価シート　エビデンス総体】

診療ガイドライン	日本うつ病学会　うつ病診療ガイドライン 2025
対象	初期治療で十分な効果が得られなかったうつ病
介入	別の抗うつ薬を追加
対照	現在の抗うつ薬継続

*1 各ドメインは“高（−2）”“中/疑い（−1）”“低（0）”の 3 段階
*2 上昇要因は“高（+2）”“中（+1）”“低（0）”の 3 段階
*3 エビデンスの強さは“強（A）”“中（B）”“弱（C）”“非常に弱（D）”の 4 段階．エビデンスの強さについて，RCT は“強（A）”からスタート，観察研究は“弱（C）”からスタート
*4 重要性はアウトカムの重要性（1〜9）

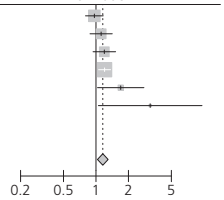
エビデンス総体

アウトカム	研究デザイン/研究数	バイアスリスク*1	非一貫性*1	不精確性*1	非直接性*1	その他（出版バイアスなど）*1	上昇要因（観察研究）*2	リスク人数（アウトカム率）						効果指標（種類）	効果指標統合値	95%信頼区間	エビデンスの強さ*3	重要性*4	コメント
								対照群分母	対照群分子	（%）	介入群分母	介入群分子	（%）						
反応率	RCT/6	−1	0	0	−1	−1	0	1010	440	43.6	987	508	51.5	RR	1.16	1.03：1.30	弱（C）	9	
寛解率	RCT/6	−1	0	−1	−1	−1	0	1010	263	26.0	987	347	35.2	RR	1.32	1.12：1.54	弱（C）	7	
抑うつ症状の軽減	RCT/2	−1	0	0	−1	−1	0	83	NA	NA	79	NA	NA	SMD	−0.52	−1.53：0.48	弱（C）	7	
すべての理由による治療中断率	RCT/6	0	−1	−1	−1	0	0	1010	67	6.6	987	70	7.1	RR	1.08	0.78：1.48	弱（C）	7	
有害事象による治療中断率	RCT/3	0	0	−1	−1	0	0	121	2	1.7	111	5	4.5	RR	2.34	0.08：66.56	弱（C）	7	

反応率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Licht 2002	66	98	69	99	0.97	[0.80; 1.17]
Xiao 2020	48	68	43	68	1.12	[0.88; 1.42]
Kessler 2018a/2018b	94	241	78	239	1.20	[0.94; 1.52]
Kato 2018	273	537	233	551	1.20	[1.06; 1.37]
Ferreri 2001	20	32	14	38	1.70	[1.03; 2.78]
Carpenter 2002	7	11	3	15	3.18	[1.05; 9.62]
Random effects model		987		1010	1.16	[1.03; 1.30]

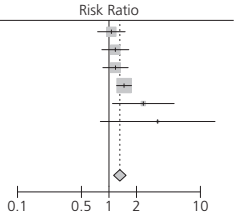
Heterogeneity: I^2=47.2%, τ^2=0.0050, p=0.0915



寛解率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Xiao 2020	34	68	32	68	1.06	[0.75; 1.50]
Licht 2002	43	98	37	99	1.17	[0.84; 1.65]
Kessler 2018a/2018b	63	241	53	239	1.18	[0.86; 1.62]
Kato 2018	188	537	132	551	1.46	[1.21; 1.76]
Ferreri 2001	14	32	7	38	2.38	[1.09; 5.16]
Carpenter 2002	5	11	2	15	3.41	[0.80; 14.44]
Random effects model		987		1010	1.32	[1.12; 1.54]

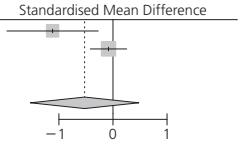
Heterogeneity: I^2=32.7%, τ^2=0.0064, p=0.1909



抑うつ症状の軽減

Study	Experimental Total	Control Total	SMD	95% CI
Carpenter 2002	11	15	−1.12	[−1.96; −0.27]
Xiao 2020	68	68	−0.08	[−0.42; 0.25]
Random effects model	79	83	−0.52	[−1.53; 0.48]

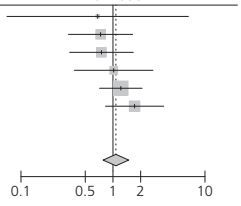
Heterogeneity: I^2=79.9%, τ^2=0.4273, p=0.0257



すべての理由による治療中断率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Carpenter 2002	1	11	2	15	0.68	[0.07; 6.61]
Kato 2018	10	537	14	551	0.73	[0.33; 1.64]
Xiao 2020	9	68	12	68	0.75	[0.34; 1.66]
Ferreri 2001	6	32	7	38	1.02	[0.38; 2.72]
Kessler 2018a/2018b	27	241	22	239	1.22	[0.71; 2.08]
Licht 2002	17	98	10	99	1.72	[0.83; 3.56]
Random effects model		987		1010	1.08	[0.78; 1.48]

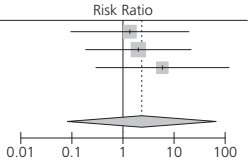
Heterogeneity: I^2=0.0%, τ^2=0, p=0.6056



有害事象による治療中断率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Carpenter 2002	1	11	1	15	1.36	[0.10; 19.50]
Xiao 2020	2	68	1	68	2.00	[0.19; 21.54]
Ferreri 2001	2	32	0	38	5.92	[0.29; 119.00]
Random effects model		111		121	2.34	[0.08; 66.56]

Heterogeneity: I^2=0.0%, τ^2=0, p=0.7622



	判断						
問題	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
望ましい効果	わずか	小	中	大		さまざま	わからない
望ましくない効果	大	中	小	わずか		さまざま	わからない
エビデンスの確実性	非常に低	低	中	高			わからない
価値観	重要な不確実性，ばらつきあり	重要な不確実性，ばらつきの可能性あり	重要な不確実性，ばらつきはおそらくなし	重要な不確実性，ばらつきなし			
効果のバランス	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	わからない
必要資源量	大きなコスト	中等度のコスト	無視できる程度のコスト	中等度の節減	大きな節減	さまざま	わからない
必要資源のエビデンスの確実性	非常に低	低	中	高			採用研究なし
費用対効果	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	採用研究なし
公平性	減る	おそらく減る	おそらく影響なし	おそらく増える	増える	さまざま	わからない
容認性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
実行可能性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない

CQ 9-6

【評価シート　エビデンス総体】

診療ガイドライン	日本うつ病学会　うつ病診療ガイドライン 2025
対象	初期治療で十分な効果が得られなかったうつ病
介入	抗精神病薬を追加
対照	現在の抗うつ薬継続

*1 各ドメインは“高（−2）”“中/疑い（−1）”“低（0）”の 3 段階
*2 上昇要因は“高（+2）”“中（+1）”“低（0）”の 3 段階
*3 エビデンスの強さは“強（A）”“中（B）”“弱（C）”“非常に弱（D）”の 4 段階．エビデンスの強さについて，RCT は“強（A）”からスタート，観察研究は“弱（C）”からスタート
*4 重要性はアウトカムの重要性（1〜9）

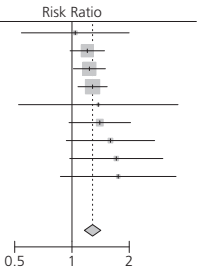
エビデンス総体

アウトカム	研究デザイン/研究数	バイアスリスク*1	非一貫性*1	不精確性*1	非直接性*1	その他（出版バイアスなど）*1	上昇要因（観察研究）*2	リスク人数（アウトカム率）						効果指標（種類）	効果指標統合値	95%信頼区間	エビデンスの強さ*3	重要性*4	コメント
								対照群分母	対照群分子	（%）	介入群分母	介入群分子	（%）						
反応率	RCT/9	−1	0	0	−1	0	0	1023	340	33.2	1551	748	48.2	RR	1.28	1.16：1.42	弱（C）	9	
寛解率	RCT/9	−1	0	0	−1	−1	0	1102	264	24.0	1630	583	35.8	RR	1.34	1.16：1.55	弱（C）	7	
抑うつ症状の軽減	RCT/5	−1	0	0	−1	0	0	546	NA	NA	717	NA	NA	SMD	−0.57	−1.13：−0.00	弱（C）	7	
すべての理由による治療中断率	RCT/10	0	−1	−1	−1	0	0	1113	160	14.4	1642	280	17.1	RR	1.16	0.91：1.48	弱（C）	7	
有害事象による治療中断率	RCT/8	0	0	0	−1	0	0	933	22	2.4	1574	161	10.2	RR	3.42	2.19：5.35	弱（C）	7	

