

Online Appendix

Supplementary Table 1.

Search strategy and number of hits per database, current to 6 January 2018.

		CENTAL	EMBASE (Ovid)	Medline (Ovid)	PsycINFO (Ovid)
#1	(adverse outcome* or complication* or drug fatalit* or drug hypersensitivity or drug reaction* or drug safety or drug tolerance or patient safety or safety or side effect* or contraindication*):ti,ab	28,005	1,971,280	1,317,397	104,037
#2	(safety or adverse or tolerability or tolerance or tolderat* or harm or harms or harmful or injur* or damage* or impair* or complication* or risk or risks):ti, ab	52,167	6,085,797	4,307,670	675,292
#3	(side effect* or treatment emergency or undesirable effect*):ti,ab	10,866	318,633	216,094	29,905
#4	(suicid* or death*):mp	53,745	1,245,150	844,745	142,528
#5	(agitat* or constipate* or delusion* or diarrh* or dissociat* or dizz* or drug month or hallucinate* or headache* or hypoten* or hyperten* or insommi* or manic or mania or hypomani* or nausea* or seizur* or sleep* or drows* or urin* or vomit* or temor*):ti,ab,sh	17,598	2,740,691	1,752,836	209,777
#6	ae.fs [Floating subheading: adverse drug reaction]	3,103	1,233,595	1,723,054	-
#7	ct.fs [Floating subheading: contraindications]	440	1,051,958	402,339	8,885
#8	#1 or #2 or #3 or #4 or #5 or #6 or #7	60,051	9,970,310	8,076,658	919,166
#9	*Ketamine/ae, to	2,980	2,449	877	-
#10	(ketamine* or ketaject or ketalar or ketanest or ketaset or ketalean):ti,ab,id,sh	2,971	37,161	17,121	2,855
#11	#9 or #10	2,982	37,161	17,121	2,855
#12	(depression or depressive or mood disorder* or affective disorder* or bipolar or psychosis or psychotic disorder* or schizophren* or schizoaffect* or eating disorder* or anorexi* or bulimi* or EDNOS or anxi* or anxiety disorder* or general* anxiety or GAD or obsess* or OCD or trauma* or PTSD or personality disorder* or paranoi* or schizoid or schizotyp* or anti\$social or borderline or histroni* or narcissisti* or avoidant or dependent):ti,ab,id,sh,tm	14,957	2,998,007	2,260,413	785,078
#13	((animal*1 or nonhuman) not human*1 and (animal*1 or nonhuman)):sh	8,452	5,540,662	4,775,747	7,010
#14	exp *anesthesiological procedure/	25,482	321,866	-	-
#15	#13 or #14	-	5,823,254	4,775,747	7,010
#16	#8 and #11 and #12	-	5,939	2,202	808
#17	Limit #16 to human	-	4,375	1,296	490
#18	#17 not #14	505	3,256	1,296	490

Risk of Bias.

For further information, please refer to Chapter 8 of the Cochrane Handbook, available from: <https://methods.cochrane.org/bias/assessing-risk-bias-included-studies>

In short, bias refers to any systematic deviation from the ‘true’ population effect, leading to either under- or over-estimation of the intervention effect in the individual studies included in a meta-analysis. Different biases in different studies can help to explain the variation in results observed for the studies included in a meta-analysis and it is for this reason the Cochrane Collaboration encourages reviewers assess risk of bias in all studies included in any review – this process is referred to as the assessment of risk of bias in the included studies.

Sources of Bias in Randomised Controlled Trials.

Source	Explanation
Selection bias	Systematic differences in the characteristics of the intervention and control groups at baseline. Whilst the randomisation process is designed to minimise selection bias, this is only successfully achieved if participants are allocated to the intervention and control groups via a randomised process (sequence generation). Additionally, it should also be the case that researchers and others involved in the study have no foreknowledge of forthcoming allocations such that the randomisation sequence cannot be intentionally subverted (allocation concealment).
Performance bias	Systematic differences between the intervention and control groups in terms of the standard of care provided or in exposure to factors other than the intervention of interest. Blinding (or masking) of participants and study personnel is often recommended to reduce the risk that participants may be exposed to differing standards of care in either the intervention or control groups; however, blinding may not always be possible to achieve given the nature of the intervention being tested.
Detection bias	Systematic differences between the intervention and control groups in how the primary and secondary outcomes are determined. Blinding (or masking) of outcome assessors is often recommended to reduce the risk that knowledge of which group – intervention or control – a participant was assigned may affect outcome measurement or interpretation.
Attrition bias	Systematic differences between the intervention and control groups from the study; for example, differential levels of withdrawal between the intervention and control groups can lead to differential levels of incomplete outcome data between the groups due to differential reasons for exclusion of participant data from analyses or differential availability of participant data in the intervention and control groups.
Reporting bias	Systematic differences between reported and unreported findings. For example, where statistically significant differences between the intervention and control groups are more likely to be reported than statistically non-significant differences.
Other biases	These may include particular sources of bias relevant to certain study types; for example, carry-over effects for cross-over trials, recruitment bias in cluster-randomised trials, contamination effects, and other biases that may be of relevance to studies conducted in particular clinical settings.

The Cochrane Risk of Bias Tool.

