UNIVERSITY OF COPENHAGEN
FACULTY OF HEALTH AND MEDICAL SCIENCES

THE EVIDENCE FOR ADJUVANT INTERVENTIONS IN INTENSIVE CARE PATIENTS

– haloperidol for the management of delirium,
stress ulcer prophylaxis, and oxygen supplementation

MARIJA BARBATESKOVIC

PHD THESIS

This thesis has been submitted to the Graduate School of Health and Medical Sciences,
University of Copenhagen, on the 31st of December 2019

THE EVIDENCE FOR ADJUVANT INTERVENTIONS IN INTENSIVE CARE PATIENTS

– haloperidol for the management of delirium,
stress ulcer prophylaxis, and oxygen supplementation

MARIJA BARBATESKOVIC

PHD THESIS

Copenhagen Trial Unit, Centre for Clinical Intervention Research,
Rigshospitalet, Copenhagen University Hospital, Denmark

December 2019

Author

Marija Barbateskovic, MSc Health
Copenhagen Trial Unit, Centre for clinical intervention research
Rigshospitalet, Copenhagen University Hospital, Denmark

Supervisors

Faculty supervisor
Anders Perner, Professor, MD, PhD
Department of Intensive Care, Rigshospitalet, University of Copenhagen, Denmark
Faculty of Health and Medical Science, University of Copenhagen, Denmark

Project supervisor
Jørn Wetterslev, Chief Physician, MD, PhD
Copenhagen Trial Unit, Centre for Clinical Intervention Research
Rigshospitalet, University of Copenhagen, Denmark

Co-supervisor
Janus Christian Jakobsen, Chief Physician/Consultant, MD, PhD
Copenhagen Trial Unit, Centre for Clinical Intervention Research,
Rigshospitalet, University of Copenhagen, Denmark

Members of assessment committee

Chairperson
Lars S. Rasmussen, Professor, DMSci, PhD, MD
Department of Anaesthesia, Rigshospitalet, University of Copenhagen, Denmark
Department of Clinical Medicine, University of Copenhagen, Denmark

Representative from Danish research institution
Robin Christensen, Professor, PhD, M.Sc.
Department of Clinical Research, University of Southern Denmark

Representative from research institution abroad
Ryan Zarychanski, Associate Professor, PhD, MD
University of Manitoba, Winnipeg, Canada

ORIGINAL PAPERS

The PhD thesis is based on the following papers:

I. **Pharmacological interventions for prevention and management of delirium in intensive care patients: a systematic overview of reviews and meta-analyses.** Barbateskovic M, Krauss SR, Collet MO, Larsen LK, Jakobsen JC, Perner A, Wetterslev J. BMJ Open 2019;9(2):e024562.

II. **Haloperidol for the treatment of delirium in critically ill patients: a systematic review with meta-analysis and Trial Sequential Analysis.** Barbateskovic M, Krauss SR, Collet MO, Andersen-Ranberg NC, Mathiesen O, Jakobsen JC, Perner A, Wetterslev J. Acta Anaesthesiol Scand 2019. doi: 10.1111/aas.13501. [Epub ahead of print]

III. **Stress ulcer prophylaxis with proton pump inhibitors or histamin-2 receptor antagonists in adult intensive care patients: a systematic review with meta-analysis and Trial Sequential Analysis.** Barbateskovic M, Marker S, Granholm A, Anthon CT, Krag M, Jakobsen JC, Perner A, Wetterslev J, Møller MH. Intensive Care Med 2019; 45(2):143-158.

IV. **Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the intensive care unit.** Barbateskovic M, Schjørring OL, Krauss SR, Jakobsen JC, Meyhoff CS, Dahl RM, Rasmussen BS, Perner A, Wetterslev J. Cochrane Database of Systematic Reviews 2019, Issue 11. Art. No.: CD012631. DOI: 10.1002/14651858.CD012631.pub2

V. **Higher versus lower levels of oxygen supplementation in critically ill adults. A systematic review with meta-analysis and Trial Sequential Analysis**. Barbateskovic M, Schjørring OL, Krauss SR, Jakobsen JC, Meyhoff CS, Rasmussen BS, Perner A, Wetterslev J. [In review]

VI. **CHIMS: Clinical Heterogeneity In Meta-Analysis Score - a new tool to assess and quantify clinical heterogeneity in meta-analyses of interventions.** Barbateskovic M, Koster TM, Eck RJ, Maagaard M, Afshari A, Blokzijl F, Cronhjort M, Dieperink W, Fabritius ML, Feinberg J, French C, Gareb B, Geisler A, Granholm A, Hiemstra B, Hu R, Imberger G, Jensen BT, Jonsson AB, Karam O, Kong DZ, Korang SK, Koster G, Lai B, Liang N, Lundstrøm LH, Marker S, Meyhoff T, Nielsen EE, Nørskov AK, Petersen MW, Risom EC, Rygård SL, Safi S, Sethi N, Sjövall F, Lauridsen SV, van Bakelen N, Volbeda M, van der Horst ICC, Gluud C, Perner A, Møller MH, Keus F, Wetterslev J. [In review]

TABLE OF CONTENTS

1. PREFACE

The work presented in this PhD thesis was conducted during my employment as a PhD student at the Copenhagen Trial Unit (CTU), Rigshospitalet, from May 2015 to December 2019, in collaboration with the Centre for Research in Intensive Care (CRIC).

Several persons have contributed to making this PhD study possible. First and foremost, I would like to thank my supervisors Anders Perner, Jørn Wetterslev and Janus Christian Jakobsen for granting me the opportunity to work on the CRIC projects. Thank you all for assisting with your valuable methodological and clinical expertise. A special thanks to Jørn for your extensive support; your door was always open to me, whether for explanations and guidance regarding complex methodology or pleasant talks about politics, physics and philosophy. I have learnt so much from you, especially the importance of being thorough and the value that comes from including skilled people in research projects, as well as how to navigate in the field of research. Although you are now leaving CTU, I sincerely hope that we will continue our cooperation in the future.

I also wish to thank the collaborators on all my studies; your valuable and important clinical expertise in each CRIC programme was vital to the study. I am especially thankful to Olav Lilleholt Schjørring for your valuable insight into the oxygen literature, which has definitely made our reviews better. To all collaborators who have spent many hours on literature screening, data extraction and risk of bias evaluation – without you these projects could not have been realised.

I also wish to thank Eric Keus and Iwan van der Horst for a pleasant and rewarding research stay at the University Medical Center Groningen in January 2019. Thanks to all who contributed to the development of CHIMS; a special thanks to Thijs Koster, Ruben Eck and Mathias Maagaard for their contribution to the project, and to all who contributed to the project by testing the CHIMS tool.

The CTU has been my foundation during this study. I have been fortunate to work together with many talented people, who all have contributed with academic discussions, help with figures, software, administrative support, as well as with interesting conversations in the CTU kitchen – and especially with support during difficult times; Jane Lindschou, Ema Erkocevic Petersen, Janus Engstrøm, Mette Hansen, Snezana Djurisic, Sara Russo Krauss, Christian Gluud, Sara Klingenberg and Dimitrinka Nikolova. Also, thanks to Heidi Graff, Marie Oxenbøll and Laura Larsen for sharing the PhD journey with me.

I would also like to thank *Innovation Fund Denmark* for granting the financial support for my PhD study, to *Rigshospitalets Forskningspulje* for granting me an additional six months of salary and to CTU for financial support during my maternity leave.

Finally, my deepest gratitude goes to my family. Especially to my husband Patrick for supporting me through the PhD process, for listening, offering advice, proof reading and for allowing me to work long nights for many months while taking care of our children.

Marija Barbateskovic, December 2019

2. LIST OF ABBREVIATIONS

Alpha = α	Risk of type 1 error
BIOSIS-Previews	BioSciences Information Service-Previous
CHIMS	Clinical Heterogeneity In A Meta-analysis Score
CINAHL	Cumulative Index to Nursing & Allied Health Literature
CI	Confidence interval
C. difficile infection	Clostridium difficile infection
CRIC	Centre for Research in Intensive Care
CTU	Copenhagen Trial Unit
D^2	Diversity
Embase	Excerpta Medica Database
FiO_2	Fraction of inspired oxygen
GRADE	Grading of Recommendations Assessment, Development and Evaluation
I^2	Inconsistency factor
ICC	Intraclass correlation coefficients
ICU	Intensive care unit
LILACS	Latin American Caribbean Health Sciences Literature
MEDLINE	Medical literature Analysis and Retrieval System Online
QTc prolongation	QT interval corrected for heart rate
PaO_2	arterial oxygen tension
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-analysis
RR	Relative risk
SaO_2	Arterial oxygen saturation
SAPS II	Simplified acute physiology score II
SpO_2	Peripheral oxygen saturation
TSA	Trial Sequential Analysis

3. SUMMARY

Background

To manage and preserve vital functions in the critically ill patient, adjuvant interventions, also described as intervention modalities, are applied to support organ dysfunction in addition to core therapy targeting the underlying disease. Examples of adjuvant interventions are oxygen supplementation, fluids, dialysis, analgesia, sedation, muscle relaxation, stress ulcer prophylaxis and delirium management. Characteristically, these adjuvant interventions have each been approved for specific indications in the non-intensive care unit setting and their use has subsequently been adapted in the intensive care unit (ICU) setting. The overall aim of this PhD study was to assess the evidence of important and frequently administered adjuvant interventions used in intensive care; haloperidol for the management of delirium, stress ulcer prophylaxis, and oxygen supplementation.

Haloperidol for the management of delirium – Study I and II

To assess the evidence of pharmacological interventions for delirium, we conducted a systematic overview of reviews of the pharmacological prevention and management of delirium in the ICU setting, as well as a systematic review of the use of haloperidol for managing (treating) delirium in the critical care setting.

Stress ulcer prophylaxis – Study III

To assess the evidence of stress ulcer prophylaxis in the ICU setting, we conducted a systematic review assessing the effects of proton pump inhibitors or histamin-2 receptor antagonists.

Oxygen supplementation – Study IV and V

To assess the evidence of the effect of applying higher versus lower levels of oxygen supplementation in patients admitted to the ICU and in critically ill patients, we conducted two systematic reviews; the first systematic review focused exclusively on ICU patients, the second one included all critically ill patients, defined as all patients who are at high risk of dying or who have actual or potential life-threatening health problems irrespective of setting.

Clinical heterogeneity in meta-analyses – Study VI

To increase our knowledge on clinical heterogeneity in meta-analyses, and to support our view that no single method exists to assess clinical heterogeneity in meta-analyses, we developed the Clinical Heterogeneity In Meta-analysis Score (CHIMS) and tested the interrater scale reliability and agreement.

Methods

Systematic overview – Study I

The systematic overview of reviews included reviews identified by searching major international medical databases. Two authors independently screened the retrieved titles for inclusion, extracted data and assessed risk of bias. We categorised reviews reported according to Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) as systematic and assessed risk of bias using ROBIS (a tool to assess risk of bias in systematic reviews).

Systematic reviews – Study II-V

The four systematic reviews of interventions (Study II-V) included randomised clinical trials identified by searching major medical international databases, including trial registers. Two authors independently screened the retrieved titles for inclusion, extracted data and assessed risk of bias. We meta-analysed the data and used Trial Sequential Analysis (TSA) to calculate the meta-analytic required information size,

considering risk of random errors due to sparse data, multiple outcomes, and multiple testing on accumulating data. We appraised the certainty of evidence and our confidence in the effect estimates using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach.