反応率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Fava 2012/Mischoulon 2012	10	56	29	169	1.04	[0.54; 2.00]
El-Khalili 2010	160	298	66	148	1.20	[0.98; 1.48]
Bauer 2009	185	330	74	163	1.23	[1.02; 1.50]
Durgam 2016	265	550	101	269	1.28	[1.08; 1.53]
Reeves 2008	6	12	4	11	1.38	[0.52; 3.61]
Mahmoud 2007	49	141	33	133	1.40	[0.97; 2.03]
Keitner 2009	34	64	11	33	1.59	[0.93; 2.72]
Papakostas 2015	25	71	14	68	1.71	[0.97; 3.00]
McIntyre 2007	14	29	8	29	1.75	[0.87; 3.52]
Random effects model		1551		1023	1.28	[1.16; 1.42]

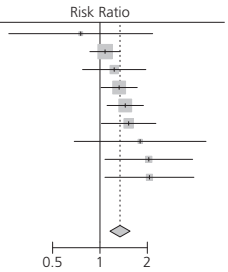
Heterogeneity: I^2=0.0%, τ^2=0, p=0.8973



寛解率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Fava 2012/Mischoulon 2012	4	56	16	169	0.75	[0.26; 2.16]
Durgam 2016	174	550	79	269	1.08	[0.86; 1.35]
Papakostas 2015	27	71	21	68	1.23	[0.77; 1.96]
Bauer 2009	134	330	50	163	1.32	[1.02; 1.73]
El-Khalili 2010	137	298	47	148	1.45	[1.11; 1.89]
Lenze 2015	40	91	26	90	1.52	[1.02; 2.27]
McIntyre 2007	9	29	5	29	1.80	[0.69; 4.72]
Mahmoud 2007	26	141	12	133	2.04	[1.08; 3.88]
Keitner 2009	32	64	8	33	2.06	[1.08; 3.95]
Random effects model		1630		1102	1.34	[1.16; 1.55]

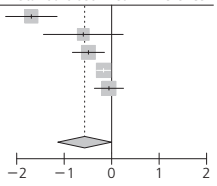
Heterogeneity: I^2=14.1%, τ^2=0.0091, p=0.3167



抑うつ症状の軽減

Study	Experimental Total	Control Total	SMD	95% CI
Moica 2018	36	36	−1.69	[−2.23; −1.15]
Reeves 2008	12	11	−0.60	[−1.44; 0.24]
Papakostas 2015	71	68	−0.49	[−0.83; −0.15]
Durgam 2016	544	264	−0.17	[−0.32; −0.02]
Fava 2012/Mischoulon 2012	54	167	−0.06	[−0.36; 0.25]
Random effects model	717	546	−0.57	[−1.13; −0.00]

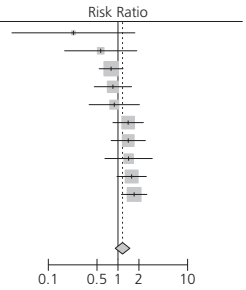
Heterogeneity: I^2=87.6%, τ^2=0.3617, p<0.0001



すべての理由による治療中断率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Reeves 2008	1	12	4	11	0.23	[0.03; 1.75]
Lenze 2015	4	91	7	90	0.57	[0.17; 1.86]
Durgam 2016	57	550	35	269	0.80	[0.54; 1.18]
McIntyre 2007	11	29	13	29	0.85	[0.46; 1.57]
Keitner 2009	12	64	7	33	0.88	[0.38; 2.03]
Bauer 2009	51	330	18	163	1.40	[0.85; 2.32]
Papakostas 2015	22	71	15	68	1.40	[0.80; 2.47]
Fava 2012/Mischoulon 2012	8	56	17	169	1.42	[0.65; 3.11]
Mahmoud 2007	35	141	21	133	1.57	[0.97; 2.56]
El-Khalili 2010	79	298	23	148	1.71	[1.12; 2.60]
Random effects model		1642		1113	1.16	[0.91; 1.48]

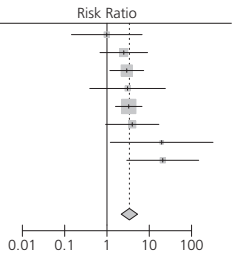
Heterogeneity: I^2=38.6%, τ^2=0.0494, p=0.1009



有害事象による治療中断率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Lenze 2015	2	91	2	90	0.99	[0.14; 6.87]
Mahmoud 2007	8	141	3	133	2.52	[0.68; 9.28]
Bauer 2009	30	330	5	163	2.96	[1.17; 7.50]
Keitner 2009	6	64	1	33	3.09	[0.39; 24.64]
Durgam 2016	54	550	8	269	3.30	[1.59; 6.84]
McIntyre 2007	8	29	2	29	4.00	[0.93; 17.25]
Papakostas 2015	10	71	0	68	20.12	[1.20; 336.73]
El-Khalili 2010	43	298	1	148	21.36	[2.97; 153.56]
Random effects model		1574		933	3.42	[2.19; 5.35]

Heterogeneity: I^2=0.0%, τ^2<0.0001, p=0.4531



	判断						
問題	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
望ましい効果	わずか	小	中	大		さまざま	わからない
望ましくない効果	大	中	小	わずか		さまざま	わからない
エビデンスの確実性	非常に低	低	中	高			わからない
価値観	重要な不確実性，ばらつきあり	重要な不確実性，ばらつきの可能性あり	重要な不確実性，ばらつきはおそらくなし	重要な不確実性，ばらつきなし			
効果のバランス	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	わからない
必要資源量	大きなコスト	中等度のコスト	無視できる程度のコスト	中等度の節減	大きな節減	さまざま	わからない
必要資源のエビデンスの確実性	非常に低	低	中	高			採用研究なし
費用対効果	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	採用研究なし
公平性	減る	おそらく減る	おそらく影響なし	おそらく増える	増える	さまざま	わからない
容認性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
実行可能性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない

CQ 9-7

【評価シート　エビデンス総体】

診療ガイドライン	日本うつ病学会　うつ病診療ガイドライン 2025
対象	初期治療で十分な効果が得られなかったうつ病
介入	別の抗うつ薬や抗精神病薬以外の薬剤を追加
対照	現在の抗うつ薬継続

*1 各ドメインは“高（−2）”“中/疑い（−1）”“低（0）”の 3 段階
*2 上昇要因は“高（+2）”“中（+1）”“低（0）”の 3 段階
*3 エビデンスの強さは“強（A）”“中（B）”“弱（C）”“非常に弱（D）”の 4 段階．エビデンスの強さについて，RCT は“強（A）”からスタート，観察研究は“弱（C）”からスタート
*4 重要性はアウトカムの重要性（1〜9）

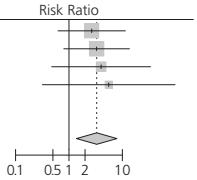
エビデンス総体

アウトカム	研究デザイン/研究数	バイアスリスク*1	非一貫性*1	不精確性*1	非直接性*1	その他（出版バイアスなど）*1	上昇要因（観察研究）*2	リスク人数（アウトカム率）						効果指標（種類）	効果指標統合値	95%信頼区間	エビデンスの強さ*3	重要性*4	コメント
								対照群分母	対照群分子	（%）	介入群分母	介入群分子	（%）						
寛解率	RCT/4	−1	0	−1	−1	0	0	89	5	5.6	119	22	18.5	RR	3.31	1.38：7.91	弱（C）	7	
抑うつ症状の軽減	RCT/8	−1	0	0	−1	−1	0	200	NA	NA	218	NA	NA	SMD	−0.89	−1.52：−0.27	弱（C）	7	
反応率	RCT/7	−1	0	−1	−1	0	0	221	60	27.1	280	97	34.6	RR	1.29	0.99：1.69	弱（C）	9	
すべての理由による治療中断率	RCT/6	0	0	−1	−1	0	0	170	20	11.8	233	37	15.9	RR	1.16	0.68：1.99	弱（C）	7	
有害事象による治療中断率	RCT/5	0	0	−1	−1	0	0	163	7	4.3	227	15	6.6	RR	1.05	0.25：4.40	弱（C）	7	