Source	Support for Judgement	Review Authors' Judgement
Selection bias		
Random sequence generation	Describe the method(s) used to generate the allocation sequence in sufficient detail to allow an assessment of whether (or not) the method would result in comparable groups.	Selection bias (biased allocation to the intervention and control groups) due to inadequate generation of the randomisation sequence. Bias can be rated as “high”, “unclear”, or “low”.
Allocation sequence	Describe the method(s) used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, participant enrolment.	Selection bias (biased allocation to the intervention and control groups) due to inadequate concealment of allocation prior to assignment. Bias can be rated as “high”, “unclear”, or “low”.
Performance bias		
Blinding of participants and personnel	Describe all measures used, if any, to blind study participants and clinical personnel from knowledge of which intervention a participant received. Provide information relating to whether the intended blinding was effective.	Performance bias due to knowledge of the allocated interventions by participants and personnel during the study. Bias can be rated as “high”, “unclear”, or “low”.
Detection bias		
Blinding of outcome assessors	Describe all measures used, if any, to blind outcome assessors from knowledge of which intervention a participant received. Provide information relating to whether the intended blinding was effective.	Detection bias due to knowledge of the allocated interventions by outcome assessors. Bias can be rated as “high”, “unclear”, or “low”.
Attrition bias		
Incomplete outcome data	Describe the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. State whether attrition and exclusions were reported, the numbers in each analysed group (compared with the total number of participants randomized to each group), reasons for attrition and exclusions, if reported, and any re-inclusions in the analyses. Assessments should be made for each main outcome (or class of outcomes).	Attrition bias due to the amount, nature, and/or handling of incomplete outcome data. Bias can be rated as “high”, “unclear”, or “low”.
Reporting bias		
Selective reporting	State how the possibility of selective outcome reporting was examined by the review authors, and what was found.	Reporting bias due to selective outcome reporting. Bias can be rated as “high”, “unclear”, or “low”.
Other bias		
Other sources of bias	State any important concerns about bias not addressed in the other domains in the tool.	Bias due to problems not covered elsewhere in the table. Bias can be rated as “high”, “unclear”, or “low”.

Supplementary Table 2.

Risk of bias table.

Study ID	Adequate Sequence Generation	Allocation Concealment	Participant and Personnel Blinding	Outcome Assessor Blinding	Incomplete Outcome Data	Selective Reporting Bias	Other Bias
Anderson 2017 ^{a1}	<p>Quote: "Participants were randomly assigned (1:1)...us[ing] permuted block randomization" (p. 367).</p> <p>Judgement: Use of permuted block randomisation likely to have minimised bias in the sequence generation process.</p> <p>Rating: low risk.</p>	<p>Quote: "Randomization was done by the Christie Hospital Clinical Trials Co-ordination Unit...The randomization code... was provided to the local site pharmacies" (p. 367).</p> <p>Judgement: Randomisation by an offsite researcher is likely to have ensured treatment allocation was successfully concealed.</p> <p>Rating: low risk.</p>	<p>Quote: "...patients...were masked to treatment allocation" (p. 367). Additionally, "[A]ssessment and ECT treatment teams were masked to treatment allocation, although the anaesthetists administering the study medication were not" (p. 367).</p> <p>Judgement: Although participants and other clinical personnel were blind to treatment allocation, anaesthetists were not.</p> <p>Rating: unclear risk.</p>	<p>Quote: "Researchers responsible for outcome assessment did not attend ECT sessions" (p. 367).</p> <p>Judgement: Blinding of outcome assessors is likely to have been successfully achieved using this method.</p> <p>Rating: low risk.</p>	<p>Quote: "Statistical analysis was based on a modified intention-to-treat population defined as all patients who received the first ECT, depending on the availability of data. Missing efficacy data were filled in by prorating on the provision that at least 70% of items in the same scale had been completed" (p. 369).</p> <p>Judgement: Data on the primary outcome in this review, suicidality as rated by the MADRS suicidality item, was based on available case data only.</p> <p>Rating: high risk.</p>	<p>Judgement: No reason to suspect selective outcome reporting as all measures outlined in the study methods section were reported.</p> <p>Rating: low risk.</p>	<p>Judgment: no other bias apparent.</p> <p>Rating: low risk.</p>
George 2017 ^{a2}	<p>Quote: "A randomization list was created, with the position of the control treatment randomly assigned by a computer-generated random number sequence" (p. 1201).</p> <p>Judgement: Use of a computer-generated sequence is likely to have minimised bias.</p> <p>Rating: low risk.</p>	<p>Quote: "The randomization list was kept in a locked cupboard only accessible to the anaesthetist drawing up the study drug" (p. 1201).</p> <p>Judgement: suggests allocation status was able to be successfully concealed.</p> <p>Rating: low risk.</p>	<p>Quote: "An anaesthetist investigator sequentially assigned participants to the randomization list, drew up the medications for each treatment, and blinded the injection volume by taping over the barrel of the syringe. This anaesthetist was not involved in any treatment sessions, contact with participants, or study assessments...Participants...were aware that the RCT included one session with a control treatment but did not know the control treatment was inserted within the first three sessions" (p. 1201).</p> <p>Judgement: suggests blinding of participants and study personnel was able to be successfully achieved.</p> <p>Rating: low risk.</p>	<p>Quote: "Raters...were aware that the RCT included one session with a control treatment but did not know the control treatment was inserted within the first three sessions" (p. 1201).</p> <p>Judgement: suggests blinding out outcome assessors was able to be successfully achieved.</p> <p>Rating: low risk.</p>	<p>Quote: "An intention-to-treat analysis of MADRS scores for all participants enrolled in the RCT was performed" (p. 1202).</p> <p>Judgement: data on MADRS outcomes for all 16 participants are reported, with 0.0% drop out.</p> <p>Rating: low risk.</p>	<p>Judgement: No reason to suspect selective outcome reporting. All measures outlined in study protocol were measured.</p> <p>Rating: low risk.</p>	<p>Judgment: no other bias apparent.</p> <p>Rating: low risk.</p>