Clinical heterogeneity in meta-analyses – Study VI
The items and domains of CHIMS, is built upon a consensus work by Gagnier et al. We drafted the CHIMS tool, reviewed and revised it three times, and pilot-rated meta-analyses in two rounds. We tested CHIMS for interrater scale reliability and agreement in three groups.

Results
Haloperidol for the management of delirium – Study I and II
We found 378 reviews on pharmacological prevention and management of delirium. Only one systematic review on delirium prevention was found to be systematic according to PRISMA and no systematic review investigating management of manifest delirium was found.

Meta-analyses showed no evidence of a difference on mortality, delirium severity, QTc prolongation, delirium resolution, extrapyramidal symptoms (very low certainty evidence), when comparing haloperidol with any intervention. Only one trial reported on days alive without delirium and cognitive function. No trials reported adequately on serious adverse events and no trials reported on quality of life.

Stress ulcer prophylaxis – Study III
We did not find evidence of a difference on mortality (high certainty of evidence), although gastrointestinal bleeding was reduced with stress ulcer prophylaxis (high certainty of evidence). Only approximately 5% of patients receiving placebo or no prophylaxis had a bleeding episode. Meta-analyses did not find evidence of a difference on serious adverse events, myocardial ischemia, pneumonia, clostridium (C) difficile infection (moderate to very low certainty of evidence), when comparing the use of stress ulcer prophylaxis with placebo or no prophylaxis. No trials reported on quality of life.

Oxygen supplementation – Study IV and V
We found an increased risk on mortality in the traditional meta-analysis; however, TSA indicated that evidence is insufficient to confirm or refute a 20% relative change in mortality, when comparing higher versus lower oxygen supplementation (very low certainty of evidence) in ICU patients. Meta-analysis and TSA indicated an increase in serious adverse events (very low certainty of evidence). Due to insufficient data, the effects on lung injury and sepsis were inconclusive (very low certainty of evidence). No trials reported on quality of life, acute myocardial infarction, or stroke.

We did not find evidence of a difference on mortality (low certainty of evidence) in critically ill patients. A subgroup analysis of trials conducted in the ICU setting did not show evidence of a difference on mortality.
Meta-analyses showed no evidence of a difference on serious adverse events, quality of life, lung injury, sepsis, or cardiovascular events.

Clinical heterogeneity in meta-analyses – Study VI
We developed CHIMS that covers four domains (setting, population, intervention, and outcome) with a total of 11 items. Each item is scored 0, 1 or 2, and are accumulated to a minimum summary score of 0 points or a maximum summary score of 22 points. Results of the reliability tests of CHIMS found it to be a reliable tool for assessing and quantifying clinical heterogeneity in meta-analyses.

Conclusion
Overall, the evidence was sparse or even absent, of low quality, with insufficient information sizes within the adjuvant interventions assessed; haloperidol for the management of delirium, stress ulcer prophylaxis, and oxygen supplementation for patients admitted to the ICU.

The evidence for the use of haloperidol to manage delirium in ICU patients and in critically ill patients is sparse, of low quality and inconclusive, due to very sparse data.

There is high certainty of evidence that stress ulcer prophylaxis does not reduce mortality, although there is high certainty of evidence that gastrointestinal bleeding and clinically important bleeding is reduced with stress ulcer prophylaxis in ICU patients. It is worth noting, however, that approximately 95% of patients in the ICU will not experience a bleeding period, when stress ulcer prophylaxis is not administered. This may indicate overutilization of medication. In relation to this, one trial indicated excess mortality when using pantoprazole for the most severely ill patients, with Simplified Acute Physiology Score (SAPS) II score greater than 53.

With low to very low certainty of evidence, it is inconclusive whether higher versus lower fraction of inspired oxygen or targets of arterial oxygenation have beneficial or harmful effects in ICU patients and in critically ill patients.

Patients admitted to the ICU are clinically heterogeneous. This heterogeneity of the population, together with a general high clinical heterogeneity between the trials conducted in this setting, may challenge the interpretation of randomised clinical trials and meta-analyses.

There is a need to map out the evidence for treatment modalities used in the intensive care setting to help identify knowledge gaps and to plan future research.

4. SUMMARY IN DANISH (DANSK RESUMÉ)

Baggrund

Behandling og opretholdelse af vitale funktioner hos den kritisk syge patient, udføres ved hjælp af adjuverende interventioner, også beskrevet som interventionsmetoder anvendt til at understøtte organdysfunktion, i tillæg til kerneterapi rettet mod den underliggende sygdom. Eksempler på adjuverende interventioner er ilttilskud, væske, dialyse, analgesi, sedation, muskelafslapning, stress ulcus profylakse og håndtering af delirium. Karakteristisk for disse adjuverende interventioner er, at de hver især er godkendt til specifikke indikationer uden for intensivregi, hvorefter de er adapteret til forhold i intensivterapi. Det overordnede formål med dette PhD projekt var at vurdere evidensen for brugen af livsvigtige og ofte administrerede adjuverende interventioner i intensivregi; haloperidol til håndtering af delirium, stress ulcus profylakse og ilttilskud. Dette blev opnået ved at udarbejde systematiske litteraturoversigter (engelsk: systematic reviews) med metaanalyser og forsøgssekventielle analyser (engelsk: Trial Sequential Analysis).

Haloperidol til håndtering af delirium – Studie I og II

For at vurdere evidensen for brugen af farmakologiske interventioner til forebyggelse og håndtering af delirium, udarbejdede vi en systematisk litteraturgennemgang af oversigtsartikler af farmakologisk forebyggelse og håndtering (behandling) af delirium i intensivterapi samt en systematisk litteraturgennemgang af haloperidol til håndtering af delirium i intensivterapi.

Stress ulcus profylakse – Studie III

For at vurdere evidensen for brugen af stress ulcus profylakse i intensivterapi, udarbejdede vi en systematisk litteraturgennemgang af effekten af protonpumpehæmmere eller histamin-2-receptorantagonister.

Ilttilskud – Studie IV og V

For at vurdere evidensen for brugen af højere versus lavere niveauer af ilttilskud, foretog vi to undersøgelser; først en systematisk litteraturgennemgang der fokuserer på intensivterapi og dernæst en systematisk litteraturgennemgang af samme intervention hos alle kritisk syge patienter uanset indlæggelsessted, defineret som alle patienter, der er i høj risiko for at dø, eller som har faktiske eller potentielle livstruende helbredsproblemer.

Klinisk heterogenitet i metaanalyser – Studie VI

For at øge vores viden om klinisk heterogenitet i metaanalyser og afhjælpe at der ikke findes en enkelt metode til vurdering af klinisk heterogenitet i metaanalyser, udviklede vi Clinical Heterogeneity In a Meta-analysis Score (CHIMS) og testede skalaens inter-observatør reliabilitet og overensstemmelse.

Metode

Systematisk litteraturgennemgang af oversigtsartikler – Studie I

Den systematiske litteraturgennemgang af oversigtsartikler inkluderede oversigtsartikler identificeret fra søgning i store internationale medicinske databaser. Vi kategoriserede oversigterne i henhold til Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) – et sæt af 27 emner som systematiske litteraturoversigter anbefales at rapportere i henhold til. Vi vurderede risiko for bias ved hjælp af ROBIS (et værktøj til vurdering af risiko for bias i systematiske oversigtsartikler).

Systematiske oversigter – Studie II-V

De fire systematiske oversigter af interventioner inkluderede randomiserede kliniske forsøg, der blev identificeret ved søgning i store internationale medicinske databaser, herunder forsøgsregistre. To forfattere

screenede uafhængigt af hinanden søgeresultaterne, ekstraherede data og vurderede risikoen for bias. Vi metaanalyserede data og brugte forsøgssekventiel analyse til at beregne den informationsstørrelse metaanalyser kræver, under hensyntagen til risikoen for tilfældige fejl på grund af sparsomme data, flere effektmål og flere test på akkumulerende data. Vi anvendte Grading of Recommendations Assessment, Development and Evaluation (GRADE) til vurdering og sammenfatning af evidensens kvalitet og styrke.

Klinisk heterogenitet i metaanalyser – Studie VI
Elementer og domæner i CHIMS, bygger på et konsensusarbejde af Gagnier et al. Vi udarbejdede CHIMS-værktøjet, gennemgik og reviderede det tre gange, og pilotbedømte metaanalyser i to runder, hvorefter vi testede CHIMS for inter-observatør reliabilitet og overensstemmelse i tre grupper.

Resultater
Haloperidol til håndtering af delirium – Studie I og II
Vi fandt 378 oversigtsartikler om farmakologisk forebyggelse og håndtering af manifest delirium. Kun en oversigtsartikel om deliriumforebyggelse blev vurderet til at være systematisk ifølge PRISMA. Der blev ikke fundet systematiske oversigtsartikler der undersøgte håndtering af manifest delirium.

Metaanalyse fandt ingen forskel på mortalitet, delirium sværhedsgrad, QTc-forlængelse, remission af delirium, ekstrapyramidale symptomer (evidens af meget lav sikkerhed) ved sammenligning af haloperidol med kontrolinterventioner hos kritisk syge patienter. Kun et forsøg rapporterede på 'dage i live uden delirium' og 'kognitiv funktion'. Ingen forsøg rapporterede alvorlige bivirkninger tilstrækkeligt og ingen forsøg rapporterede på livskvalitet.

Stress ulcus profylakse – Studie II
Vi fandt ingen forskel på dødelighed (evidens af høj sikkerhed) i stress ulcus profylakse gruppen sammenlignet med kontrolgruppen, på trods af at gastrointestinal blødning blev reduceret med stress ulcus profylakse (evidens af høj sikkerhed). Kun ca. 5% af patienterne der fik placebo eller ingen profylakse havde en blødningsepisode. I Metaanalyserne fandt vi ingen forskel på alvorlige bivirkninger, myokardie iskæmi, lungebetændelse, C. difficile infection (evidens af meget lav sikkerhed), ved brugen af stress ulcus profylakse versus placebo eller ingen brug af profylakse. Ingen forsøg rapporterede på livskvalitet.

Ilttilskudt – Studie IV og V
Vi fandt øget risiko for død i den traditionelle metaanalyse, men den forsøgssekventielle analyse indikerede at der var utilstrækkeligt med data til at bekræfte eller afkræfte en relativ ændring på 20% ved sammenligning af højere versus lavere niveauer af supplerende ilt i intensivterapi (evidens af meget lav sikkerhed). Metaanalyse og forsøgssekventiel analyse indikerede en stigning i alvorlige bivirkninger (evidens af meget lav sikkerhed). På grund af utilstrækkelige data er virkningerne på lungeskade og sepsis inkonklusive (evidens af meget lav sikkerhed). Ingen forsøg rapporterede på livskvalitet, akut myokardieinfarkt eller apopleksi.

Vi fandt ingen forskel på mortalitet (evidens af lav sikkerhed) hos kritisk syge patienter. En subgruppeanalyse af forsøg udført i intensivregi fandt ingen forskel på mortalitet. Metaanalyse fandt ingen forskel på alvorlige bivirkninger, livskvalitet, lungeskader, sepsis og kardiovaskulære tilfælde.

Klinisk heterogenitet i metaanalyser – Studie VI
Vi udviklede CHIMS som dækker over fire domæner (miljø, population, intervention og effektmål) med i alt 11 elementer. Hvert element scores 0, 1 eller 2, som derefter akkumuleres til en sammenfattende score på minimum 0 point og maksimum 22 point. Resultaterne af reliabilitetsanalyserne fandt at CHIMS er et pålideligt værktøj til vurdering og kvantificering af klinisk heterogenitet i metaanalyser.