寛解率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Joffe 1993	6	18	2	16	2.67	[0.62; 11.39]
Joffe 1993	7	17	2	16	3.29	[0.80; 13.57]
Patkar 2006	4	30	1	30	4.00	[0.47; 33.73]
Mozaffari-Khosravi 2013	5	54	0	27	5.55	[0.32; 96.78]
Random effects model		119		89	3.31	[1.38; 7.91]

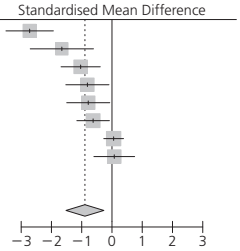
Heterogeneity: I^2=0.0%, τ^2=0, p=0.9707



抑うつ症状の軽減

Study	Experimental Total	Control Total	SMD	95% CI
Jahangard 2018	25	25	−2.72	[−3.50; −1.94]
Nemets 2002	10	10	−1.66	[−2.70; −0.61]
Mowla 2011	20	22	−1.03	[−1.68; −0.39]
Joffe 1993	17	16	−0.81	[−1.53; −0.10]
Joffe 1993	17	16	−0.78	[−1.50; −0.07]
Mozaffari-Khosravi 2013	41	21	−0.62	[−1.16; −0.09]
Ravindran 2008a	72	72	0.06	[−0.27; 0.38]
Stein 1993	16	18	0.08	[−0.60; 0.75]
Random effects model	218	200	−0.89	[−1.52; −0.27]

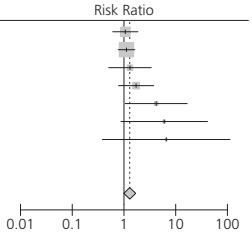
Heterogeneity: I^2=86.9%, τ^2=0.6911, p<0.0001



反応率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Appelberg 2001	17	51	16	51	1.06	[0.61; 1.86]
Ravindran 2008a	34	73	30	72	1.12	[0.77; 1.61]
Peet 2002	16	52	4	17	1.31	[0.51; 3.38]
Patkar 2006	12	30	7	30	1.71	[0.78; 3.75]
Baumann 1996	6	10	2	14	4.20	[1.06; 16.68]
Nemets 2002	6	10	1	10	6.00	[0.87; 41.21]
Mozaffari-Khosravi 2013	6	54	0	27	6.56	[0.38; 112.24]
Random effects model		280		221	1.29	[0.99; 1.69]

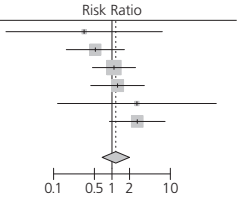
Heterogeneity: I^2=25.6%, τ^2<0.0001, p=0.2332



すべての理由による治療中断率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Nemets 2002	0	10	1	10	0.33	[0.02; 7.28]
Peet 2002	6	52	4	18	0.52	[0.17; 1.63]
Mozaffari-Khosravi 2013	13	54	6	27	1.08	[0.46; 2.53]
Mowla 2011	6	26	5	27	1.25	[0.43; 3.59]
Joffe 1993	1	18	0	16	2.68	[0.12; 61.27]
Ravindran 2008a	11	73	4	72	2.71	[0.91; 8.12]
Random effects model		233		170	1.16	[0.68; 1.99]

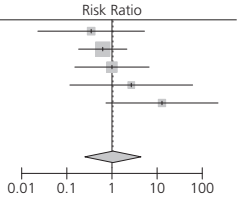
Heterogeneity: I^2=2.7%, τ^2=0.0501, p=0.3994



有害事象による治療中断率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Peet 2002	1	52	1	18	0.35	[0.02; 5.25]
Mozaffari-Khosravi 2013	5	54	4	27	0.62	[0.18; 2.14]
Patkar 2006	2	30	2	30	1.00	[0.15; 6.64]
Joffe 1993	1	18	0	16	2.68	[0.12; 61.27]
Ravindran 2008a	6	73	0	72	12.82	[0.74; 223.48]
Random effects model		227		163	1.05	[0.25; 4.40]

Heterogeneity: I^2=12.7%, τ^2=0.2194, p=0.3331



	判断						
問題	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
望ましい効果	わずか	小	中	大		さまざま	わからない
望ましくない効果	大	中	小	わずか		さまざま	わからない
エビデンスの確実性	非常に低	低	中	高			わからない
価値観	重要な不確実性，ばらつきあり	重要な不確実性，ばらつきの可能性あり	重要な不確実性，ばらつきはおそらくなし	重要な不確実性，ばらつきなし			
効果のバランス	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	わからない
必要資源量	大きなコスト	中等度のコスト	無視できる程度のコスト	中等度の節減	大きな節減	さまざま	わからない
必要資源のエビデンスの確実性	非常に低	低	中	高			採用研究なし
費用対効果	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	採用研究なし
公平性	減る	おそらく減る	おそらく影響なし	おそらく増える	増える	さまざま	わからない
容認性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
実行可能性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない

CQ 9-8

【評価シート　エビデンス総体】

診療ガイドライン	日本うつ病学会　うつ病診療ガイドライン 2025
対象	初期治療で十分な効果が得られなかったうつ病
介入	精神療法の追加
対照	現在の抗うつ薬継続

*1 各ドメインは“高（−2）”“中/疑い（−1）”“低（0）”の 3 段階
*2 上昇要因は“高（+2）”“中（+1）”“低（0）”の 3 段階
*3 エビデンスの強さは“強（A）”“中（B）”“弱（C）”“非常に弱（D）”の 4 段階．エビデンスの強さについて，RCT は“強（A）”からスタート，観察研究は“弱（C）”からスタート
*4 重要性はアウトカムの重要性（1〜9）

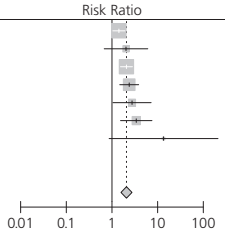
エビデンス総体

アウトカム	研究デザイン/研究数	バイアスリスク*1	非一貫性*1	不精確性*1	非直接性*1	その他（出版バイアスなど）*1	上昇要因（観察研究）*2	リスク人数（アウトカム率）						効果指標（種類）	効果指標統合値	95% 信頼区間	エビデンスの強さ*3	重要性*4	コメント
								対照群分母	対照群分子	（%）	介入群分母	介入群分子	（%）						
反応率	RCT/7	−1	0	−1	0	−1	0	415	103	24.82	411	212	51.58	RR	2.08	1.57：2.75	弱（C）	9	
寛解率	RCT/12	−1	0	0	0	−1	0	782	173	22.12	983	330	33.57	RR	1.59	1.23：2.05	弱（C）	7	
抑うつ症状の軽減	RCT/17	−1	0	−1	0	0	0	493	NA	NA	506	NA	NA	SMD	−1.06	−1.37：0.76	弱（C）	7	
すべての理由による治療中断率	RCT/18	0	−1	−1	0	0	0	909	130	14.30	1129	156	13.81	RR	0.99	0.79：1.22	弱（C）	7	
有害事象による治療中断率	RCT/2	0	−1	−1	0	0	0	192	4	2.083	395	3	0.759	RR	0.37	0.00：6475.17	弱（C）	7	

反応率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Schramm 2007	44	65	31	65	1.42	[1.05; 1.93]
Souza 2016	6	17	4	23	2.03	[0.68; 6.09]
Wiles 2013/2016	95	234	46	235	2.07	[1.53; 2.81]
Nakagawa 2017	31	40	13	40	2.38	[1.48; 3.84]
Nakao 2018	11	20	4	20	2.75	[1.05; 7.20]
Watkins 2011a	17	21	5	21	3.40	[1.54; 7.51]
Wiles 2008	8	14	0	11	13.48	[0.87; 209.89]
Random effects model		411		415	2.08	[1.57; 2.75]