Grunebaum 2017 ^{a3}	<p>Quote: "We used a permuted block design with 1:1 assignment between treatment and block size randomized between 4 and 5 with equal probability" (p. 178).</p> <p>Judgement: Use of permuted block randomisation likely to have minimized bias on the sequence generation process.</p> <p>Rating: low risk.</p>	<p>Judgment: no specific details provided.</p> <p>Rating: unclear risk.</p>	<p>Quote: "Patients...were blind to treatment...participants...answered the question of whether they thought the infusion was midazolam, ketamine, or 'no idea'. Of the participants randomized to ketamine, five of seven (71.4%) correctly guessed their infusion drug...versus seven of nine (77.8%) randomized to midazolam" (p.178 and 180). "...[P]sychiatrists...were blind to treatment" (p. 178).</p> <p>Judgement: Although participants were blinded to treatment assignment, over half were able to correctly guess which arm they had been assigned to, suggesting blinding was not successful. Clinical personnel were successfully blinded to treatment assignment.</p> <p>Rating: unclear risk.</p>	<p>Quote: "...[A]ssessors were blind to treatment...assessors answered the question of whether they thought the infusion was midazolam, ketamine, or 'no idea'. Clinical assessors guessed correctly after four of seven (57.1%) ketamine and five of nine (55.6%) midazolam infusions" (p.178 and 180).</p> <p>Judgement: Although outcome assessors were blinded to treatment assignment, over half were able to correctly guess which arm participants had been assigned to, suggesting blinding was not successful.</p> <p>Rating: unclear risk.</p>	<p>Quote: "The intent to treat analysis included all randomized participants" (p. 178).</p> <p>Judgement: Use of the intent to treat analysis is likely to have reduced risk of bias.</p> <p>Rating: low risk.</p>	<p>Judgement: No reason to suspect selective outcome reporting as all measures outlined in the study methods section were reported.</p> <p>Rating: low risk.</p>	<p>Judgment: no other bias apparent.</p> <p>Rating: low risk.</p>
Grunebaum 2018 ^{a4}	<p>Quote: "We used a permuted block design with 1:1 assignment between treatment [groups] and block size randomized between 4 and 5 with equal probability" (p. 3).</p> <p>Judgement: Use of a permuted block design is likely to have minimised bias in the generation of the randomisation sequence.</p> <p>Rating: low risk.</p>	<p>Judgment: no specific details provided.</p> <p>Rating: unclear risk.</p>	<p>Quote: "Patients were blind to treatment...study personnel were blind to treatment" (p. 3).</p> <p>Judgement: Although no specific information was reported on the method(s) used to ensure participants and clinical personnel were blind to treatment allocation, the authors do state that both participants and clinical personnel had been blind to treatment assignment.</p> <p>Rating: low risk.</p>	<p>Judgment: no specific details provided.</p> <p>Rating: unclear risk.</p>	<p>Judgment: no specific details provided.</p> <p>Rating: unclear risk.</p>	<p>Judgement: No reason to suspect selective outcome reporting. All measures reported in the study methods section were reported.</p> <p>Rating: low risk.</p>	<p>Judgment: no other bias apparent.</p> <p>Rating: low risk.</p>
Hu 2016 ^{a5}	<p>Quote: "...[P]atients were randomized according to a table of random numbers..." (p. 624).</p> <p>Judgement: Whilst use of a random numbers table is</p>	<p>Judgment: no specific details provided.</p> <p>Rating: unclear risk.</p>	<p>Quote: "The solutions were provided in identical 50-ml syringes. Ketamine forms a clear solution when dissolved in 0.9% saline." Additionally, "[t]he anesthesiologist was blind to the</p>	<p>Quote: "Two raters...blind to the study protocol and treatment assignment independently assessed patients..." (p.625).</p>	<p>Quote: "The analyses were conducted in the modified intent-to-treat sample, i.e., including patients with a baseline and ≥ 1 follow-up assessment. Continuous and categorical outcomes</p>	<p>Judgement: No reason to suspect selective outcome reporting. All measures outlined in the study methods</p>	<p>Judgment: no other bias apparent.</p> <p>Rating: low risk.</p>

	likely to have minimised bias in the generation of the random sequence. However, it is unclear whether the random numbers table used in this study was closed or open. Rating: unclear risk.		group membership of patients" (p. 625). Judgement: Suggests blinding could have been convincingly achieved. Rating: low risk.	Judgement: Suggests blinding was achieved. Rating: low risk.	were analysed as the last observation carried forward" (p. 625). Judgement: Reads as available case analysis. The last observation carried forward method has been shown to introduce bias. Rating: high risk.	section were reported. Rating: low risk.	
Kudoh 2002 ^{a6}	Quote: "The randomization was performed on computer-generated codes" (p. 115). Judgement: Use of a computer-generated sequence is likely to have minimised risk of bias. Rating: low risk.	Judgment: no specific details provided. Rating: unclear risk.	Judgment: no specific details provided. Rating: unclear risk.	Judgment: no specific details provided. Rating: unclear risk.	Judgment: no specific details provided. Rating: unclear risk.	Judgement: No reason to suspect selective outcome reporting. All measures outlined in the study methods section were reported. Rating: low risk.	Judgment: no other bias apparent. Rating: low risk.
Loo 2016 ^{a7}	Quote: "The randomization list (position of placebo treatment) was allocated according to a computer-generated random number sequence" (p. 50). Judgement: use of a computer-generated sequence likely to have minimised bias. Rating: low risk.	Quote: "The randomization list was kept in a folder in a locked room, and the anaesthetists were instructed to keep treatment allocation concealed" (p. 50). Judgement: suggests allocation status was able to be successfully concealed. Rating: low risk.	Quote: "Participants...were aware that one treatment was a control, but were blind to its position, including that it was place within the first three sessions" (p. 50). Judgement: although participants may not have been informed as to which phase was the control phase, they were nonetheless informed that one phase would constitute a control phase and may have been able to guess which phase this was. Additionally, no specific information was provided on clinical personnel blinding. Rating: unclear risk.	Quote: "[R]aters... were aware that one treatment was a control, but were blind to its position, including that it was place within the first three sessions" (p. 50). Judgement: although participants may not have been informed as to which phase was the control phase, they were nonetheless informed that one phase would constitute a control phase and may have been able to guess which phase this was. Rating: unclear risk.	Quote: "Data from all 15 participants were used in the analyses" (p. 51). Judgement: data from all randomised participants used, with no loss to follow-up. Rating: low risk.	Judgement: No reason to suspect selective outcome reporting. All measures outlined in study protocol were measured. Rating: low risk.	Judgment: no other bias apparent. Rating: low risk.
Murrough 2013 ^{a8}	Quote: "Randomly assigned in a 2:1 ratio..." (p. 1135). Judgement: No further specifics on how the random sequence was generated,	Quote: "The study research pharmacist prepared sealed envelopes that	Quote: "...patients...were masked to treatment assignment...all other study personnel, including investigators, anesthesiologists	Quote: "...raters...and data analysts were masked to treatment assignment" (p. 1135).	Quote: "Modified intention-to-treat analyses included all randomly assigned patients with baseline measurement and	Judgement: No reason to suspect selective outcome reporting. All measures outlined in	Judgment: no other bias apparent. Rating: low risk.