Konklusion

Overordnet set var evidensen sparsom eller endda fraværende, af lav sikkerhed, med utilstrækkelige informationsstørrelser inden for de vurderede adjuverende interventione: haloperidol til håndtering af delirium, stress ulcus profylakse og supplerende ilt til patienter indlagt på intensivafdeling.

Evidensen for brug af haloperidol til håndtering af delirium hos intensivpatienter og kritisk syge patienter er sparsom, af lav kvalitet og inkonklusiv på grund af sparsomme data.

Der er høj evidens for at stress ulcus profylakse ikke reducerer dødeligheden, selvom der er høj evidens for at gastrointestinal blødning og klinisk vigtig gastrointestinal blødning reduceres med stress ulcus profylakse hos intensivpatienter. Det er imidlertid værd at bemærke, at ca. 95% af patienterne i intensivregi ikke oplever en blødningsperiode, når de ikke behandles med stress ulcus profylakse. Dette kan indikere overmedicinering. Det er i denne forbindelse også værd at bemærke, at et forsøg har indikeret øget mortalitet hos de mest alvorligt syge patienter, med Simplified Acute Physiology Score (SAPS) II score større end 53, når de får pantoprazole.

Med lav til meget lav sikkerhed for evidensen, er det inkonklusivt hvorvidt højere versus lavere niveau af ilttilskud har gavnlige eller skadelige virkninger hos intensivpatienter og kritisk syge patienter.

Patienter indlagt på intensivafdeling er klinisk heterogene. Denne heterogenitet i populationen, sammenholdt med en generel høj klinisk heterogenitet mellem forsøgene udført i intensivafdelinger og hos kritisk syge, kan vanskeliggøre tolkning af randomiserede kliniske forsøg og metaanalyser.

Der er behov for at kortlægge evidensen for behandlingsmetoder anvendt i intensivmedicin, for at bistå med at identificere lakunerne i tilgængelig viden og til planlægning af fremtidig forskning.

5. BACKGROUND

5.1. Evidence-based medicine

Evidence-based medicine is the conscious, explicit and judicious use of the current best evidence, when making decisions about the health care of patients.[1] It is practiced by combining clinical expertise with the best available clinical evidence from systematic research. More specifically, it involves the formulation of a clear clinical patient-centred question, the subsequent search through the literature to identify relevant papers which can be used to answer the question asked, the evaluation of the evidence for its validity by critically appraising the results, and the implementation of the findings in clinical practice.[2, 3] However, identifying the best research evidence can be difficult. In general, results from observational studies, non-randomised trials or single randomised trials should not form the basis for using interventions in evidence-based medicine. Observational studies may provide information about risk factors and provide information on the prevalence on e.g. the use of a specific intervention or the incidence of participants developing a specific disease; therefore, observational studies may serve as the base for randomised clinical trials. Rather, evidence based medicine should be based on systematic reviews with meta-analysis of randomised trials with Grading of Recommendations Assessment, Development and Evaluation (GRADE), which is regarded as the highest level of evidence.[4] In short, in evidence-based medicine we need sufficient, real data, collected in an unbiased way, addressing patient centered and important outcomes. Several potential errors should be avoided, such as systematic errors, random errors, and non-patient centered outcomes, in order to obtain the best evidence. However, until the best obtainable evidence is present, we must try to filter the best available evidence by appraising risk of bias, considering insufficient information, and lack of adequate or patient centered outcomes.

The systematic review process involves several steps, where decisions are to be taken during the process, which may impact the conclusion and implications.

5.2. Systematic reviews

Systematic reviews of randomised clinical trials are considered the most valid way of examining the benefits and harms of interventions, and to summarise the best available evidence.[4]

Cochrane reviews are considered the gold standard of systematic reviews. The methods of such reviews are pre-published in a protocol, where the following core methods are described and defined as: scope and aim, inclusion criteria, method of literature search, data collection, risk of bias assessment, meta-analyses, and summary of findings.

5.3. Systematic errors in randomised trials

The methodological quality of randomised clinical trials included in the systematic review may impact the effect estimates of the interventions, which may alter the results of the meta-analyses and conclusion of the review.[5] Randomised clinical trials with inadequate methods may be associated with bias and tend to exaggerate the effect of interventions.[5-8] Bias, also described as systematic errors, may ultimately mislead health-care decision making, if not accounted for. Therefore, evaluation of the risk of bias in the included trials is an essential component of a systematic review. Assessment of the risk of bias in each included trial is performed using the Cochrane risk of bias tool. Until recently, Cochrane advocated the use of the Cochrane risk of bias tool originally published in the Cochrane Handbook in 2008 and updated in 2011. The risk of bias tool for randomised trials includes the following domains, with empirical evidence supporting an association with systematic error: generation of the allocation sequence, allocation concealment, blinding of participants and personnel, blinding of outcome assessor, incomplete outcome data, selective outcome reporting, and other sources of bias, e.g. vested interests.[9, 10] Risk of bias on each domain is assessed as low, unclear or high. The domains of sequence generation, allocation concealment and selective outcome reporting does not vary between outcomes for a trial, whilst blinding of participants and personnel, blinding of outcome assessment and incomplete outcome data may differ between outcomes in a trial, and should therefore be assessed separately for each outcome. Cochrane has recently updated the Cochrane risk-of bias-tool (RoB2)[11] – differences between the two tools are summarised in table 1.

Table 1. Differences between Cochranes risk of bias tool 1 and 2

RoB 1	RoB 2
Random sequence generation (selection bias)	Bias arising from the randomisation process
Allocation concealment (selection bias)	
Blinding of participants and personnel (performance bias)	Bias due to deviations from intended interventions
Blinding of outcome assessor (detection bias)	Bias in measurement of the outcome
Incomplete outcome data (attrition bias)	Bias due to missing outcome data
Selective reporting (reporting bias)	Bias in selection of the reported result
Other	N/A
N/A	Overall bias

5.4. Meta-analysis

Meta-analysis is the statistical procedure of combining the results from two or more trials. The advantage of the meta-analysis is improvement in precision; however, it may also mislead seriously if bias, heterogeneity (variation across trials) and study design are not considered carefully. The meta-analysis result itself is a precision weighted average of the effect estimates from the included trials. The weighting is based on the inverse variance within each trial in the fixed effect model and the variance in each trial added to the between trial variance in the random effects model. During the process of preparing the meta-analysis, many decisions are made including choice of meta-analytic model(s) and (1-α)Confidence Interval (CI) for the pooled intervention effect (e.g. 95% CI), which may influence the results and thus have an impact on the conclusion.

5.5. Clinical heterogeneity in meta-analyses

Heterogeneity in meta-analyses refers to any kind of variability across trials. In a meta-analysis, no two trials will be completely identical. Therefore, systematic reviews include assessment of the variability across trials in order to judge whether meta-analysis make sense. It is crucial to assess the presence of heterogeneity when conducting meta-analyses, as the heterogeneity may affect the interpretation of the results, which may further affect the generalisability of the conclusion.

Different types of heterogeneity across the included trials in systematic reviews may occur. Clinical heterogeneity can be characterised as variability in settings, participants, characteristics of interventions and comparators, use of co-interventions, and the types and timing of outcome assessments.[12, 13] Methodological heterogeneity is characterised by variability in trial design and risk of bias, whilst statistical heterogeneity is characterised by variability in treatment effects between trials.[14]

Methodological heterogeneity is evaluated by assessing risk of bias in and across the included trials.[9] Statistical heterogeneity is assessed by visual inspection of the forest plots and by the Chi^2-statictic and I^2 statistic.[14] Investigation of clinical heterogeneity may be performed by subgroup analyses and meta-regression, which may be used to explore whether intervention effect is associated with different populations or intervention characteristic at trial level (ecological association), such as dose or duration.[14, 15] Although subgroup analyses and meta-regression analyses may detect differences in treatment effect size associated with trial characteristics, this is not conducted consistently,[16] and the overall clinical heterogeneity is usually not assessed or quantified, except for some general statements of clinical or trial heterogeneity being abundant or pronounced, etc.

5.6. Assessment of random errors in meta-analyses

As mentioned previously, systematic reviews with meta-analysis of randomised clinical trials are considered the best available evidence.[4, 17] However, the best available evidence may not be equal to sufficient or best obtainable evidence.[18-21] Most Cochrane reviews have been shown to be underpowered; Turner and colleagues showed that only 22% of 14,886 meta-analyses from Cochrane reviews had 80% power to detect or refute a 30% relative risk reduction.[22] About 50% of meta-analyses within critical care suggesting a beneficial effect have been shown to fulfil criteria for being true positive due to a too low amount of information, and 29% of meta-analyses suggesting a harmful effect have been shown to fulfil criteria for being true harmful due to a too low amount of information.[23] Furthermore, large pooled intervention effects observed in the "early positive meta-analyses" tend to vanish.[20, 24] The risk of type 1 error increases along repeated unadjusted confidence intervals (CIs) and significance testing on accumulation data, e.g. when meta-analyses are updated.[25] Thus, a meta-analysis should include a large information size (= sum of included trials sample sizes) adjusted for heterogeneity, if a random effects meta-analysis is applied.

Trial Sequential Analysis (TSA) can be applied to assess the risk of random errors and a meta-analytic required information size can be calculated.[26] The required information size accommodate the unweighted cumulated event proportion in the control group, the assumption or anticipation of a plausible relative risk reduction or relative risk increase, the heterogeneity variance of the meta-analysis, and the chosen risks of type 1 and type 2 errors.[26, 27]

Trial sequential monitoring boundaries can be constructed based on the required information size, which enables one to determine the statistical inference concerning cumulative meta-analysis that has not yet reached the required information size.[26, 28-30] Firm evidence for benefits or harms, disregarding risk of bias, may be established if the trial sequential monitoring boundary is crossed before reaching the required information size, which suggest that further trials may be redundant. Conversely, if the boundary is not reached, one may conclude that it is necessary to include additional trials and participants before a certain intervention effect can be detected or rejected.[31] TSA may also assess when a specific anticipated intervention effect can be rejected, which turn up when the cumulative Z-score crosses the trial sequential monitoring boundaries for futility.

5.7. Patient centered outcomes

When designing clinical trials in order to compare the effects of interventions, the selection of appropriate outcomes is pivotal. The outcomes need to be relevant and important to patients and the public, healthcare professionals and others making decisions about heath care.[32] However, patient-important outcomes such as mortality, functional disability and quality of life are rarely primary outcomes in randomised clinical trials of critically ill patients,[33] although critical illness is associated with long-term physical and psychical sequela that may impact functional status and quality of life.[34-36] Furthermore, published trial reports have been shown to underreport adverse events with a median of 64% compared to unpublished trials.[37] In the ICU setting, patients experience numerous serious adverse events, which logistically may be impossible to register. The reporting of serious adverse events in trials of ICU patients is therefore per default underreported.

To improve reporting of harm in systematic reviews, it is recommended to adhere to the PRISMA harms checklist.[38] In addition to the 27 PRISMA items, the harms checklist recommends how to address harm in a systematic review, including definition of title, definition of harms addressed, how to handle data and synthesis of results. Nonetheless, only 38% of registered systematic review protocols include adverse events as outcomes and only 65% of the published reviews report fully on their pre-registered outcomes.[39] It is also recommended to report patient-reported outcomes.[40] The use of self-reported outcomes in the ICU setting may be challenged by patients not being capable of providing estimates and proxy estimates may therefore be used. This challenges the interpretation as estimates on e.g. quality of life obtained from a survivor of critical illness may not be similar to those obtained from their proxy responders.[41, 42]

5.8. Grading the certainty of evidence

GRADE is an approach to grading quality of evidence and strength of potential recommendations in healthcare. After the evidence has been collected and summarised (e.g. with meta-analysis), the quality of evidence for each important outcome is appraised and presented in a summary of findings table.