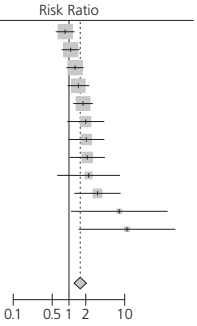
Heterogeneity: I^2=35.4%, τ^2=0.0458, p=0.1580



寛解率

Study	Experimental Events	Experimental Total	Control Events	Control Total	RR	95% CI
Kocsis 2009/Klein 2011	52	195	30	96	0.85	[0.59; 1.24]
Kocsis 2009/Klein 2011	67	200	30	96	1.07	[0.75; 1.53]
Reynolds 2010	35	60	29	64	1.29	[0.91; 1.82]
Schramm 2007	31	65	21	65	1.48	[0.96; 2.28]
Wiles 2013/2016	57	234	32	235	1.79	[1.21; 2.65]
Lynch 2007_study 2	15	21	5	14	2.00	[0.94; 4.25]
Paykel 1999/Scott 2000	19	80	9	78	2.06	[0.99; 4.27]
Nakagawa 2017	17	40	8	40	2.12	[1.04; 4.35]
Souza 2016	5	17	3	23	2.25	[0.62; 8.17]
Watkins 2011a	13	21	4	21	3.25	[1.27; 8.35]
Nakao 2018	8	20	1	20	8.00	[1.10; 58.19]
Town 2017/2020	11	30	1	30	11.00	[1.51; 79.96]
Random effects model		983		782	1.59	[1.23; 2.05]

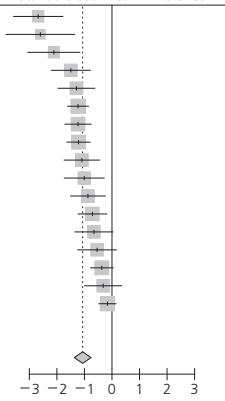
Heterogeneity: I^2=55.1%, τ^2=0.0908, p=0.0108



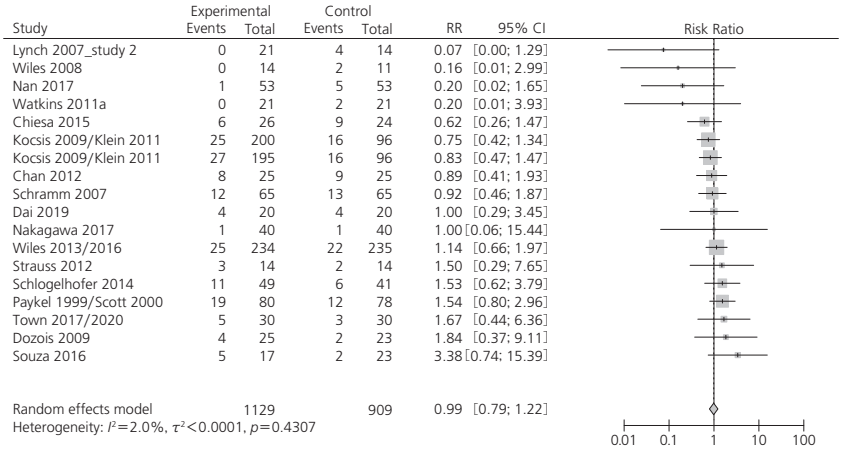
抑うつ症状の軽減

Study	Experimental Total	Control Total	SMD	95% CI
Embling 2002	19	19	−2.68	[−3.58; −1.78]
Dunn 1979	10	10	−2.60	[−3.85; −1.35]
Strauss 2012	14	14	−2.11	[−3.06; −1.16]
Nakao 2018	20	20	−1.49	[−2.20; −0.78]
Watkins 2011a	21	21	−1.29	[−1.96; −0.62]
Schramm 2007	63	61	−1.23	[−1.62; −0.85]
Nakagawa 2017	40	40	−1.23	[−1.71; −0.75]
Nan 2017	52	48	−1.22	[−1.64; −0.79]
Chiesa 2015	23	20	−1.09	[−1.74; −0.45]
Chan 2012	17	16	−1.01	[−1.74; −0.28]
Dozois 2009	21	21	−0.87	[−1.51; −0.24]
Town 2017/2020	30	30	−0.71	[−1.23; −0.19]
Souza 2016	16	18	−0.66	[−1.35; 0.04]
Dai 2019	16	16	−0.54	[−1.25; 0.17]
Schlogelhofer 2014	49	41	−0.37	[−0.79; 0.05]
Baert 2010_study 2	15	20	−0.32	[−0.99; 0.36]
Paykel 1999/Scott 2000	80	78	−0.16	[−0.48; 0.15]
Random effects model	506	493	−1.06	[−1.37; −0.76]

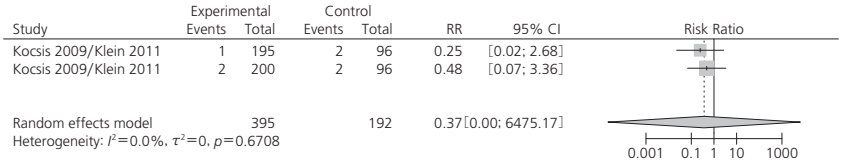
Heterogeneity: I^2=77.7%, τ^2=0.2990, p<0.0001



すべての理由による治療中断率



有害事象による治療中断率



	判断						
問題	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
望ましい効果	わずか	小	中	大		さまざま	わからない
望ましくない効果	大	中	小	わずか		さまざま	わからない
エビデンスの確実性	非常に低	低	中	高			わからない
価値観	重要な不確実性，ばらつきあり	重要な不確実性，ばらつきの可能性あり	重要な不確実性，ばらつきはおそらくなし	重要な不確実性，ばらつきなし			
効果のバランス	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	わからない
必要資源量	大きなコスト	中等度のコスト	無視できる程度のコスト	中等度の節減	大きな節減	さまざま	わからない
必要資源のエビデンスの確実性	非常に低	低	中	高			採用研究なし
費用対効果	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	採用研究なし
公平性	減る	おそらく減る	おそらく影響なし	おそらく増える	増える	さまざま	わからない
容認性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
実行可能性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない

CQ 10-3

【評価シート　エビデンス総体】

診療ガイドライン	日本うつ病学会　うつ病診療ガイドライン 2025
対象	2 剤以上の抗うつ薬治療に不応なうつ病外来患者
介入	抗精神病薬の上乗せ
対照	プラセボの追加

*1 各ドメインは“高（−2）”“中/疑い（−1）”“低（0）”の 3 段階
*2 上昇要因は“高（+2）”“中（+1）”“低（0）”の 3 段階
*3 エビデンスの強さは“強（A）”“中（B）”“弱（C）”“非常に弱（D）”の 4 段階．エビデンスの強さについて，RCT は“強（A）”からスタート，観察研究は“弱（C）”からスタート
*4 重要性はアウトカムの重要性（1〜9）

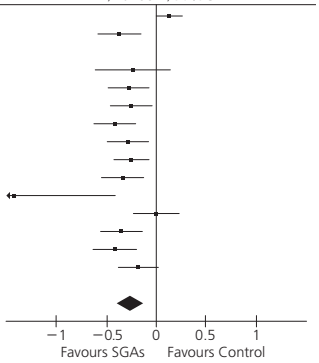
エビデンス総体

アウトカム	研究デザイン/研究数	バイアスリスク*1	非一貫性*1	不精確性*1	非直接性*1	その他（出版バイアスなど）*1	上昇要因（観察研究）*2	リスク人数（アウトカム率）						効果指標（種類）	効果指標統合値	95% 信頼区間	エビデンスの強さ*3	重要性*4	コメント
								対照群分母	対照群分子	（%）	介入群分母	介入群分子	（%）						
抑うつ症状の軽減	RCT/15	−1	−1	0	0	−1			2653			2653		SMD	−0.25	−0.36；−0.14	中（B）	8	
反応率	RCT/13	−1	0	0	0	−1		480	2196	21.9	740	2374	31.2	RR	1.36	1.24；1.51	中（B）	9	ARR＝0.09，NNT＝11
寛解率	RCT/13	−1	−1	0	0	−1		455	2627	17.3	643	2808	22.9	RR	1.35	1.15；1.59	中（B）	9	ARR＝0.06，NNT＝18
少なくとも 1 つの有害事象の発現頻度	RCT/13	−1	−1	0	0	−1		1357	2676	50.7	1705	2868	59.4	RR	1.23	1.11；1.35	中（B）	8	ARR＝0.09，NNH＝11
有害事象による治療中断率	RCT/13	−1	−1	0	0	−1		50	2675	1.9	132	2866	4.6	RR	2.06	1.41；3.01	中（B）	7	ARR＝0.03，NNH＝36
すべての理由による治療中断率	RCT/13	−1	−1	0	0	−1		263	2285	11.5	373	2469	15.1	RR	1.23	1.03；1.46	中（B）	7	ARR＝0.04，NNH＝28