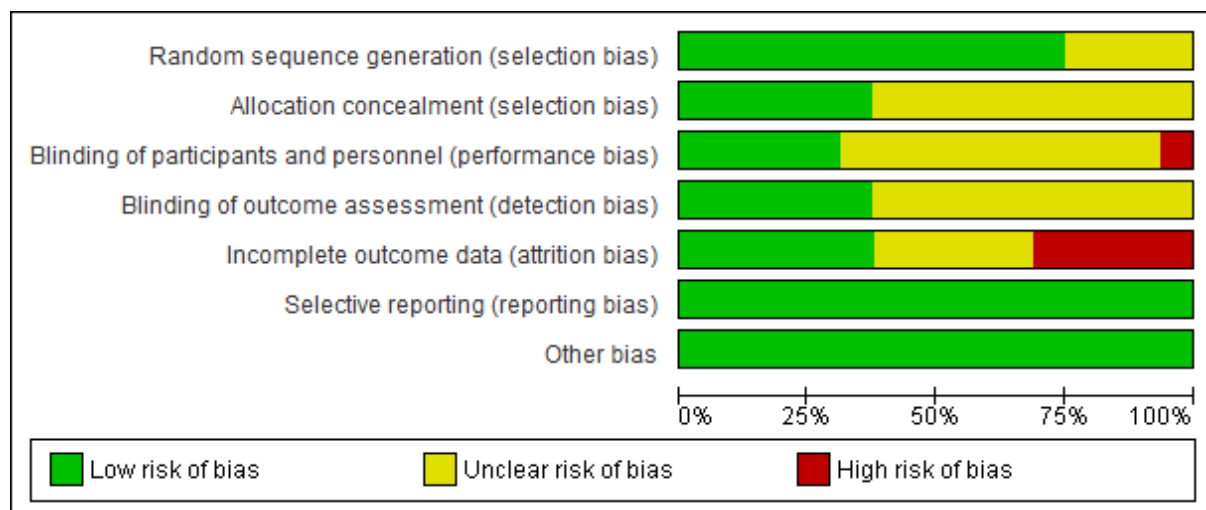
	<p>therefore, it is difficult to judge to what degree there was bias in the generation of the random sequence.</p> <p>Rating: unclear risk.</p>	<p>contained the drug identity" (p. 1135).</p> <p>Judgement: Use of sealed envelopes would suggest that allocation was successfully concealed.</p> <p>Rating: low risk.</p>	<p>...were masked to treatment assignment" (p. 1135).</p> <p>Judgement: Suggests participants and clinical personnel were blind to treatment assignment.</p> <p>Rating: low risk.</p>	<p>Judgement: Suggests outcome assessors were blind to treatment assignment.</p> <p>Rating: low risk.</p>	<p>at least one post baseline measurement" (p. 1136).</p> <p>Judgement: Reads as available case data analysis.</p> <p>Rating: unclear risk.</p>	<p>the study methods section were reported.</p> <p>Rating: low risk.</p>	
Murrough 2015 ^{a9}	<p>Quote: "The randomization scheme was generated by the research pharmacy using permuted blocks of size six..." (p. 3572).</p> <p>Judgement: Use of permuted blocks of equal size is likely to have minimized bias in the generation of the random sequence.</p> <p>Rating: low risk.</p>	<p>Judgment: no specific details provided.</p> <p>Rating: unclear risk.</p>	<p>Quote: "...all study investigators, anesthesiologists...were blind to treatment assignment" (p. 3572).</p> <p>Judgement: Suggests blinding was successfully achieved. No specific information on blinding of participants were provided, however.</p> <p>Rating: unclear risk.</p>	<p>Quote: "...all...raters were blind to treatment assignment" (p. 3572).</p> <p>Judgement: Suggests blinding was successfully achieved.</p> <p>Rating: low risk.</p>	<p>Quote: "The intention-to-treat sample included all participants who were randomized and completed at least one post-treatment assessment" (p. 3573).</p> <p>Judgement: Reads as complete case analysis as participants were required to have completed at least one post-intervention assessment.</p> <p>Rating: high risk.</p>	<p>Judgement: No reason to suspect selective outcome reporting. All measures outlined in the study methods section were reported.</p> <p>Rating: low risk.</p>	<p>Judgment: no other bias apparent.</p> <p>Rating: low risk.</p>
Ray-Griffith 2017 ^{a10}	<p>Quote: "Randomization occurred using the 'urn' method to ketamine or methohexital anesthesia before the initial treatment" (p. 267).</p> <p>Judgement: Although the urn design approaches complete randomization as the sample size increases, the small size of this trial may have precluded complete randomisation from being achieved. Nevertheless, it is acknowledged that that urn design is not as vulnerable to experimental bias as other small study randomisation procedures commonly used</p> <p>Rating: low risk.</p>	<p>Judgment: no specific details provided.</p> <p>Rating: unclear risk.</p>	<p>Quote: "The treatment team, consisting of the attending psychiatrist and anesthesiologist administering ECT, were not blinded to the anesthetic agent..." (p. 267).</p> <p>Judgement: No blinding of clinical personnel ensured.</p> <p>Rating: high risk.</p>	<p>Quote: "Two raters completed all assessments blinded to the anesthetic agent" (p. 267).</p> <p>Judgement: Blinding of outcome assessors was ensured.</p> <p>Rating: low risk.</p>	<p>Judgment: no specific details provided.</p> <p>Rating: unclear risk.</p>	<p>Judgement: No reason to suspect selective outcome reporting. All measures outlined in the study methods section were reported.</p> <p>Rating: low risk.</p>	<p>Judgment: no other bias apparent.</p> <p>Rating: low risk.</p>
Sos 2013 ^{a11}	<p>Quote: "...randomised by a flip of a coin" (p. 59).</p>	<p>Judgment: no specific details provided.</p>	<p>Judgment: no specific details provided.</p>	<p>Judgment: no specific details provided.</p>	<p>Quote: "...twenty-seven patients received the intended treatment and</p>	<p>Judgement: No reason to suspect selective outcome</p>	<p>Judgment: no other bias apparent.</p>