The GRADE approach appraises the quality of a body of evidence based on the extent to which one can be confident that an estimate of effect or association reflects the item being assessed.
In the GRADE approach, randomised trials start as high-quality evidence supporting estimates of intervention effects. Five factors may lead to downgrading the certainty; the quality measure of a body of evidence considers risk of bias within the trials, the directness of the evidence, heterogeneity of the data, precision of effect estimates, and risk of publication bias.[43] Ultimately, the quality of evidence for each outcome falls into one of four categories, from high, over moderate and low, to very low certainty of evidence.[43]

5.9. Overviews of reviews

Approximately 22 new systematic reviews are published every day.[44] Overviews may be used to filter this information load. It is a discipline within evidence synthesis, which has also progressed to improve faster access to information and may be valuable in healthcare decision-making.

Overviews of reviews are reports that include multiple reviews and/or systematic reviews reporting on a specific topic, that integrates information from multiple related reviews to provide a comprehensive synthesis of all evidence derived from systematic reviews related to a specific research question. They are also known as overviews, (systematic) review of (systematic) reviews, umbrella reviews, meta-reviews, synthesis of systematic reviews and summary of systematic reviews.[45] Overviews are preferred, when the aim is to collect all systematic (and potentially unbiased) information from systematic reviews. They may also

be valuable, when evidence relating to a specific topic exist, but is conflicting, hereby bringing reviews together in a transparent and systematic way, thus supporting decision making by gathering, appraising and systematically analysing the evidence. An essential and important aspect of the overview is how systematic reviews are defined, as this will lay the basis for the overall summarized evidence. Overall, the methodology of overviews is very similar to that of systematic reviews of randomised trials but differ in the way that the overview collects evidence from synthesized evidence. Analyses of outcome data in an overview may either be descriptive (presenting data in the overview exactly as they are presented in the SR) or outcome data can be analysed in a way that differs from analyses conducted in the systematic review.[46]

Overview of reviews is also a discipline within Cochrane, where the most common question addressed by overviews is how to examine the evidence from systematic reviews of different interventions for the same condition or population.[46]

5.10. Network meta-analyses

Another way to get an overview of interventions is via network meta-analyses, which syntheses networks of direct and indirect comparisons of interventions to assess the effect of more than two interventions for the same condition. Estimates from different direct comparisons with a common comparator, that have not been compared directly in a clinical trial, can in a network meta-analysis yield an indirect comparison of these interventions.[47] Although network meta-analyses include the possibility of comparing and ranking all available treatment options in one analysis, it also comes with challenges. The validity of the network meta-analysis relies on the fulfilment of the transitivity assumption – that the included trials are similar in factors that may affect the relative effects.[47, 48] Another challenge is problems with multiplicity due to multiple comparisons and due to the assessment of many outcomes.[48]

5.11. Critical and intensive care

Intensive care, also designated critical care, is a multidisciplinary and interprofessional specialty devoted to the management of patients having, or at risk of developing, acute, life-threatening organ dysfunction.[49] A variety of technologies providing support for a failing organ system, particularly the lungs, cardiovascular system, immunological system, and kidneys, are used. The central point of expertise is the pathophysiological support for organ dysfunction and the primary goal of intensive care is to prevent further physiologic deterioration while the underlying disease is treated and resolves.[49] Compared to a ward based care, where patients do not require organ support (but may need intravenous fluids/medicine or oxygen by face mask), and a high dependency unit where patients need single organ support such as the need for renal replacement therapy, inotropes or invasive blood pressure monitoring, the ICU admit patients requiring that two or more organ are supported or patients who are in need of mechanical ventilation. Another difference is the staff-to-patient ratio in the ICU, which is high due to close monitoring and multiple additional interventions. The ICU has been defined as an organized system for the provision of care to critically ill patients, that provides intensive and specialized medical and nursing care, has an enhanced capacity for monitoring, and has multiple modalities of physiologic organ support to sustain life during a period of acute organ system insufficiency.[49] There is wide international variation because access to resources is highly variable – especially taking into account the considerable international variation in capacity to provide health care. The amount of ICU beds per country differs considerably, with less beds in developing countries.[50] Furthermore, in the developing world, ICU beds are almost exclusively found in large referral hospitals where diagnostic blood tests may be unavailable and microbiological investigations are rare.[51-53]

5.12. Management of the critically ill and intensive care patient

Admission to the ICU may be planned or be an emergency admission. Mortality rates are high and survival to discharge vary depending on the reason for admission and physiological impairments. Patients are admitted to the ICU for various reasons. Common indications for ICU admission include hypotension unresponsive to fluid resuscitation (e.g. sepsis), myocardial infarction, cardiac arrest, requirement for advanced respiratory support (e.g. severe asthma, chronic obstructive pulmonary disease exacerbation), requirement for sedation, head injury, severe liver disease, advanced post-operative monitoring due to comorbidities or severity of surgery, and requirement for renal support.[54]

Initial resuscitation of the critically ill patient follows an ABCDE approach: airway, breathing, circulation, disability, and exposure. The initial treatment of critical illness is to support the patient's vital functions by respiratory monitoring and support, circulatory support, neurological monitoring and support, and renal support – which may, at least initially, take priority over establishing a precise diagnosis.[55] Using life threatening shock as an example, patients need immediate treatment rather than diagnosis of the cause, as the principles of management may require some of the same adjuvant interventions, whether the shock results from myocardial infarction or a gastrointestinal bleed. Likewise, although the actual management may differ, the principles of handling other life-threatening organ failures, e.g. respiratory failure, do not depend on a precise initial diagnosis.[55] However, it will always be important to initiate exact diagnosing to reach a precise diagnosis as soon as possible, to finetune the interventions used for the specific condition in the patient.

Post initial period, during the period of recover, medical treatment of the underlying disease is continued as well as the organ support, to maintain physiology in as normal a state as possible, allowing time for the actual treatment of the underlying condition to work.

5.13. Adjuvant interventions

Management and preservation of the vital functions in the critically ill patient is done using adjuvant interventions, also described as intervention modalities applied to support organ dysfunction, in addition to core therapy targeting the underlying disease (e.g. antibiotics for primary infection or percutaneous coronary intervention (PCI) and coronary artery bypass graft (CABG) in acute coronary syndrome or surgery for abdominal compartment syndrome). Examples of adjuvant interventions are oxygen supplementation, fluids, dialysis, analgesia, sedation, muscle relaxation, stress ulcer prophylaxis and management of delirium. In the intensive care setting, the adjuvant interventions are not applied according to etiologic diagnosis, but to some degree according to the main patient group, e.g. trauma, cardiac arrest, stroke, or acquired syndrome, e.g. sepsis, delirium. The adjuvant interventions are therefore applied regardless of the underlying disease or syndrome and are thus applied due to relatively wide indications. Characteristically, these adjuvant interventions have each been approved for specific indications in the non-ICU setting and their use has subsequently been adopted in the ICU setting.

Safe and effective medical therapy is crucial to ensure optimal patient care and outcome. However, acute organ dysfunction during critical illness may affect the intervention effects in the ICU patients, so data generated in settings outside the ICU may not be directly or unaltered transposed to the ICU, where the patients stage of critical or severe illness is definitely higher and the duration of the use of the medication is usually short. In addition, polypharmacy, which is common during ICU admission and which is associated with medication-associated adverse events, may impact the effectiveness of the applied interventions.[56] Nonetheless, inadequate data to guide prescribing decisions in this setting has resulted in the "off-label" use of nearly 50% of the medications prescribed in the ICU.[56, 57]

An example is the treatment of delirium, which is very frequent in ICU patients, and is often managed with pharmacological drugs of which haloperidol is mostly used, although it is not approved for this indication.[58] Another example is the application of stress ulcer prophylaxis with proton pump inhibitors, which is administered to a majority of the ICU patients (73%[59]) to prevent gastrointestinal bleeding.[59] In addition, oxygen supplementation therapy is applied to prevent or treat hypoxaemia. Although oxygen supplementation beyond 21% (more than air) is regarded as an interventional medicinal product,[60] no formal approval applies and its use in the ICU settings is based on descriptive studies and small randomised trials.[61-68]

The effect of adjuvant interventions in the ICU settings should be evaluated as thoroughly as each of the core interventions, and high-quality evidence should form the basis for its use.

5.14. Heterogeneity in the intensive care setting

Designing trials in the critical care setting is challenging due to the heterogeneous patient population in the ICU, since patients are admitted with, or develop, several clinical disorders (i.e. sepsis, acute respiratory distress syndrome, multi-trauma). Moreover, although adjuvant interventions are used in most of the patients, there is a wide variability of the use across patients in the ICU in e.g. sedation, nutrition, transfusion strategies and fluid balance. Hence, in order to detect even moderate differences in survival, trials require a large sample size.[69] Different barriers that may be encountered when performing randomised trials in the ICU setting, such as maintaining blinding and encountering definite heterogeneity in the population, has led some opinion leaders to question the randomised trial design in the ICU setting.[70-72]

6. AIMS AND HYPOTHESES

The overall aim of this PhD study was to assess benefits and harms of important and frequently administered adjuvant interventions, haloperidol for the management of delirium, stress ulcer prophylaxis, and oxygen supplementation to patients admitted to the ICU, by performing systematic reviews with meta-analyses and TSA.

We hypothesised that evidence was sparse, of low quality, with insufficient information sizes, or even absent within the adjuvant interventions: haloperidol for the management of delirium, stress ulcer prophylaxis, and oxygen supplementation in patients admitted to the ICU. The specific aims and hypotheses of the conducted sub-studies were:

Haloperidol for the management of delirium
Study I: to systematically and critically assess the quantity and quality of the available reviews and meta-analyses of randomised clinical trials on the effects of pharmacological prevention and management of delirium in ICU patients. We hypothesised to find a high number of reviews with heterogeneous quality.

Study II: to assess the benefits and harms of haloperidol versus placebo or any intervention for the treatment of delirium in critically ill patients. We hypothesised an increase in mortality, serious adverse reactions/events and QTc prolongation; a reduction in delirium duration and severity; and a beneficial effect on quality of life and cognitive status.

Stress ulcer prophylaxis
Study III: to assess the benefits and harms of stress ulcer prophylaxis with proton pump inhibitors or histamin-2 receptor antagonists versus placebo or no prophylaxis in adult ICU patients. We hypothesised an absence of effect on mortality, a reduction of GI bleeding, and an increase of infectious adverse events and myocardial ischemia.

Oxygen supplementation
Study IV: to assess the benefits and harms of higher versus lower inspiratory oxygen fraction or targets of arterial oxygenation in adults in ICUs. We hypothesised a decrease in mortality, serious adverse events, quality of life, lung injury, acute myocardial infarction, stroke, and sepsis.

Study V: to assess the benefits and harms of higher versus lower inspiratory oxygen fractions or targets of arterial oxygenation in critically ill adult patients. We hypothesised a decrease in mortality, serious adverse events, quality of life, lung injury, sepsis, and cardiovascular events.