抑うつ症状の軽減

Study or Subgroup	SGAs Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95%CI
Bauer 2019	−11.4	8.52	444	−12.6	8.26	441	9.0%	0.14[0.01, 0.27]
Berman 2007	−8.8	8.52	181	−5.8	8.26	172	7.6%	−0.36[−0.57, −0.15]
Chaput 2008	−6	0	10	−1.8	0	11		Not estimable
Corya 2006	−14.1	8.52	55	−12.2	8.26	60	4.9%	−0.23[−0.59, −0.14]
Hobart 2018a	−10.4	8.3	192	−8.1	8.5	202	7.8%	−0.27[−0.47, −0.07]
Hobart 2018b	−6	5.5	191	−4.6	5.7	206	7.8%	−0.25[−0.45, −0.05]
Kamijima 2013	−7.6	5.6	197	−5.3	5.6	195	7.8%	−0.41[−0.61, −0.21]
Kamijima 2018	−9.2	7.2	208	−7.2	7.1	203	7.9%	−0.28[−0.47, −0.08]
Kato 2023	−8.5	7.4	248	−6.7	7.3	243	8.2%	−0.24[−0.42, −0.07]
Marcus 2008	−8.5	8.52	185	−5.7	8.26	184	7.6%	−0.33[−0.54, −0.13]
Shelton 2001	−13.6	8.52	10	−1.2	8.26	10	1.1%	−1.42[−2.42, −0.41]
Shelton 2005	−8.71	8.46	146	−8.81	8.34	142	7.2%	0.01[−0.22, 0.24]
Thase 2007	−12.6	10.3	198	−9.2	9.7	203	7.8%	−0.34[−0.54, −0.14]
Thase 2015a	−8.36	7.8	175	−5.15	7.8	178	7.5%	−0.41[−0.62, −0.20]
Thase 2015b	−8.29	10.9	213	−6.33	10.7	203	7.9%	−0.18[−0.37, 0.01]
Total (95%CI)			2653			2653	100.0%	−0.25[−0.36, −0.14]

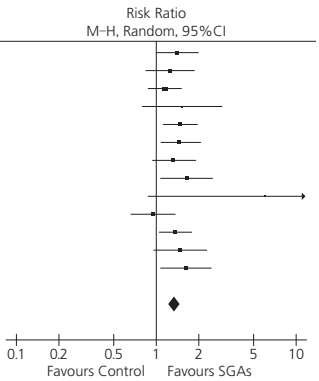
Heterogeneity: Tau²=0.03; Chi²=50.03, df=13 (P<0.00001); I²=74%
Test for overall effect: Z=4.43 (P<0.00001)



反応率

Study or Subgroup	SGAs Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95%CI
Berman 2007	61	181	41	172	8.6%	1.41[1.01, 1.98]
Corya 2006	100	231	19	56	6.3%	1.28[0.86, 1.89]
Hobart 2018a	72	191	66	202	13.4%	1.15[0.88, 1.51]
Hobart 2018b	20	191	14	205	2.3%	1.53[0.80, 2.95]
Kamijima 2013	83	197	55	195	12.6%	1.49[1.13, 1.97]
Kamijima 2018	78	208	52	203	11.3%	1.46[1.09, 1.96]
Kato 2023	63	248	46	243	8.6%	1.34[0.96, 1.88]
Marcus 2008	47	185	28	184	5.5%	1.67[1.10, 2.54]
Shelton 2001	6	10	1	10	0.3%	6.00[0.87, 41.21]
Shelton 2005	40	146	41	142	7.1%	0.95[0.66, 1.37]
Thase 2007	80	198	60	203	13.2%	1.37[1.04, 1.79]
Thase 2015a	41	175	28	178	5.2%	1.49[0.97, 2.30]
Thase 2015b	49	213	29	203	5.6%	1.61[1.06, 2.44]
Total (95%CI)		2374		2196	100.0%	1.36[1.24, 1.51]
Total events	740		480			

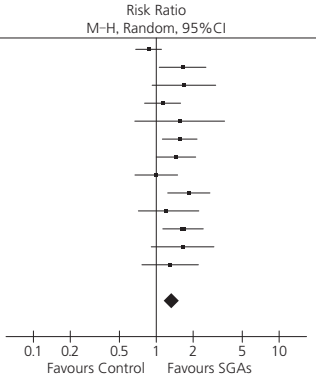
Heterogeneity: Tau²=0.00; Chi²=10.08, df=12 (P=0.61); I²=0%
Test for overall effect: Z=6.18 (P<0.00001)



寛解率

Study or Subgroup	SGAs Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95%CI
Bauer 2019	95	444	110	441	12.3%	0.86[0.67, 1.09]
Berman 2007	47	181	27	172	7.7%	1.65[1.08, 2.53]
Corya 2006	69	231	10	56	5.0%	1.67[0.92, 3.03]
Hobart 2018a	56	191	52	202	10.1%	1.14[0.83, 1.57]
Hobart 2018b	13	191	9	205	3.0%	1.55[0.68, 3.54]
Kamijima 2013	64	197	40	195	9.6%	1.58[1.13, 2.23]
Kamijima 2018	61	208	41	203	9.5%	1.45[1.03, 2.05]
Kato 2023	44	248	43	243	8.6%	1.00[0.68, 1.47]
Marcus 2008	60	185	32	184	8.7%	1.86[1.28, 2.72]
Shelton 2005	24	146	19	142	5.5%	1.23[0.70, 2.14]
Thase 2007	54	198	34	203	8.6%	1.63[1.11, 2.39]
Thase 2015a	26	175	16	178	5.1%	1.65[0.92, 2.97]
Thase 2015b	30	213	22	203	6.1%	1.30[0.78, 2.18]
Total (95%CI)		2808		2627	100.0%	1.35[1.15, 1.59]
Total events	643		455			

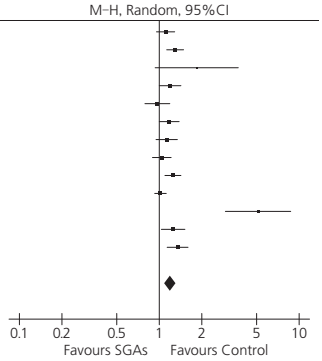
Heterogeneity: Tau²=0.04; Chi²=23.18, df=12 (P=0.03); I²=48%
Test for overall effect: Z=3.17 (P=0.0002)



少なくとも 1 つの有害事象の発現頻度

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95%CI
Bauer 2019	247	444	218	441	9.4%	1.13[0.99, 1.28]
Berman 2007	149	182	110	176	9.2%	1.31[1.15, 1.50]
Corya 2006	61	243	8	60	1.7%	1.88[0.95, 3.72]
Hobart 2018a	115	192	100	202	8.0%	1.21[1.01, 1.45]
Hobart 2018b	100	197	107	206	7.8%	0.98[0.81, 1.18]
Kamijima 2013	141	197	117	195	8.9%	1.19[1.03, 1.38]
Kamijima 2018	128	209	108	203	8.3%	1.15[0.97, 1.36]
Kato 2023	155	250	144	244	9.0%	1.05[0.91, 1.21]
Marcus 2008	154	191	120	190	9.3%	1.28[1.12, 1.45]
Shelton 2005	129	146	119	142	10.1%	1.05[0.96, 1.16]
Thase 2007	70	200	14	206	2.5%	5.15[3.00, 8.84]
Thase 2015a	111	188	89	191	7.7%	1.27[1.04, 1.54]
Thase 2015b	145	229	103	220	8.2%	1.35[1.14, 1.61]
Total (95%CI)		2868		2676	100.0%	1.23[1.11, 1.35]
Total events	1705		1357			