	<p>Judgement: Coin flip describes a random component in the sequence generation process.</p> <p>Rating: low risk.</p>	<p>Rating: unclear risk.</p>	<p>Rating: unclear risk.</p>	<p>Rating: unclear risk.</p>	<p>were included in all analyses (intention-to-treat)" (p. 59).</p> <p>Judgement: Use of intention-to-treat analyses with no drop-out from either the intervention or control means there was little bias.</p> <p>Rating: low risk.</p>	<p>reporting. All measures outlined in the study methods section were reported.</p> <p>Rating: low risk.</p>	<p>Rating: low risk.</p>
Su 2017 ^{a12}	<p>Judgment: no specific details provided.</p> <p>Rating: unclear risk.</p>	<p>Judgment: no specific details provided.</p> <p>Rating: unclear risk.</p>	<p>Judgment: no specific details provided.</p> <p>Rating: unclear risk.</p>	<p>Judgment: no specific details provided.</p> <p>Rating: unclear risk.</p>	<p>Quote: "All available data on individuals were used" (p. 2485).</p> <p>Judgement: Analysing data only on the basis of those with complete data suggests the use of available case data.</p> <p>Rating: high risk.</p>	<p>Judgement: No reason to suspect selective outcome reporting. All measures outlined in the study methods section were reported.</p> <p>Rating: low risk.</p>	<p>Judgment: no other bias apparent.</p> <p>Rating: low risk.</p>
Zarate 2012 ^{a13}	<p>Quote: "Patients were randomly assigned to the order in which they received the two infusions via a random-numbers chart" (p. 940).</p> <p>Judgement: Use of a random numbers table likely to have minimised bias in the generation of the randomisation sequence.</p> <p>Rating: low risk.</p>	<p>Judgment: no specific details provided.</p> <p>Rating: unclear risk.</p>	<p>Quote: "Study solutions were supplied in identical 50-mL syringes containing either 0.9% saline or ketamine with the additional volume of saline to total 50 mL; when dissolved in 0.9% saline, ketamine forms a clear solution" (p. 940). In addition, "[a]ll staff, including the anaesthesiologist were blind to whether placebo or drug was being administered" (p. 940).</p> <p>Judgement: the use of identical looking syringes is likely to have ensured blinding of participants to treatment allocation could be successfully achieved. In addition, all study personnel were blind to treatment allocation.</p> <p>Rating: low risk.</p>	<p>Quote: "Patient ratings were performed by research nurses and/or psychologists who trained together to establish reliability (p. 941).</p> <p>Judgement: However, no specific information was provided on whether these staff were also blind to treatment allocation.</p> <p>Rating: unclear risk.</p>	<p>Quote: "The primary intent-to-treat analysis included all available data" (p. 941).</p> <p>Judgement: Use of intention-to-treat analyses with no drop-out from either the intervention or control means there was little bias.</p> <p>Rating: low risk.</p>	<p>Judgement: No reason to suspect selective outcome reporting. All measures outlined in the study protocol were reported.</p> <p>Rating: low risk.</p>	<p>Judgment: no other bias apparent.</p> <p>Rating: low risk.</p>
Canuso, 2014 ^{a14}	<p>Quote: "[E]ligible participants were randomly assigned (1:1) to receive either intranasal esketamine (84 mg) or matching placebo...Randomization was</p>	<p>Quote: "[I]nvestigators were not provided with the randomization codes." (p. 622).</p>	<p>Judgement: The trial is described throughout as "double-blind"; however, no specific details are provided as to how blinding was achieved.</p>	<p>Judgement: The trial is described throughout as "double-blind"; however, no specific details are provided as</p>	<p>Quote: "All randomized participants who received at least one dose of study medication in the double-blind phase were included in the safety analysis set.</p>	<p>Judgement: No reason to suspect selective outcome reporting. All measures outlined in</p>	<p>Judgment: no other bias apparent.</p> <p>Rating: low risk.</p>