Clinical heterogeneity in meta-analyses
Study VI: to develop a tool for assessing and quantifying clinical heterogeneity in meta-analyses of interventions, and to test the reliability of the tool. In a supplementary exploratory analysis, we aimed to estimate the association, if any, between clinical and statistical heterogeneity. We hypothesised that we would be able to develop a tool aiming to assess and quantify clinical heterogeneity in meta-analyses with high interrater scale reliability and agreement.

7. METHODS

This section briefly describes the methods for each of the papers. Detailed descriptions of the methods for each of the papers are reported in the published protocols.[73-78]

7.1. Study I – Systematic overview of reviews

Study I is a systematic overview of reviews and meta-analyses performed according to the pre-published protocol.[73]

Eligibility criteria

We included all reviews and meta-analyses of pharmacological interventions for the prevention and management (treatment) of delirium in patients admitted to the ICU. We defined a systematic review as a review positively fulfilling the PRISMA reporting guidelines and host hoc decided to define a group of reviews failing on a maximum of two arbitrary PRISMA criteria as semi-systematic reviews.

Search methods for identification of reviews

We searched the Cochrane Library, Medical literature Analysis and Retrieval System Online (MEDLINE), Excerpta Medica Database (Embase), Science Citation Index-Expanded, BioSciences Information Service (BIOSIS)-Previews, Cumulative Index to Nursing & Allied Health Literature (CINAHL), and Latin American Caribbean Health Sciences Literature (LILACS) in order to identify reviews eligible for inclusion.

Data collection, data extraction and risk of bias assessment

Four independent authors (the author of this thesis and co-authors) selected the reviews and extracted data. Reviews containing a methods section and/or a literature search were checked against the PRISMA criteria. Hereafter, we assessed the methodological quality with the ROBIS tool (a tool for assessing risk of bias in systematic reviews) of the reviews failing on a maximum of two arbitrary PRISMA criteria.

Data synthesis

Data were presented descriptively.

7.2. Study II, III, IV and V – Systematic reviews of randomised trials

Study II, III and IV and V are systematic reviews with meta-analyses and TSA performed according to the pre-published protocols.[74-77] We used the recommendations by the Cochrane Collaboration[9] and the eight-step procedure for better validation of meta-analytic results in systematic reviews,[79] and reported the manuscripts following PRISMA.

Eligibility criteria

We included randomised clinical trials irrespective of publication status, publication date and language in all four papers.

Outcomes

Table 2 summarises the prespecified outcomes.

Table 2. Prespecified outcomes for study II to V

Study II	
Primary outcomes:	All-cause mortality
	Proportion of participants with one or more serious adverse reaction* (composite outcome)
Secondary outcomes:	Days alive without delirium within 28 days
	Quality of life
	Cognitive function
	Delirium severity
Exploratory outcome:	QT prolongation
Study III	
Primary outcomes:	All-cause mortality
	Any gastrointestinal bleeding
Secondary outcomes:	Proportion of participants with one or more serious adverse event* (composite outcome)
	Quality of life
	Myocardial ischaemia
	Hospital-acquired pneumonia
	C difficile infection
Study IV	
Primary outcomes:	All-cause mortality
	Proportion of participants with one or more serious adverse event* (composite outcome)
	Quality of life
Secondary outcomes:	Lung injury** (composite outcome)
	Acute myocardial infarction diagnosed after randomisation
	Stroke diagnosed after randomisation
	Sepsis diagnosed after randomisation
Study V	
Primary outcomes:	All-cause mortality
	Proportion of participants with one or more serious adverse event* (composite outcome)
Secondary outcomes:	Quality of life
	Lung injury** (composite outcome)
	Sepsis occurring after randomisation
	Cardiovascular events*** (composite outcome)

*defined as any untoward medical occurrence that resulted in death, was life-threatening, required hospitalisation or prolongation of existing hospitalisation, and resulted in persistent or significant disability or jeopardised the patient[80]

**defined as either ALI/ARDS, pulmonary fibrosis or pneumonia, or as defined by trialists

***defined as either myocardial infarction, stroke, peripheral arterial thrombosis, deep vein thrombosis, pulmonary embolism, or as defined by trialists

For the composite outcomes, we estimated the reported proportion of participants with one or more serious adverse events, lung injuries and cardiovascular events, in two ways:

1. by choosing the one specific event with the highest proportion reported in each trial that addresses the lowest possible proportion of participants with one or more events.
2. by cumulating all reported events, assuming that participants only experience one event (the number of participants in each group will constitute a maximum), address the highest possible reported proportion of participants with one or more events.

Search methods for the identification of trials

We searched the Cochrane Library, MEDLINE, Embase, Science Citation Index-Expanded including Conference Proceedings Citation Index, BIOSIS-Previews, Cumulative Index to Nursing & Allied Health Literature (CINAHL), LILACS in order to identify trials eligible for inclusion.

Furthermore, we searched for on-going and unpublished trials in the following registers: ClinicalTrials.gov; WHO International Clinical Trials Registry Platform (ICTRP); EU clinical trial register and Australian New Zealand Clinical Trials Registry (ANZCTR).

The websites of the U.S. Food and Drug Administration (FDA), the European Medicines Agency (EMA) and websites of medical companies including pharmaceutical trial registers were searched for unpublished trials for Paper II and III. Ultimately, we searched the reference lists of the included trials and previous meta-analyses to identify further relevant trials.

Data collection and data extraction and risk of bias assessment

Two independent authors (the author of this thesis and a co-author) selected the trials, extracted data and assessed risk of bias. The risk of bias assessment was conducted using the Cochrane risk of bias assessment tool, specifically assessing the following domains: 1) random sequence generation; 2) allocation concealment; 3) blinding of participants and personnel; 4) blinding of outcome assessment; 5) incomplete outcome data; 6) selective outcome reporting; and 7) other bias, including early stopping and bias due to vested financial interest or academic bias. We adjudicated trials as ‘overall low risk of bias’ when all bias domains were adjudicated as low risk of bias. Conversely, trials were adjudicated as ‘overall high risk of bias’ when unclear or high risk of bias was adjudicated in one or more domains.

The corresponding author of the included trial reports were contacted when a bias domain was adjudicated unclear risk of bias; if no reply was received, the bias domain remined unclear. Furthermore, we asked about insufficiently reported data and asked for additional data on unreported outcomes corresponding to our pre-specified outcomes.

Data synthesis

We used conventional meta-analytic statistics to calculate pooled effects estimates of each outcome using Review Manager (RevMan).[81] Relative risks (RR) with 95% confidence intervals (CI) were calculated for dichotomous outcomes. For continuous outcomes, we used end-scores, mean difference or standardised mean difference, and reported 95% CI.

Intervention effects were assessed with both fixed effect model meta-analysis and random effects model meta-analysis. We used the model presenting the more conservative point estimate of the two, which is the point estimate closest to no effect. The estimate with the widest CI was used, if the estimate from the two models were approximately equal.[79]

We adjusted the threshold for significance in the meta-analyses according to the amount of co-primary and co-secondary outcomes.[79] Furthermore, we used Bayes factor, with the anticipated intervention effect as the alternative hypothesis, to assess the ratio of the probability of the data (in the meta-analysis) given the alternative hypothesis, divided by the probability of the data given the nul-hypothesis.[79]

Pre-specified subgroup analyses aiming to assess the impact of clinical heterogeneity were performed.
We also performed analyses to assess the potential impact of systematic errors (bias), by conducting subgroup analysis of trials with overall low risk of bias.

We used TSA to calculate the meta-analytic required information size considering risk of random errors due to sparse data, multiple outcomes, and multiple testing of accumulating data.[25, 26, 82] We used an alpha corresponding to the adjusted threshold for significance, a power of 90% (beta 10%), a diversity as suggested by the trials in the meta-analysis, and used an a priori relative risk reduction or relative risk increase as anticipated intervention effects.

The GRADE approach was used to assess the overall certainty of evidence for all pre-defined outcomes.[43] We used the GRADEpro Guideline Development Tool (GDT) software to create the summary of findings tables.[83] We appraised the certainty of evidence and our confidence in the effect estimates, considering risk of bias, inconsistency, indirectness, imprecision and publication bias. We rated the overall certainty of evidence for all pre-specified outcomes as high, moderate, low or very low.

7.3. Study VI - Clinical heterogeneity in meta-analyses

Study VI is a methodological study, where we developed a tool (CHIMS) to assess and quantify clinical heterogeneity in meta-analyses. The study was performed according to the pre-published protocols and reported following the Guidelines for Reporting Reliability and Agreement Studies (GRRAS).[78, 84]

Development of CHIMS

We constructed CHIMS based on work by Gagnier and colleagues, who designate a list of clinical variables suggested for investigating clinical heterogeneity in meta-analyses.[12, 13] The pilot phase included the draft and review of the CHIMS tool, test scoring of meta-analyses, writing a manual, and another round of test scoring.

Assessment of interrater scale reliability and agreement

We evaluated the final CHIMS tool for interrater scale reliability and agreement by scoring 60 meta-analyses. Two independent evaluators involved in the development of CHIMS and two independent evaluators not involved in the development of CHIMS scored 20 ICU meta-analyses and 20 non-ICU meta-analyses. In addition, a sample of 20 meta-analyses were CHIMS scored by two of the review's original authors. Finally, the evaluators pairwise agreed upon each item scored and thereby achieved a total consensus score.

Statistical analyses/data synthesis

We analysed the interrater reliabilities of the summarised total CHIMS with intraclass correlation coefficients (ICC) for co-developers of CHIMS, for non-developers of CHIMS and for pairs of original review authors. For pairs of original review authors, we similarly analysed interrater reliability within the four domains of CHIMS. We also calculated quadratic weighted kappa values for the agreement between the categorical classification of CHIMS (low: 0-11; moderate 12-18; high 19-22).

We performed linear regression for associations between the raters' summarised total CHIMS. We analysed the possible difference between the distributions of consensus CHIMS in ICU and non-ICU meta-analyses using the Mann-Whitney test.

Finally, we calculated quadratic weighted kappa values for the agreement between the categorical classification of CHIMS and the categorical classification of I^2 (low $I^2 \leq 30\%$; moderate $I^2 > 30\%$ to $\leq 60\%$; high $I^2 > 60\%$) modified from Higgins et al.[85] in the 60 meta-analyses. Linear regression was also used to analyse the possible association between the consensus CHIMS and I^2 in 60 meta-analyses.

We classified agreement as suggested by Landis and Koch: values less than 0 indicated poor, 0-0.20 slight, 0.21-0.40 fair, 0.41-0.60 moderate, 0.61-0.80 substantial, and 0.81-1.00 almost perfect agreement.[86]

All ICC and kappa values were presented with 95% CI. We used Statistical Package for Social Sciences (SPSS) version 17 (SPSS Statistics for Windows, Chicago: SPSS Inc.) for the analysis of scale reliability and http://vassarstats.net/kappa.html to calculate kappa values.

8. PROJECT SUMMARIES

This section briefly highlights the main findings of the papers included in this thesis. Detailed presentations are reported in the six papers.

8.1. Study I

We screened 3745 titles/abstracts and included 378 reviews reporting on either prevention or management of delirium. A total of 57 reviews included a method and/or a literature search section and was thus checked against the PRISMA criteria. The 378 reviews were composed of 369 narrative reviews, eight semi-systematic reviews (defined as a review failing on a maximum of two arbitrary PRISMA criteria) and one systematic review fulfilling all 27 PRISMA criteria. Only the systematic review was overall low risk of bias, but only included trials of overall high risk of bias; the remaining eight semi-systematic reviews were overall high risk of bias.