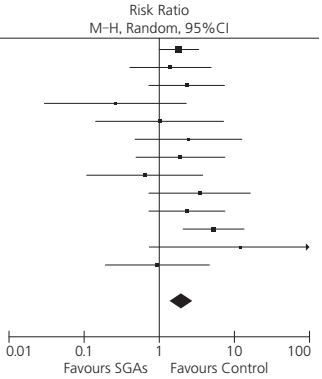
Heterogeneity: $Tau^2=0.02$; $Chi^2=52.53$, df=12 (P<0.00001); $I^2=77\%$
Test for overall effect: Z=4.15 (P<0.0001)



有害事象による治療中断率

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95%CI
Bauer 2019	28	444	15	441	24.8%	1.85[1.00, 3.42]
Berman 2007	6	182	4	176	8.1%	1.45[0.42, 5.05]
Corya 2006	29	243	3	60	9.3%	2.39[0.75, 7.57]
Hobart 2018a	1	192	4	202	2.9%	0.26[0.03, 2.33]
Hobart 2018b	2	197	2	206	3.6%	1.05[0.15, 7.35]
Kamijima 2013	5	197	2	195	5.0%	2.47[0.49, 12.60]
Kamijima 2018	6	209	3	203	6.9%	1.94[0.49, 7.66]
Kato 2023	2	248	3	243	4.2%	0.65[0.11, 3.88]
Marcus 2008	7	191	2	190	5.4%	3.48[0.73, 16.55]
Shelton 2005	10	146	4	142	9.6%	2.43[0.78, 7.57]
Thase 2007	27	200	5	206	13.3%	5.56[2.19, 14.16]
Thase 2015a	6	188	0	191	1.7%	13.21[0.75, 232.79]
Thase 2015b	3	229	3	220	5.2%	0.96[0.20, 4.71]
Total (95%CI)		2866		2675	100.0%	2.06[1.41, 3.01]
Total events	132		50			

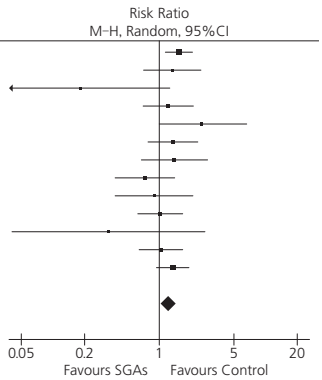
Heterogeneity: $Tau^2=0.05$; $Chi^2=13.46$, df=12 (P=0.34); $I^2=11\%$
Test for overall effect: Z=3.75 (P=0.0002)



すべての理由による治療中断率

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95%CI
Bauer 2019	95	444	61	441	19.6%	1.55[1.15, 2.07]
Berman 2007	22	182	16	176	7.0%	1.33[0.72, 2.45]
Chaput 2008	1	10	6	11	0.8%	0.18[0.03, 1.27]
Corya 2006	60	243	12	60	8.2%	1.23[0.71, 2.14]
Hobart 2018a	15	192	8	202	3.3%	2.63[1.04, 6.64]
Hobart 2018b	26	197	20	206	8.3%	1.36[0.78, 2.35]
Kamijima 2013	17	197	12	195	5.3%	1.40[0.69, 2.86]
Kamijima 2018	15	209	20	203	6.4%	0.73[0.38, 1.38]
Kato 2023	10	248	11	243	4.0%	0.89[0.39, 2.06]
Marcus 2008	29	191	28	190	10.3%	1.03[0.64, 1.66]
Shelton 2001	1	10	3	10	0.7%	0.33[0.04, 2.69]
Shelton 2005	30	146	28	142	10.9%	1.04[0.66, 1.65]
Thase 2007	52	200	40	206	15.2%	1.34[0.93, 1.93]
Total (95%CI)		2469		2285	100.0%	1.23[1.03, 1.46]
Total events	373		263			

Heterogeneity: $Tau^2=0.02$; $Chi^2=14.78$, df=12 (P=0.25); $I^2=19\%$
Test for overall effect: Z=2.28 (P=0.02)



	判断						
問題	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
望ましい効果	わずか	小	中	大		さまざま	わからない
望ましくない効果	大	中	小	わずか		さまざま	わからない
エビデンスの確実性	非常に低	低	中	高			わからない
価値観	重要な不確実性，ばらつきあり	重要な不確実性，ばらつきの可能性あり	重要な不確実性，ばらつきはおそらくなし	重要な不確実性，ばらつきなし			
効果のバランス	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	わからない
必要資源量	大きなコスト	中等度のコスト	無視できる程度のコスト	中等度の節減	大きな節減	さまざま	わからない
必要資源のエビデンスの確実性	非常に低	低	中	高			採用研究なし
費用対効果	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	採用研究なし
公平性	減る	おそらく減る	おそらく影響なし	おそらく増える	増える	さまざま	わからない
容認性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
実行可能性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない

CQ 10-4

【評価シート　エビデンス総体】

診療ガイドライン	日本うつ病学会　うつ病診療ガイドライン 2025
対象	2 剤以上の抗うつ薬治療に不応なうつ病外来患者
介入	気分安定薬の追加
対照	プラセボまたは気分安定薬以外の追加

*1 各ドメインは“高（−2）”“中/疑い（−1）”“低（0）”の 3 段階
*2 上昇要因は“高（+2）”“中（+1）”“低（0）”の 3 段階
*3 エビデンスの強さは“強（A）”“中（B）”“弱（C）”“非常に弱（D）”の 4 段階．エビデンスの強さについて，RCT は“強（A）”からスタート，観察研究は“弱（C）”からスタート
*4 重要性はアウトカムの重要性（1〜9）

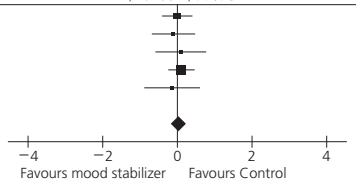
エビデンス総体

アウトカム	研究デザイン/研究数	バイアスリスク*1	非一貫性*1	不精確性*1	非直接性*1	その他（出版バイアスなど）*1	上昇要因（観察研究）*2	リスク人数（平均値、標準偏差）						効果指標（種類）	効果指標統合値	95%信頼区間	エビデンスの強さ*3	重要性*4	コメント
								対照群分母	対照群分子	（%）*2	介入群分母	介入群分子	（%）*2						
抑うつ症状の軽減	RCT/5	−1	0	−2	−1	−1			174			177		SMD	0.02	−0.19；0.23	弱（C）	8	
反応率	RCT/4	−1	0	−2	−1	−1		41	152	27.0	33	150	22.0	RR	0.83	0.56；1.23	弱（C）	9	ARR＝0.05，NNT＝21
寛解率	RCT/1	−1	0	−2	−1	−1		18	73	24.7	11	69	15.94	RR	0.65	0.33；1.27	非常に弱（D）	9	ARR＝0.087，NNT＝12
有害事象による治療中断率	RCT/2	−1	0	−1	−1	−1		7	90	7.8	18	86	20.93	RR	2.55	1.15；5.64	弱（C）	7	ARR＝0.132，NNH＝8
すべての理由による治療中断率	RCT/4	−1	0	−2	−1	−1		24	104	23.1	17	110	15.45	RR	0.73	0.43；1.25	弱（C）	7	ARR＝0.109，NNH＝10
少なくとも 1 つの有害事象の発現頻度	RCT/2	−1	0	−2	−1	−1		36	90	40.0	56	86	65.12	RR	5.19	0.09；307.74	弱（C）	8	ARR＝0.251，NNH＝4

抑うつ症状の軽減

Study or Subgroup	Mood Stabilizer Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95%CI
Barbee 2011	−9.31	11.21	48	−9.37	8.33	48	27.4%	0.01[−0.39, 0.41]
Girlanda 2014	−7	8.2	27	−6.1	5.7	22	13.8%	−0.12[−0.69, 0.44]
Nierenberg 2003	−5.5	5.6	18	−5.8	5.6	17	10.0%	0.05[−0.61, 0.72]
Nierenberg 2006	−1.6	8.2	69	−2.3	5.7	73	40.5%	0.10[−0.23, 0.43]
Santos 2008	−10.5	8.52	15	−9.26	8.26	14	8.3%	−0.14[−0.87, 0.59]
Total (95%CI)			177			174	100.0%	0.02[−0.19, 0.23]