	<p>balanced using randomly permuted blocks and stratified by study center and type of standard-of-care antidepressant (i.e., monotherapy or antidepressant plus augmentation therapy). A computerized system was used for randomization." (p. 622).</p> <p>Judgement: Use of a computer-generated sequence, generated through the use of a permuted block design, is likely to have minimised bias.</p> <p>Rating: low risk.</p>	<p>Judgement: suggests allocation could have been concealed.</p> <p>Rating: low risk.</p>	<p>Rating: unclear risk.</p>	<p>to how blinding was achieved.</p> <p>Rating: unclear risk.</p>	<p>Efficacy data were analyzed in an intent-to-treat analysis set, which included all participants in the safety analysis set who had MADRS scores at baseline and at 4 hours postbaseline on day 1." (p. 624).</p> <p>Judgement: Reads as available case analysis. Data would suggest that, during the double-blind phase, there was 0.0% drop-out. However, during the naturalistic follow-up phase, there was a 25.7% drop-out.</p> <p>Rating: unclear risk.</p>	<p>the study protocol were reported.</p> <p>Rating: low risk.</p>	
Ionescu, 2019 ^{a15}	<p>Quote: "Patients were randomized (immediately following the pre-infusion phase)...[g]roup allocation was completed by a computer-generated randomization algorithm." (p.518).</p> <p>Judgement: Use of a computer-generated sequence is likely to have minimised bias.</p> <p>Rating: low risk.</p>	<p>Quote: "The randomization list was maintained in a locked cabinet by a senior anaesthesiologist. All clinicians, patients, and raters were blind to the randomization assignments." (p. 581).</p> <p>Judgement: would suggest allocation could have been successfully concealed.</p> <p>Rating: low risk.</p>	<p>Quote: "All clinicians, patients, and raters were blind to the randomization assignments." (p. 518). However, there is no clear information as to whether participants themselves were blind to treatment allocation.</p> <p>Judgement: rated as unclear given that no information was provided as to whether the participants themselves were blind to treatment allocation.</p> <p>Rating: unclear risk.</p>	<p>Judgment: no specific details provided.</p> <p>Rating: unclear risk.</p>	<p>Quote: "For the analysis of depression (primary outcome; HDRS total score) and SI (secondary outcome; C-SSRS SI score and C-SSRS SI intensity rating), the intent-to-treat (ITT) model was utilized to include all patients." (p. 518).</p> <p>Judgement: Use of the intent to treat analysis is likely to have reduced risk of bias.</p> <p>Rating: low risk.</p>	<p>Judgement: No reason to suspect selective outcome reporting. All measures outlined in the study protocol were reported.</p> <p>Rating: low risk.</p>	<p>Judgment: no other bias apparent.</p> <p>Rating: low risk.</p>

References to included studies.

- a1 Anderson I, Blamire A, Branton T, et al. Ketamine augmentation of electroconvulsive therapy to improve neuropsychological and clinical outcomes in depression (Ketamine-ECT): A multicentre, double-blind, randomised, parallel-group, superiority trial. *The Lancet Psychiatry* 2017; **4**: 365-77.
- a2 George D, Gálvez V, Martin D, et al. Pilot randomized controlled trial of titrated subcutaneous ketamine in older patients with treatment-resistant depression. *Am J Geriatr Psychiatry* 2017; **25**: 1199-209.
- a3 Grunebaum M, Ellis S, Keilp J, et al. Ketamine versus midazolam in bipolar depression with suicidal thoughts: A pilot midazolam-controlled randomized clinical trial. *Bipolar Disord* 2017; **19**: 176-83.
- a4 Grunebaum M, Galfalvy H, Choo T-H, et al. Ketamine for rapid reduction of suicidal thoughts in major depression: A midazolam-controlled randomized clinical trial. *Am J Psychiatry* 2018; **175**: 327-55.
- a5 Hu Y-D, Xiang Y-T, Fang J-X, et al. Single i.v. ketamine augmentation of newly initiated escitalopram for major depression: Results from a randomized, placebo-controlled 4-week study. *Psychol Med* 2016; **46**: 623-35.
- a6 Kudoh A, Takahira Y, Katagai H, Takazawa T. Small-dose ketamine improves the postoperative state of depressed patients. *Anesth Analg* 2002; **95**: 114-8.
- a7 Loo C, Gálvez V, O'Keefe E, et al. Placebo-controlled pilot trial testing dose titration and intravenous, intramuscular and subcutaneous routes for ketamine in depression. *Acta Psychiatr Scand* 2016; **134**: 48-56.
- a8 Murrough J, Iosifescu D, Chang L, et al. Antidepressant efficacy of ketamine in treatment-resistant major depression: A two-site randomized controlled trial. *Am J Psychiatry* 2013; **170**: 1134-42.
- a9 Murrough J, Soleimani L, DeWilde K, et al. Ketamine for rapid reduction of suicidal ideation: A randomized controlled trial. *Psychol Med* 2015; **45**: 3571-80.
- a10 Ray-Griffith S, Eads L, Han X, Golden K, Stowe Z. A randomized pilot study comparing ketamine and methohexital anesthesia for electroconvulsive therapy in patients with depression. *J ECT* 2017; **33**: 268-71.
- a11 Sos P, Klirova M, Novak T, Kohutova B, Horacek J, Palenicek T. Relationship of ketamine's antidepressant and psychotomimetic effects in unipolar depression. *Activitas Nervosa Superior Rediviva* 2013; **55**: 57-63.
- a12 Su T-P, Chen M-H, Li C-T, et al. Dose-related effects of adjunctive ketamine in Taiwanese patients with treatment-resistant depression. *Neuropsychopharmacol* 2017; **42**: 2482-92.
- a13 Zarate Jr. C, Brutsche N, Ibrahim L, et al. Replication of ketamine's antidepressant efficacy in bipolar depression: A randomized controlled add-on trial. *Biol Psychiatry* 2012; **71**: 939-46.
- a14 Canuso CM, Singh JB, Fedgchin M, et al. Efficacy and safety of intranasal esketamine for the rapid reduction of symptoms of depression and suicidality in patients at imminent risk for suicide: Results of a double-blind, randomized, placebo-controlled study. *Am J Psychiatry* 2018; **175**: 620-30.
- a15 Ionescu DF, Bentley KH, Eikermann M, et al. Repeat-dose ketamine augmentation for treatment-resistant depression with chronic suicidal ideation: A randomized, double blind, placebo controlled trial. *J Affect Disord* 2019; **243**: 516-24.