The systematic review and eight semi-systematic reviews all assessed prevention of delirium and all assessed the effects of alpha-2-agonists. None of these reviews found evidence of a reduction in mortality (systematic review RR 0.99, 95% CI 0.79 to 1.24). Furthermore, the systematic review and three semi-systematic reviews found no evidence of a difference for the prevention of delirium (systematic review RR 0.85, 0.63 to 1.14) with alpha-2-agonists; conversely, four semi-systematic reviews found a beneficial effect. Serious adverse events, quality of life, non-serious adverse events and cognitive function were not assessed in either the systematic review, nor any of the eight semi-systematic reviews. We did not identify any systematic or semi-systematic reviews investigating the effect of other pharmacological interventions for the prevention of delirium.

For the management of delirium, we did not identify any systematic or semi-systematic review investigating the effect of pharmacological agents. Of all 378 reviews, 60% stated that haloperidol was indicated for the management of delirium.

As haloperidol is the preferred pharmacological agent used for the management of delirium in the ICU setting, and because no systematic review assessing the effects of haloperidol was identified in the overview, we conducted a systematic review assessing the benefits and harms of haloperidol versus any intervention in critically ill patients (defined as patients who are at high risk of dying or who have actual or potential life-threatening health problems irrespective of setting).

8.2. Study II

We screened 3863 titles/abstracts and included eight randomised clinical trials with 11 comparisons and with a total of 951 patients. The 11 comparisons compared haloperidol with placebo in two, dexmedetomidine in one, morphine in one, benzodiazepine in one, ondansetron in two and antipsychotics in four trials. Five comparisons used haloperidol as escape drug. Critically ill patients were included in all trials and was comprised of patients admitted to the ICU in five trials (seven comparisons), cardiac surgical patients in two trials (two comparisons) and medical patients in one trial (two comparisons). One trial was overall low risk of bias, the remaining were overall high risk of bias.

Three trials used haloperidol as escape drug; excluding these and regardless of risk of bias, we did not find evidence of a difference on all-cause mortality (RR 1.01; 95% CI 0.33-3.06; I^2=0%; 3 trials; 4 comparisons; 112 participants; very low certainty of evidence) or delirium severity (standardised mean difference (SMD) -0.15; 95% CI -0.61-0.30; I^2=27%; 3 trials; 4 comparisons; 134 participants; very low certainty of evidence), when comparing haloperidol with control interventions. The corresponding summary estimates in all trials showed similar results. Subgroup analysis of trials conducted in the intensive care setting versus all other settings shoved no interaction (P = 0.87 and P = 0.45, respectively).

Excluding trials using haloperidol as escape drug, no trials reported adequately on serious adverse reactions/events; one trial reported on days alive without delirium, cognitive function or QTc prolongation; and no trials reported on quality of life.

8.3. Study III

We screened 10,054 titles/abstracts and included 42 trials with a total of 6899 ICU patients. Three trials were overall low risk of bias; the remainder were overall high risk of bias.

We did not find evidence of a difference in trials with overall low risk of bias on mortality (RR 1.03, 95% CI 0.94–1.14; I^2=0%; TSA-adjusted CI 0.94–1.14; 3 trials; 3557 participants; high certainty of evidence), when comparing stress ulcer prophylaxis with placebo or no prophylaxis. The corresponding summary estimates in all trials showed similar results (RR 1.01; 95% CI; 0.93-1.10; TSA-adjusted CI 0.93-1.10, 28 trials; 5656 participants; moderate certainty of evidence). Subgroup analysis of proton pump inhibitors versus histamin-2 receptor antagonist shoved no interaction (P = 0.51).

Meta-analysis and TSA showed evidence of a difference in trials with overall low risk of bias on the occurrence of any gastrointestinal bleeding (RR 0.60, 95% CI 0.47–0.77; I^2 = 0%; TSA-adjusted CI 0.36–1.00; 3 trials; 3596 participants; high certainty of evidence). The corresponding summary estimates in all trials showed similar results (RR 0.52; 95% CI 0.45-0.61; I^2=43%; TSA-adjusted CI 0.39-0.68; 39 trials; 6627 participants; low certainty of evidence). Meta-analysis indicated that clinically important gastrointestinal bleeding was reduced (RR 0.63, 95% CI 0.48–0.81), but the TSA-adjusted CI 0.35–1.13 indicated lack of firm evidence. Subgroup analysis of proton pump inhibitors versus histamin-2 receptor antagonist shoved no interaction (P = 0.38).

Meta-analysis of trials with overall low risk of bias showed no evidence of a difference on the estimated highest reported proportion of serious adverse events (RR 1.03; 95% CI 0.94-1.14; I^2 = 0%; TSA adjusted CI 0.94-1.14; 3 trials; 3587 participants; low certainty of evidence), pneumonia (RR 1.01; 95% CI 0.87-1.18; I^2 = 0%; TSA adjusted CI 0.77-1.33; 3 trials; 3596 participants; moderate certainty of evidence), and C. difficile infection (RR 0.84; 95% CI 0.48-1.47; I^2 = 0%; 3 trials; 3596 participants; low certainty of evidence). The corresponding summary estimates in all trials showed similar results.

Only one trial reported on myocardial ischemia and no trials reported on quality of life.

8.4. Study IV

We screened 32,813 titles/abstracts and included 10 trials with a total of 1458 participants. All trials were overall high risk of bias; two trials were low risk of bias on all domains except for blinding of participants and personnel.

Meta-analysis indicated evidence of increased mortality (RR 1.18; 95% CI 1.01 to 1.37; I^2 = 0%; 4 trials; 1135 participants; very low certainty of evidence) when comparing higher with lower levels of oxygen supplementation. TSA showed that the required information size was not reached, and the CI included a 20% relative risk increase. Therefore, TSA could not confirm or refute a 20% increase in mortality.

Meta-analysis indicated evidence of harm on serious adverse events (estimated highest proportion of specific serious adverse events in each trial RR 1.13; 95% CI 1.04 to 1.23; I^2 = 0%; 1234 participants; 6 trials; very low certainty of evidence) from higher levels of oxygen supplementation. TSA showed that the cumulative Z-curve crossed the trial sequential monitory boundary for harm.

Meta-analysis showed no evidence of a difference on lung injury (estimated highest reported proportion of lung injury RR 1.03; 95% CI 0.78 to 1.36; I^2 = 0%; 1167 participants; 5 trials; very low certainty of evidence). TSA showed that the cumulative Z-curve did not cross any boundaries for harm or benefit, nor the trial sequential boundary for futility, indicating insufficient information to confirm or reject a 20% relative change.

Only one trial reported on the effects on sepsis and no trials reported on quality of life, acute myocardial infarction, or stroke.

8.5. Study V

We screened 35,402 titles/abstracts and included 50 trials with a total of 21,014 participants. All trials were overall high risk of bias; five trials were low risk of bias on all domains except for blinding of participants, personnel, and outcome assessors.

Meta-analysis and TSA in trials with overall low risk of bias except for blinding, showed no evidence of a difference on mortality (RR 0.98, 95% CI 0.89-1.09, I^2 = 0%; TSA-adjusted CI 0.86-1.12; 8 trials; 16,156 participants; low certainty of evidence), when comparing higher with lower levels of oxygen supplementation. The corresponding summary estimates in all trials showed similar results (RR 1.04; 95% CI 0.96-1.13; I^2=2%; TSA-adjusted CI 0.96-1.13, 34 trials; very low certainty evidence). Subgroup analysis of trials conducted in the intensive care setting versus all other settings, shoved no interaction (P = 0.71).

Meta-analysis and TSA in trials with overall low risk of bias except for blinding, showed no evidence of a difference on the proportion of participants with one or more serious adverse events (RR 0.99, 95% CI 0.89-1.12, I^2=0%; TSA-adjusted CI 0.83-1.19; 3 trials, 8056 participants; low certainty of evidence). The corresponding summary estimates in all trials showed similar results (RR 1.03; 95% CI 0.95-1.13; I^2=17%; TSA-adjusted CI 0.91-1.18; 6 trials; 8874 participants; low certainty of evidence). Subgroup analysis of trials conducted in the intensive care setting versus all other settings, shoved no interaction (P = 0.15).

Meta-analysis and TSA of all trials regardless of risk of bias, showed no evidence of a difference on quality of life (mean difference (MD) 0.37; 95% CI -1.55-2.29; I^2=57%; TSA-adjusted CI -2.41-3.16; 6 trials, 7445 participants; very low certainty of evidence), estimated highest reported proportion of lung injury (RR 0.93; 95% CI 0.76-1.12; I^2 = 0%; TSA-adjusted CI 0.64-1.32; 10 trials; 9227 participants; very low certainty of evidence), sepsis (RR 1.64; 95% CI 0.96-2.80; I^2=0%; 4 trials; 1307 participants; very low certainty of evidence), and the estimated highest reported proportion of cardiovascular events (RR 1.06; 95% CI 0.86-1.31; I^2=11%; TSA-adjusted CI 0.45-2.51; 16 trials; 16,607 participants; very low certainty of evidence).

8.6. Study VI

We developed the first tool for assessing and quantifying clinical heterogeneity in meta-analyses of interventions - CHIMS.

CHIMS measures clinical heterogeneity on a scale that includes four domains with 11 items overall:

- Setting (time of conduct/country development status/unit type)
- Population (age; sex; patient inclusion criteria/baseline disease severity, co-morbidities)
- Intervention (intervention intensity/strength/duration of intervention; timing; control intervention; co-interventions)
- Outcome (definition of outcome; timing of outcome assessment)

Each item is scored 0 to 2 points, where 0 points corresponds to low clinical heterogeneity, 1 point corresponds to moderate (or unknown/undescribed) clinical heterogeneity, and 2 points corresponds to high clinical heterogeneity.

Assessment of clinical heterogeneity in a meta-analysis using CHIMS is completed in two steps: 1) two authors independently assess clinical heterogeneity in the four domains, 2) after agreeing upon scores of individual items, a consensus score is achieved.

We tested the interrater scale reliability and agreement of CHIMS, and found almost perfect interrater scale reliability (ICC 0.94, 95% CI 0.85-0.98 for average measures and ICC 0.89, 95% CI 0.75-0.96 for single measures) and substantial agreement (kappa 0.72, 95% CI 0.42-1.00) in pairs of original review authors; almost perfect to substantial interrater scale reliability (ICC 0.85, 95% 0.72-0.92 for average measures and ICC 0.74, 95% CI 0.56-0.85 for single measures) and substantial agreement (kappa 0.61, 95% CI 0.18-1.00) in the pair of co-developers; and substantial to moderate interrater scale reliability (ICC 0.74, 95% 0.51-0.86 for average measures and ICC 0.59, 95% CI 0.34-0.76 for single measures) and moderate agreement (kappa 0.41, 95% CI 0.14-0.69) in the pair of non-developers.

We found higher clinical heterogeneity in ICU-meta-analyses (median consensus CHIMS between co-developers = 18) compared with non-ICU meta-analyses (median consensus CHIMS between co-developers = 12).

In a supplementary exploratory analysis, we found no linear association between clinical heterogeneity measured with CHIMS and statistical heterogeneity measured with I^2.

9. DISCUSSION

9.1. Principal findings

This PhD project examined the evidence gathered about three specific adjuvant interventions applied to patients admitted to the ICU. We investigated the benefits and harms of management of delirium, stress ulcer prophylaxis and oxygen supplementation by conducting systematic reviews with meta-analysis and TSA.