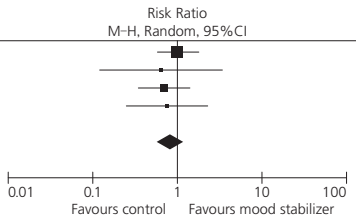
Heterogeneity: Tau²=0.00; Chi²=0.68, df=4 (P=0.95); I²=0%
Test for overall effect: Z=0.17 (P=0.87)



反応率

Study or Subgroup	Mood Stabilizer Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95%CI
Barbee 2011	16	48	16	48	48.3%	1.00[0.57, 1.76]
Nierenberg 2003	2	18	3	17	5.6%	0.63[0.12, 3.32]
Nierenberg 2006	11	69	17	73	33.1%	0.68[0.35, 1.36]
Santos 2008	4	15	5	14	12.9%	0.75[0.25, 2.23]
Total (95%CI)		150		152	100.0%	0.83[0.56, 1.23]
Total events	33		41			

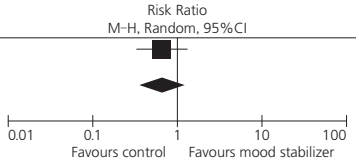
Heterogeneity: Tau²=0.00; Chi²=0.87, df=3 (P=0.83); I²=0%
Test for overall effect: Z=0.94 (P=0.35)



寛解率

Study or Subgroup	Mood Stabilizer Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95%CI
Nierenberg 2006	11	69	18	73	100.0%	0.65[0.33, 1.27]
Total (95%CI)		69		73	100.0%	0.65[0.33, 1.27]
Total events	11		18			

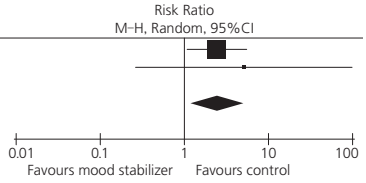
Heterogeneity: Not applicable
Test for overall effect: Z=1.27 (P=0.20)



有害事象による治療中断率

Study or Subgroup	Mood Stabilizer Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95%CI
Nierenberg 2006	16	69	7	73	92.8%	2.42[1.06, 5.52]
Santos 2008	2	17	0	17	7.2%	5.00[0.26, 97.00]
Total (95%CI)		86		90	100.0%	2.55[1.15, 5.64]
Total events	18		7			

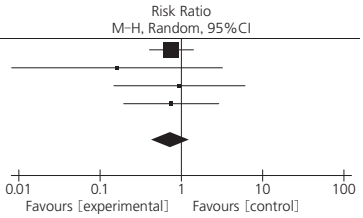
Heterogeneity: Tau²=0.00; Chi²=0.22, df=1 (P=0.64); I²=0%
Test for overall effect: Z=2.31 (P=0.02)



すべての理由による治療中断率

Study or Subgroup	Mood Stabilizer Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95%CI
Barbee 2011	12	48	16	48	72.1%	0.75[0.40, 1.41]
Girlanda 2014	0	27	2	22	3.2%	0.16[0.01, 3.25]
Nierenberg 2003	2	18	2	17	8.5%	0.94[0.15, 5.97]
Santos 2008	3	17	4	17	16.1%	0.75[0.20, 2.86]
Total (95%CI)		110		104	100.0%	0.73[0.43, 1.25]
Total events	17		24			

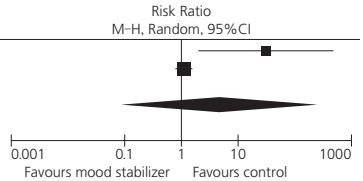
Heterogeneity: $Tau^2=0.00$; $Chi^2=1.06$, df=3 (P=0.79); $I^2=0\%$
Test for overall effect: Z=1.16 (P=0.25)



少なくとも 1 つの有害事象の発現頻度

Study or Subgroup	Mood Stabilizer Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95%CI
Santos 2008	16	17	0	17	44.5%	33.00[2.14, 509.33]
Nierenberg 2006	40	69	36	73	55.5%	1.18[0.86, 1.60]
Total (95%CI)		86		90	100.0%	5.19[0.09, 307.74]
Total events	56		36			

Heterogeneity: $Tau^2=7.80$; $Chi^2=8.90$, df=1 (P=0.003); $I^2=89\%$
Test for overall effect: Z=0.79 (P=0.43)



	判断						
問題	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
望ましい効果	わずか	小	中	大		さまざま	わからない
望ましくない効果	大	中	小	わずか		さまざま	わからない
エビデンスの確実性	非常に低	低	中	高			わからない
価値観	重要な不確実性，ばらつきあり	重要な不確実性，ばらつきの可能性あり	重要な不確実性，ばらつきはおそらくなし	重要な不確実性，ばらつきなし			
効果のバランス	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	わからない
必要資源量	大きなコスト	中等度のコスト	無視できる程度のコスト	中等度の節減	大きな節減	さまざま	わからない
必要資源のエビデンスの確実性	非常に低	低	中	高			採用研究なし
費用対効果	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	採用研究なし
公平性	減る	おそらく減る	おそらく影響なし	おそらく増える	増える	さまざま	わからない
容認性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
実行可能性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない

CQ 10-5

【評価シート　エビデンス総体】

診療ガイドライン	日本うつ病学会　うつ病診療ガイドライン 2025
対象	2 剤以上の抗うつ薬治療に不応なうつ病外来患者
介入	精神療法による付加的治療
対照	通常治療

*1 各ドメインは“高（−2）”“中/疑い（−1）”“低（0）”の 3 段階
*2 上昇要因は“高（+2）”“中（+1）”“低（0）”の 3 段階
*3 エビデンスの強さは“強（A）”“中（B）”“弱（C）”“非常に弱（D）”の 4 段階．エビデンスの強さについて，RCT は“強（A）”からスタート，観察研究は“弱（C）”からスタート
*4 重要性はアウトカムの重要性（1〜9）

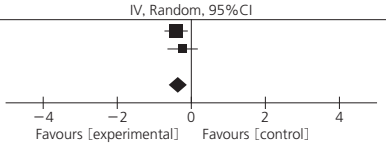
エビデンス総体

アウトカム	研究デザイン/研究数	バイアスリスク*1	非一貫性*1	不精確性*1	非直接性*1	その他（出版バイアスなど）*1	上昇要因（観察研究）*2	リスク人数（アウトカム率）						効果指標（種類）	効果指標統合値	95%信頼区間	エビデンスの強さ*3	重要性*4	コメント
								対照群分母	対照群分子	（%）	介入群分母	介入群分子	（%）						
抑うつ症状の軽減	RCT/2	−1	0	−1	−1	−1			132			140		SMD	−0.35	−0.59：−0.11	弱（C）	8	
反応率	RCT/2	−1	0	−1	−1	−1		24	132	18.18	44	140	31.43	RR	1.69	1.09：2.62	弱（C）	9	ARR＝0.132，NNT＝8
寛解率	RCT/2	−1	0	−2	−1	−1		15	132	11.36	25	140	17.86	RR	1.61	0.89：2.90	弱（C）	9	ARR＝0.065，NNT＝16
すべての理由による治療中断率	RCT/2	−1	0	−2	−1	−1		22	148	14.86	21	154	13.64	RR	0.92	0.53：1.60	弱（C）	7	ARR＝−0.012，NNH＝82
健康 QOL	RCT/1	−1	0	−2	−1	−1			46			53		SMD	0.33	−0.07：0.72	非常に弱（D）	8	
社会機能の改善度	RCT/1	−1	0	−1	−1	−1			46			53		SMD	0.50	0.10：0.90	非常に弱（D）	9	

抑うつ症状の軽減

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95%CI
Eisendrath 2016	11.6	3.4	87	13	3.4	86	63.4%	−0.41[−0.71, −0.11]
Fonagy 2015	16.4	6.2	53	17.9	6.5	46	36.6%	−0.23[−0.63, 0.16]
Total (95%CI)			140			132	100.0%	−0.35[−0.59, −0.11]