Supplementary Figure 1.

Judgements about each risk of bias item presented as percentages across all included studies.

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Anderson 2017	+	+	?	+	-	+	+
Canuso 2018	+	+	?	?	?	+	+
George 2017	+	+	+	+	+	+	+
Grunebaum 2017	+	?	?	?	+	+	+
Grunebaum 2018	+	?	+	?	?	+	+
Hu 2016	?	?	+	+	-	+	+
Ionescu 2019	+	+	?	?	+	+	+
Kudoh 2002	+	?	?	?	?	+	+
Loo 2016	+	+	?	?	+	+	+
Murrough 2013	?	+	+	+	?	+	+
Murrough 2015	+	?	?	+	-	+	+
Ray-Griffith 2017	+	?	-	+	?	+	+
Sos 2013	+	?	?	?	+	+	+
Su 2017a	?	?	?	?	-	+	+
Su 2017b	?	?	?	?	-	+	+
Zarate Jr. 2012	+	?	+	?	+	+	+

Supplementary Figure 2.

Judgements about each risk of bias item for each included study.

Supplementary Information: GRADE evidence summary

For further information on the historical background and development of the GRADE evidence summary, please see:

http://www.gradeworkinggroup.org/docs/Criteria_for_using_GRADE_2016-04-05.pdf




In short, the GRADE recommendations for synthesising the quality of evidence from systematic reviews and meta-analyses were developed to provide some objectivity to the process for rating the overall quality of the evidence obtained from a systematic review. Suggested criteria for stating that the GRADE approach was used, include:

1. The certainty in the evidence (also known as the quality of the evidence, or the confidence in the estimates) should be defined consistently with the definitions as used by the GRADE Working Group;^{1,2}
2. Explicit consideration should be given to each of the GRADE domains (summarised below) for assessing the certainty in the evidence;
3. The overall certainty in the evidence should be assessed for each important outcome using four or three categories (i.e., high, moderate, low, or very low) and definitions for each of these categories should be consistent with the definitions as used by the GRADE Working Group;^{1,2}
4. Evidence summaries and evidence to decision criteria should be used as the basis of judgement about the certainty in the evidence and the strength of recommendations. Ideally, evidence profiles should be used to assess the certainty in the evidence and these should be based on systematic reviews. At a minimum, the evidence that was assessed and the methods used to identify and appraise the evidence should be clearly described;
5. Explicit consideration should be given to each of the GRADE criteria for determining the direction and strength of a recommendation or decision. Ideally, GRADE evidence to decision frameworks should be used to document the considered research evidence, additional considerations and judgements transparently;
6. The strength of recommendations should be assessed using two categories (for or against) and definitions for each category such as strong and weak/conditional that are consistent with the definitions as used by the GRADE Working Group.^{1,2}

References:

1. Andrews JC, Schünemann HJ, Oxman AD, Pottie K, Meerpohl JJ, Coello PA, Rind D, Montori VM, Brito JP, Norris S, Elbarbary M, Post P, Nasser M, Shukla V, Jaeschke R, Brozek J, Djulbegovic B, Guyatt G. GRADE guidelines: 15. Going from evidence to recommendation—determinants of a recommendation's direction and strength. *J Clin Epidemiol.* 66: 726-35, 2013.
2. Andrews J, Guyatt G, Oxman AD, Alderson P, Dahm P, Falck-Ytter Y, Nasser M, Meerpohl J, Post PN, Kunz R, Brozek J, Vist G, Rind D, Akl EA, Schünemann HJ. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. *J Clin Epidemiol.* 66: 719-25, 2013.

Supplementary Table 3.
GRADE evidence summary for each time point.

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	ketamine	placebo or other active comparators	Relative (95% CI)	Absolute (95% CI)		
Immediate effect (within 4 hours)												
10	randomised trials	not serious	very serious ^a	not serious	not serious	risk of selection bias ^b	183	149	-	SMD 0.51 SD lower (1.00 lower to 0.03 lower)	 LOW	IMPORTANT
Very ultra rapid effect (greater than 4 but less than/equal to 12 hours)												
0												
Ultra rapid effect (greater than 12 but less than/equal to 24 hours)												
10	randomised trials	not serious	serious ^c	not serious	not serious	risk of selection bias	192	201	-	SMD 0.63 SD lower (0.99 lower to 0.26 lower)	 LOW	IMPORTANT
Rapid effect (greater than 24 but less than/equal to 72 hours)												
8	randomised trials	not serious	serious ^c	not serious	not serious	risk of selection bias	125	94	-	SMD 0.57 SD lower (0.99 lower to 0.14 lower)	 LOW	IMPORTANT
Early effect (greater than 72 but less than/equal to 2 weeks)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	ketamine	placebo or other active comparators	Relative (95% CI)	Absolute (95% CI)		
11	randomised trials	not serious	not serious	not serious	not serious	risk of selection bias	190	171	-	SMD 0.19 SD lower (0.41 lower to 0.03 higher)	⊕⊕⊕○ MODERATE	IMPORTANT
Acute effect (greater than 2 but less than/equal to 4 weeks)												
5	randomised trials	not serious	not serious	not serious	not serious	risk of selection bias	106	87	-	SMD 0.24 SD lower (0.53 lower to 0.05 higher)	⊕⊕⊕○ MODERATE	IMPORTANT
Longer-term effect (greater than 1 month)												
3	randomised trials	not serious	not serious	not serious	not serious	risk of selection bias	59	63	-	SMD 0.21 SD lower (0.58 lower to 0.16 higher)	⊕⊕⊕○ MODERATE	NOT IMPORTANT