To assess the evidence of delirium management, we conducted a systematic overview of reviews of pharmacological prevention and management of delirium in the ICU setting, as well as a systematic review of haloperidol for the management of delirium in the critical care setting. Results of the overview show that a large number of reviews on pharmacological prevention and management have been published, but only one prevention review was found to be systematic according PRISMA. Of all included 378 reviews, 60% stated that haloperidol was indicated for the management of delirium, even though we found no single systematic review investigating the effect of haloperidol for the management of delirium.

We proceeded with the conduct of a systematic review assessing the benefits and harms of haloperidol versus any intervention in critically ill patients. We only found 8 randomised clinical trials with 11 comparisons including a total of 951 patients. Just two trials (two comparisons) compared haloperidol with placebo and 3 trials (5 comparisons) used haloperidol as rescue medication. One trial was adjudicated as overall low risk of bias; however, the validity of this adjudication may not be fully justified, as the same trial used haloperidol as rescue medication. Meta-analyses showed no evidence indicating a difference, when comparing haloperidol with any intervention on mortality, delirium severity, QTc prolongation, delirium resolution, extrapyramidal symptoms. Only one trial reported on days alive without delirium and cognitive function, no trials reported adequately on serious adverse events - and no trials reported on quality of life.

To assess the evidence of stress ulcer prophylaxis in the ICU setting, we conducted a systematic review assessing the effects of proton pump inhibitors or histamin-2 receptor antagonist. We did not find evidence that showed a difference on mortality and we were able to refute a 15% relative change. Any gastrointestinal bleeding and clinical important bleeding were both reduced with stress ulcer prophylaxis. Meta-analyses showed no evidence of a difference on serious adverse events, myocardial ischemia, pneumonia, C. difficile infection. No trials reported on quality of life when comparing stress ulcer prophylaxis with placebo or no prophylaxis.

To assess the evidence of the effect of applying higher versus lower levels of oxygen supplementation in patients admitted to the ICU and critically ill patients, we conducted two systematic reviews. In the review focussing on ICU patients, we found an increased risk of mortality in the traditional meta-analysis; however, the TSA showed that no boundaries for benefit and harm, nor the trial sequential monitoring boundary for futility, was crossed, indicating that evidence was insufficient to confirm or refute a 20% relative change for the benefit or harm of higher versus lower oxygen supplementation. Meta-analysis showed an increase in serious adverse events. Due to insufficient data, the effects on quality of life, acute myocardial infarction, stroke, and sepsis are inconclusive. To assess the benefit and harms of oxygen supplementation in a broader clinical context, we conducted a similar review that included all critically ill patients. We found no evidence of a difference in mortality. The TSA anticipation of a 15% relative change showed that the boundary for futility was crossed, indicating firm evidence for lack of a 15% difference in mortality between the groups. A subgroup analysis of trials conducted in the ICU setting revealed that there was no evidence of a difference, which contrasts with the results from the review focussing on ICU patients. The contrast is caused by the inclusion of more information[26] and highlights the necessity of avoiding drawing firm conclusions (altering

guidelines) about intervention effects before the required information size has been reached or one of the trial sequential monitoring boundaries has been crossed.

To increase our knowledge of clinical heterogeneity in meta-analysis and support our view that no single method exists to assess clinical heterogeneity in meta-analysis, we developed CHIMS and tested the scale reliability and agreement of the developed score. CHIMS covers four domains (setting, population, intervention, and outcome) with a total of 11 items. It was tested by 3 groups of raters and was found to be a reliable tool for assessing and quantifying clinical heterogeneity in meta-analyses. Using CHIMS, we found higher clinical heterogeneity in meta-analyses of ICU meta-analyses compared to non-ICU meta-analyses. Moreover, we found no linear association between CHIMS and I^2.

9.2. Methodological strength and limitations of the systematic reviews

All reviews follow the methodology pre-specified in the published protocols and the minor differences between protocols and reviews are transparently reported.[73-78]

We a priori decided to report on the proportion of participants with one or more serious adverse events, which is a composite outcome. This approach was chosen to increase statistical power. However, only few trials reported on the proportion of participants with one or more serious adverse events as a composite outcome – we therefore decided to estimate the reported proportion of participants with one or more serious adverse events by:

1) choosing the one specific event with the highest proportion reported in each trial that addresses the lowest possible proportion of participants with one or more events
2) cumulating all reported events, assuming that participants only experience one event

Each component of composite outcomes may not have similar degrees of severity, and therefore could bias the results of the outcome.[4] If, for example, more serious adverse events occur in one intervention group, and other less severe serious adverse events occur in the other intervention group, then there is a risk of overlooking actual severity differences between the compared groups, when analysing the composite outcome. Potential differences can be assessed by analysing each component included in the serious adverse event outcome separately but this approach risks repetitive testing;[4] therefore, we did not use this approach. Furthermore, the analyses estimating the highest proportion of serious adverse events imply that participants included in the highest proportion also include participants having other serious adverse events. For example, if mortality is the highest proportion, then it is implied that all the participants who did not die, did not experience another serious adverse events; this analysis thus underestimates the proportion of participants with one or more serious adverse events, as participants not included in the highest proportion would be expected to experience other serious adverse events not included in the highest proportion. In addition, the analyses estimating the cumulated proportion of serious adverse events imply that all participants who experience a serious adverse events had only this specific serious adverse events, which definitely overestimates the proportion of participants with one or more serious adverse events, since a minimum of one participant would be expected to have more than one serious adverse events.

We searched for published trial results in all the major medical databases and for unpublished results in trials registers. Literature screening, data extraction and risk of bias assessment to evaluate the risk of systematic errors was performed by two independent authors. Systematic errors detected by the risk of bias assessment was sought clarified by contacting authors from all included trials. We emphasize the result of trials with overall low risk of bias, as methodological quality of trials may impact intervention effects and therefore the

conclusions and the validity of the systematic reviews;[4] this approach has recently been adopted by the Cochrane collaboration.[87]

We used TSA to control the risk of random errors due to multiple outcomes, sparse data, and multiple testing on accumulating data, in order to prevent us from drawing false firm conclusions on meta-analyses with insufficient information size. The recently published update of the Cochrane Handbook includes a chapter on prospective approaches to accumulating evidence.[88] Formal sequential statistical methods are discouraged for standard meta-analyses in most circumstances in Cochrane reviews; it may be performed as a secondary analysis – however, it may not be used for the main analyses and neither to be used to draw main conclusions.[88]

Finally, in order to summarise the certainty of the evidence of the assessed interventions, we used the GRADE approach and summarised the results in summary of findings tables.

9.3. Current evidence and implications

Delirium management

The evidence for the use of haloperidol to manage delirium in ICU patients and in critically ill patients is sparse, of low quality and inconclusive, due to a very low amount of data (very low certainty of evidence). Whether haloperidol has a beneficial, neutral, or harmful effect in ICU patients with delirium is unknown. Thus, implications for research include the need for conducting trials with overall low risk of bias comparing haloperidol with a placebo – and not allowing the use of haloperidol as rescue medication - to be able to firmly assess and conclude on the effect of haloperidol on delirium management. Furthermore, reporting of patient centred outcomes such as all-cause mortality, days alive without delirium, serious adverse events, quality of life and cognitive status post treatment should be prioritised. In the near future, a set of core outcomes to be used in trials on delirium prophylaxis and delirium management will be published, which will contribute to standardising reporting of outcomes that may increase information included in meta-analysis.[89]

Currently, two randomised clinical trials comparing haloperidol with a placebo are recruiting patients. The results of these trials will, in the coming years, shed light on the effects of haloperidol.[90, 91] We intend to update the haloperidol review when the results of the AID-ICU trial have been published.[92]

Patients with delirium challenges clinicians, as no single or bundle intervention has yet proved its effect on delirium in this setting. Clinical implications include that patients with delirium should be identified and that non-pharmacological interventions are applied as first option, although the effect on non-pharmacological interventions are also based on very low to low certainty of evidence.[93]

No pharmacological drug has proven its effectiveness for the management of delirium;[94] therefore, if a pharmacological intervention is deemed necessary then haloperidol may still be used, even though it is unknown whether haloperidol relieves delirium severity, reduces days with delirium and whether haloperidol has an effect on serious adverse events and mortality.

Stress ulcer prophylaxis

There is high certainty of evidence that stress ulcer prophylaxis does not reduce mortality by a 15% relative change, although there is high certainty of evidence that gastrointestinal bleeding is reduced with at least a 20% relative change (RR 0.60, 95% CI 0.47-0.77). Clinically important bleeding is also reduced with the use of stress ulcer prophylaxis (RR 0.63, 95% CI 0.48-0.81), which has also been shown in a meta-analysis including all critically ill patients (RR 0.62, 95% CI 0.43-0.89).[95] It is inconclusive whether stress ulcer prophylaxis has an

effect on serious adverse events, quality of life, pneumonia, myocardial ischemia and C. difficile infection (very low to moderate certainty of evidence).

Only 5.3%-6.4% of the patients admitted to the ICU experience a gastrointestinal bleeding, and only 5.4% experience clinically important bleeding. According to results from the SUP-ICU trial,[96] increased mortality may be associated with higher disease severity. However, additional research is needed to draw firm conclusions.[96-98] New trials with overall low risk of bias on the most ill patients (e.g. according to Simplified Acute Physiology Score (SAPS) II) in the ICU are needed to explore whether this patient group is at higher risk of experiencing harmful events caused by stress ulcer prophylaxis. Furthermore, it has been suggested that patients with acute kidney injury, coagulopathy, shock and chronic liver disease are at increased risk of developing clinically important bleeding;[99] further research on these groups are also needed.

Clinical implications include whether stress ulcer prophylaxis should be used as a standard treatment. It may be reasonable to argue that stress ulcer prophylaxis should not be used for all ICU patients, when only about 5% of the ICU patients bleed at all, particularly as overall all-cause mortality does not seem to be affected. Alternatively, e.g. proton pump inhibitors may be reserved for the patient population at high risk of developing stress ulcers who are actually showing signs of bleeding. It seems superfluous to administer prophylaxis to 95% of a population, who never develop the condition that prophylaxis is administered to avoid.

Oxygen

There is low to very low certainty of evidence that higher versus lower fraction of inspired oxygen, or targets of arterial oxygenation, have beneficial or harmful effects in ICU patients. The effect on mortality is less than 15% and 20% on serious adverse events, when assessing critically ill patients.