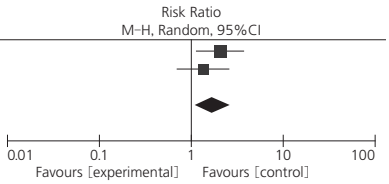
Heterogeneity: Tau²=0.00; Chi²=0.48, df=1 (P=0.49); I²=0%
Test for overall effect: Z=2.83 (P=0.005)



反応率

Study or Subgroup	Experimental Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95%CI
Eisendrath 2016	27	87	13	86	54.6%	2.05[1.14, 3.71]
Fonagy 2015	17	53	11	46	45.4%	1.34[0.70, 2.56]
Total (95%CI)		140		132	100.0%	1.69[1.09, 2.62]
Total events	44		24			

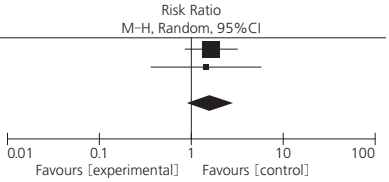
Heterogeneity: Tau²=0.00; Chi²=0.91, df=1 (P=0.34); I²=0%
Test for overall effect: Z=2.36 (P=0.02)



寛解率

Study or Subgroup	Experimental Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95%CI
Eisendrath 2016	20	87	12	86	81.7%	1.65[0.86, 3.16]
Fonagy 2015	5	53	3	46	18.3%	1.45[0.37, 5.73]
Total (95%CI)		140		132	100.0%	1.61[0.89, 2.90]
Total events	25		15			

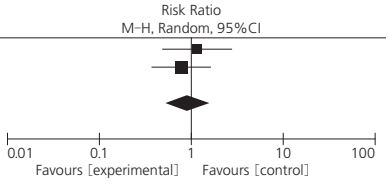
Heterogeneity: Tau²=0.00; Chi²=0.03, df=1 (P=0.87); I²=0%
Test for overall effect: Z=1.58 (P=0.11)



すべての理由による治療中断率

Study or Subgroup	Experimental Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95%CI
Fonagy 2015	10	67	8	62	41.8%	1.16[0.49, 2.74]
Eisendrath 2016	11	87	14	86	58.2%	0.78[0.37, 1.61]
Total (95%CI)		154		148	100.0%	0.92[0.53, 1.60]
Total events	21		22			

Heterogeneity: Tau²=0.00; Chi²=0.48, df=1 (P=0.49); I²=0%
Test for overall effect: Z=0.30 (P=0.76)



	判断						
問題	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
望ましい効果	わずか	小	中	大		さまざま	わからない
望ましくない効果	大	中	小	わずか		さまざま	わからない
エビデンスの確実性	非常に低	低	中	高			わからない
価値観	重要な不確実性，ばらつきあり	重要な不確実性，ばらつきの可能性あり	重要な不確実性，ばらつきはおそらくなし	重要な不確実性，ばらつきなし			
効果のバランス	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	わからない
必要資源量	大きなコスト	中等度のコスト	無視できる程度のコスト	中等度の節減	大きな節減	さまざま	わからない
必要資源のエビデンスの確実性	非常に低	低	中	高			採用研究なし
費用対効果	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	採用研究なし
公平性	減る	おそらく減る	おそらく影響なし	おそらく増える	増える	さまざま	わからない
容認性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
実行可能性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない

CQ 11-5

【評価シート　エビデンス総体】

診療ガイドライン	日本うつ病学会　うつ病診療ガイドライン 2025
対象	寛解したうつ病患者
介入	CBT を基盤とした精神療法の追加
対照	薬物治療

*1 各ドメインは“高（−2）”“中/疑い（−1）”“低（0）”の 3 段階
*2 上昇要因は“高（+2）”“中（+1）”“低（0）”の 3 段階
*3 エビデンスの強さは“強（A）”“中（B）”“弱（C）”“非常に弱（D）”の 4 段階．エビデンスの強さについて，RCT は“強（A）”からスタート，観察研究は“弱（C）”からスタート
*4 重要性はアウトカムの重要性（1〜9）

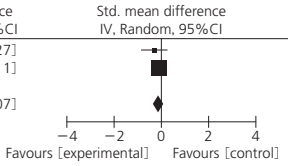
エビデンス総体

アウトカム	研究デザイン/研究数	バイアスリスク*1	非一貫性*1	不精確性*1	非直接性*1	その他（出版バイアスなど）*1	上昇要因（観察研究）*2	リスク人数（アウトカム率）						効果指標（種類）	効果指標統合値	95%信頼区間	エビデンスの強さ*3	重要性*4	コメント
								対照群分母	対照群分子	（%）	介入群分母	介入群分子	（%）						
再発の減少	RCT/17	0	0	0	0	0	0	1075	604	56.19	1208	568	47.02	RR	0.84	0.74：0.94	強（A）	9	NNT＝7
すべての理由による脱落の減少																		7	該当するアウトカムの報告なし
QOL の改善	RCT/2	0	0	−1	0	0	0							SMD	−0.13	−0.33：0.07	中（B）	7	
寛解の増加	RCT/3	0	0	−1	−1	0	0							SMD	−0.38	−0.85：0.09	弱（C）	7	
自殺関連行動の減少	RCT/4																		該当するアウトカムの報告なし
あらゆる有害事象の増加	RCT/3																	8	該当するアウトカムの報告なし
死亡率の増加	RCT/1																	6	該当するアウトカムの報告なし
経済的負担の継続	RCT/2	0	0	−1	−2	0	0	273			273						非常に弱（D）	5	統合できるアウトカムの報告なし

QOL の改善

Study or Subgroup	MBCT Mean	MBCT SD	MBCT Total	AD Mean	AD SD	AD Total	Weight	Std. mean difference IV, Random, 95%CI
Huijbers_2015	3.8	0.8	26	4	0.5	24	12.8%	−0.29[−0.85, 0.27]
Kuyken_2015	3.7	0.9	169	3.8	1	167	87.2%	−0.10[−0.32, 0.11]
Total (95%CI)			195			191	100.0%	−0.13[−0.33, 0.07]

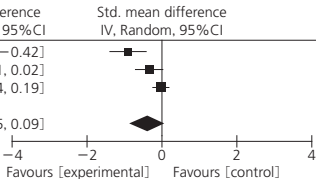
Heterogeneity: Tau²=0.00; Chi²=0.38, df=1 (P=0.54); I²=0%
Test for overall effect: Z=1.26 (P=0.21)
Test for subgroup differences; Not applicable



寛解の増加

Study or Subgroup	MBCT Mean	MBCT SD	MBCT Total	AD Mean	AD SD	AD Total	Weight	Std. mean difference IV, Random, 95%CI
Godfrin and van Heeringen_2010	8.35	10.2	39	19.28	13.72	34	28.8%	−0.90[−1.39, −0.42]
Kuyken_2008	12.61	10.39	59	17.02	14.98	58	33.1%	−0.34[−0.71, 0.02]
Kuyken_2015	11.6	10.9	169	11.9	10.7	167	38.1%	−0.03[−0.24, 0.19]
Total (95%CI)			267			259	100.0%	−0.38[−0.85, 0.09]

Heterogeneity: Tau²=0.14; Chi²=11.18, df=2 (P=0.004); I²=82%
Test for overall effect: Z=1.60 (P=0.11)



	判断						
問題	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
望ましい効果	わずか	小	中	大		さまざま	わからない
望ましくない効果	大	中	小	わずか		さまざま	わからない
エビデンスの確実性	非常に低	低	中	高			わからない
価値観	重要な不確実性，ばらつきあり	重要な不確実性，ばらつきの可能性あり	重要な不確実性，ばらつきはおそらくなし	重要な不確実性，ばらつきなし			
効果のバランス	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	わからない
必要資源量	大きなコスト	中等度のコスト	無視できる程度のコスト	中等度の節減	大きな節減	さまざま	わからない
必要資源のエビデンスの確実性	非常に低	低	中	高			採用研究なし
費用対効果	対照が優位	対照がおそらく優位	介入も対照も優位ではない	介入がおそらく優位	介入が優位	さまざま	採用研究なし
公平性	減る	おそらく減る	おそらく影響なし	おそらく増える	増える	さまざま	わからない
容認性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない
実行可能性	いいえ	おそらく，いいえ	おそらく，はい	はい		さまざま	わからない