CI: Confidence interval; SMD: Standardised mean difference

Explanation

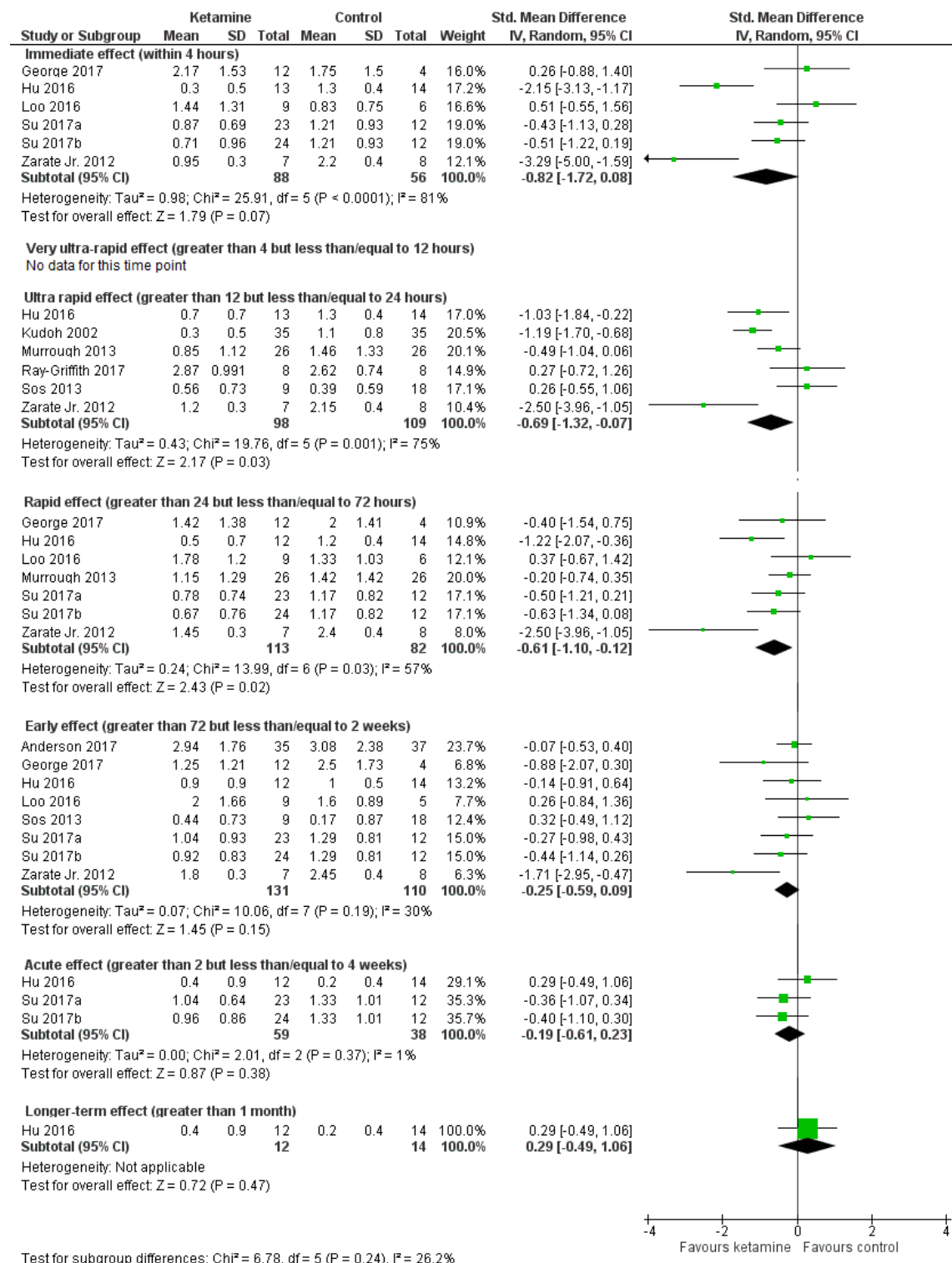
- I^2 value indicated substantial between-study heterogeneity for this outcome at this assessment point.
- A number of studies were downgraded due to the likely presence of selection bias, either as insufficient information on the method used to generate the random sequence was provided, or because insufficient information on the method(s) used to conceal allocation was reported. For a number of trials included in this analysis, therefore, risk of bias for random sequence generation and allocation concealment cannot be ruled out at this assessment point.
- I^2 value was greater than or equal to 50.0%, indicating considerable between-study heterogeneity for this outcome at this assessment point.

Explanation of reasons for downgrading

We GRADEd each pooled estimate for each relevant timepoint according to the following criteria:

- Risk of Bias:** We downgraded this domain by one level when any of the sources of Risk of Bias (as described above) were rated as "high" for any of the studies included in the pooled estimate.
- Inconsistency:** We downgraded this domain by two levels where the I^2 value indicated substantial levels of heterogeneity (i.e., $I^2 \geq 75\%$), or by one level where the I^2 value indicated considerable levels of heterogeneity (i.e., $I^2 \geq 50\%$ by $<75\%$).
- Indirectness:** We protected against indirectness in this review by ensuring all studies included in any meta-analysis reported data using a validated measure of suicidal ideation, and further, by considering the degree to which single item suicidal ideation measures correlate with full suicidal ideation scales within the discussion section of the review. We therefore did not downgrade this domain for any assessment point.
- Imprecision:** We downgraded this domain by one level where the 95% confidence interval included the null value.
- Other Considerations:** We considered here the impact of other biases as identified in the Risk of Bias assessment (details as above) and downgraded by one or more points where any other sources of bias were identified in the studies contributing to the pooled estimate.

Sensitivity analyses.



Supplementary Figure 3.

Random effects standard mean difference (SMD) and 95% confidence interval (CI) for the effect of ketamine on suicidal ideation scores, excluding those trials in which suicidal ideation was assessed using a full-scale measure of suicidal ideation.