Largely all patients admitted to the ICU receive oxygen supplementation; it may therefore be important to find even smaller differences in mortality than the 15% relative change we have disproved. As such, implications for research include new trials with overall low risk of bias, including blinding of participants, personnel, and outcome assessor. These are highly needed. In addition, it is important to aim for a clear separation of the higher and lower oxygen administration. Even when a clear separation between the groups have been defined, interpretation is still difficult. E.g. patients randomised to a high saturation does not imply that they eventually get more oxygen supplementation, as patients with “good” lung-function do not need as much oxygen to reach the same saturation as patients with “bad” lung-function. Targeting specific arterial oxygen tensions (PaO_2), arterial oxygen saturations (SaO_2) or peripheral oxygen saturation (SpO_2) may be the right way to discern and compare interventions of oxygen supplementation (a simple personalised intervention mode). Standardized delivery of different fraction of inspired oxygen (FiO_2) levels may be a too simplistic way to apply different levels of oxygen supplementation, as many patients will not need high FiO_2 levels to reach an acceptable PaO_2, SaO_2 or SpO_2 target and some patients will need higher FiO_2 levels to reach even a low PaO_2, SaO_2 or SpO_2 target. Furthermore, patient centred outcomes should be reported. Core outcome sets for ventilation trials, and for patients surviving acute respiratory failure, should be prioritised.[100, 101] Currently, five randomised trials assessing higher versus lower levels of oxygenation in the ICU setting are ongoing;[102-106] when the results of the HOT-ICU trial are published, we will update our reviews.[105, 107]

Based on a meta-analysis published by Chu et al., who found an increase in mortality with higher levels of oxygen supplementation (RR 1.14, 95% CI 1.01-1.28; high certainty evidence), clinical practice has moved toward using lower levels of oxygen compared to previously. However, our results, which include more data, may not entirely support Chu et al.’s results. Nonetheless, almost all effect estimates (all statistically insignificant CIs) from our meta-analyses indicate harmful effects when using higher levels of oxygen, which may indicate more harm with higher levels of oxygen. Therefore, clinical implication includes that

unnecessary high fractions of inspired oxygen, or targets of arterial oxygenation, should be avoided in routine clinical practice. However, as long as there is a considerable possibility that higher targets of PaO_2 compared with lower targets may benefit groups of patients, randomised clinical trials seems justified to encompass such targets and populations.

Other adjuvant interventions

Besides management of delirium, stress ulcer prophylaxis and oxygen supplementation, other adjuvant interventions are used in the ICU setting. Fluid resuscitation is used in the initial management of sepsis with the aim of improving the circulation. A fixed volume for resuscitation is recommended by The Surviving Sepsis Campaign guideline (low certainty evidence), including continuation of fluid therapy as long as hemodynamic variables improve.[108] However, a recent published systematic review with meta-analysis revealed that only few trials have been published and the evidence is of very low certainty on all assessed outcomes.[109] The accumulation of fluids is frequent and fluid removal by diuretics as well as ultrafiltration is used; it is uncertain whether forced fluid removal improves outcomes compared with less fluid removal.[110] Management of new-onset atrial fibrillation, which is common in critically ill patients, is associated with increased morbidity and mortality; however, data supporting pharmacological management strategies are of very low to low certainty of evidence.[111] Another adjuvant intervention is nutrition support, which also has low certainty of evidence.[112, 113] Thrombosis prophylaxis with low-molecular-weight heparin also come with low certainty of evidence on the effect on mortality, although it is associated with a beneficial effect (moderate certainty of evidence) on symptomatic venous thromboembolism.[114]

There may be a similar lack of firm evidence for the use of core interventions. For instance, antibiotics and combinations of different antibiotics for severe sepsis is associated with very low certainty of evidence on the effect on mortality and low certainty of evidence on other studied outcomes.[115] A study on all meta-analyses assessing interventions used in intensive care medicine has revealed that less than 1% of meta-analyses are of overall low risk of bias and only 9% of the meta-analyses were reported following the 27 PRIMA criteria.[116] Furthermore, the same group found that of 50% of the meta-analyses suggesting a beneficial effect, only 50% were true positive according to TSA; and when limiting to meta-analyses with overall low risk of bias, only one outcome had a beneficial effect and another a harmful effect.[23]

9.4. Heterogeneity

Heterogeneity, specifically in the patient population and among co-interventions used in the ICU, may challenge the interpretation of results from randomised trials. Some opinion leaders even believe that the randomised trial design should not be used in this setting, especially due to the heterogeneous population, who only has that in common that they are critically ill and therefore admitted to an ICU, but suffers from various types and degrees of organ dysfunction; it is argued that they will not respond similarly to different types of interventions.[70, 71] We advocate that claims of intervention effect heterogeneity should be supported by evidence before specific advice of restricting or extending interventions to subgroups are issued.

Heterogeneity may indeed challenge interpretation of meta-analyses, which summarises treatment effect as a precision weighted average effect. In the context of meta-analyses, as described in the introduction, heterogeneity may refer to methodological variability across trials (methodological heterogeneity), variability in effect estimates across trials (statistical heterogeneity) and variability in settings, participants, interventions and comparators, use of co-interventions, and the types and timing of outcome assessments (clinical heterogeneity). Usually, methodological heterogeneity is addressed by the risk of bias assessment and a following separation of trials, in trials with overall low risk of bias and trials with overall high risk of bias. High statistical heterogeneity may have the result that the meta-analysis of all trials is not conducted, e.g.

when the investigation of clinical heterogeneity by subgroup analyses have revealed interaction between assessed clinical differences.

To address the overall clinical heterogeneity in a meta-analysis, we developed CHIMS – a tool that may be used in the systematic review process to assess and quantify clinical heterogeneity in a meta-analysis.
The items selected to describe clinical heterogeneity in the CHIMS tool are in accordance with definitions of clinical heterogeneity from various organisations.[14, 117] To test the reliability of CHIMS, it was tested by three groups of reviewers; by pairs of co-developers of CHIMS and by pairs of non-developers of CHIMS – both pairs assessed the same 40 meta-analyses; and finally, by 20 pairs of original review authors. Interrater scale reliability and agreement ranged from moderate to almost perfect depending on the type of raters: almost perfect to substantial by pairs of original review authors, whilst co-developers achieved almost perfect to substantial, and non-developers achieved moderate to substantial interrater scale reliability and agreement.

To assess the presence of clinical heterogeneity in meta-analyses of trials in intensive care patents compared to meta-analyses of trials of all critically ill patients (of which ICU patients is a sub-group), the primary outcome, mortality, was CHIMS scored in both clinical contexts within the interventions: haloperidol for the management of delirium, stress ulcer prophylaxis and oxygen supplementation (Table 3).

Table 3. Clinical heterogeneity assessed with CHIMS in meta-analyses of three different interventions on mortality ICU patients versus critically ill patients

		Haloperidol in ICU patients [118]	Haloperidol in critically ill patients [118]	Stress ulcer prophylaxis in ICU patients [119]	Stress ulcer prophylaxis in critically ill patients [95]	Oxygen supplementation in ICU patients [120]	Oxygen supplementation in critically patients [107]
Details of meta-analysis							
Outcome		All-cause mortality	All-cause mortality	All-cause mortality	All-cause mortality	All-cause mortality	All-cause mortality
Number of trials included		3	6	25	37	8	34
Methodological heterogeneity							
Overall risk of bias		High	High	High	High	High	High
Statistical heterogeneity							
I^2 of meta-analysis		0%	0%	0%	2%	2%	0%
Statistical heterogeneity category*		Low	Low	Low	Low	Low	Low
Clinical heterogeneity assessed with CHIMS							
Setting	Years reported, performed in developed vs developing country, unit type	0	2	1	2	2	2
Population heterogeneity	Age	1	2	2	2	2	2
	Sex	1	0	1	2	2	2
	Participant incl. criteria and baseline disease severity	1	2	1	2	2	2
	Co-morbidities	1	2	2	2	2	2
Intervention heterogeneity	Intensity, strengths, or duration of intervention	2	2	2	2	2	2
	Timing	1	2	2	2	2	2
	Control intervention	2	2	2	2	1	1
	Co-intervention	2	2	1	1	2	2
Outcome heterogeneity	Definition of outcome	0	0	0	0	0	0
	Timing of outcome measurement	0	0	2	2	1	2
	CHIMS sum	11	16	16	19	18	19
	CHIMS category**	Moderate	Moderate	Moderate	High	Moderate	High

*low: $I^2 \leq 30\%$; moderate: $I^2 > 30\%$ to $\leq 60\%$; $I^2 > 60\%$

**low: CHIMS sum 0-11; moderate: CHIMS sum 12-18; high: CHIMS sum 19-22

In all three interventions, clinical heterogeneity was found to be higher in meta-analyses of critically ill patients as compared to meta-analyses of ICU patients. This strengthens the argument for the validity of CHIMS, as it may differentiate a lower clinical heterogeneity in meta-analyses of subgroups from a higher clinical heterogeneity in broader meta-analysis. According to our analyses, non-ICU meta-analyses has the

lowest clinical heterogeneity (median CHIMS between co-developers assessing 20 ICU meta-analyses: CHIMS = 12), whilst meta-analyses in ICU (mean CHIMS between two co-developers assessing 20 non-ICU meta-analyses: CHIMS = 18) and in the critically ill patients have higher clinical heterogeneity (median CHIMS from three meta-analyses in critically ill patients, Table 3: CHIMS = 19) – but does the clinical heterogeneity affect the statistical heterogeneity?

To assess the possible association between clinical heterogeneity and statistical heterogeneity, we assessed the clinical heterogeneity with CHIMS and compared it to the statistical heterogeneity expressed by the I^2 and found no linear correlation; in fact no correlation at all seemed most likely as other non-linear curve associations were also unlikely (private communication). This finding is consistent with our findings on the statistical and clinical heterogeneity of the six meta-analyses described above, where low statistical heterogeneity (I^2 0% to 2%) and moderate to high clinical heterogeneity (CHIMS 11 to 19) (Table 3) was found. Why is the clinical heterogeneity not reflected by the statistical heterogeneity? Reasons may be the influence of methodological heterogeneity (risk of bias) on the summery effect estimate, that I^2 may be a weak indicator for statistical heterogeneity,[121-123] that the I^2 , which is intimately linked to the DerSimonian-Laird model, may underestimate the heterogeneity variance.[124, 125] Choosing other and probably more adequate meta-analytic models, as e.g. the Hartung-Knapp model,[126-128] would necessitate a heterogeneity measure as D^2 which is able to estimate statistical heterogeneity in any random effects model.[27] Our sample size (60 meta-analyses assessed and amount of trials included in each meta-analysis) may be too small to show an association, or perhaps the intervention effects does not differ between populations and the variations in clinical heterogeneity (including the variation in interventions) does not influence the effect estimate.

9.5. Barriers for developing the evidence for adjuvant interventions in the ICU

Several barriers in the ICU setting challenges the design, conduct and interpretation of trial results. This may also be reflected in the meta-analyses of individual trials.

ICUs are usually small compared to other medical specialties, making it necessary to include several sites, often in an international setting, to recruit enough patients in trials to minimize random error and duration of recruitment. In addition, complex regulatory requirements, e.g. need for approvals by multiple ethics committees with different sets of requirements, multiple rules and different strategies for data management, pharmaceutical companies not required to provide placebo free of charge (and who may even decline producing and selling the placebo), are barriers.[129]

Informed consent must be obtained from patients entering the trial, but because patients admitted to the ICU often are incapable of providing the consent due to severe illness or sedation, the consent must be signed by a “surrogate decision maker” (designated proxy or a family member).[130] The emotional and logistic impact of acute hospitalization may make the decision of participating in a research project difficult for the surrogate caretaker. Time constraints may lead to loss of eligible patients that potentially could lead to bias or limit the generalizability of the results.

Blinding of participants, personnel and outcome assessors is important in clinical trials, as treatment effects may become biased if blinding has not been maintained. Compared to trials assessing the effectiveness of pharmacological agents, trials on medical devices or where the intervention is administered through a medical device such as oxygen supplementation, is often more difficult. When it is “impossible” to blind personnel, then the outcome assessor should be blinded;[131] however, in the ICU this may be a challenge as the outcome assessors are most often the personnel taking care of the patients – at least in terms of serious adverse events.

Treatment effect is reported as an average effect and heterogeneity in treatment effect is primarily examined by subgroup analyses. A limitation of the subgroup analysis is that is has reduced power due to the inclusion of fewer patients compared to the full trial cohort. However, the solution of just performing larger trials, thereby increasing the power/credibility of the subgroup analyses, extends the trial's duration and cost as more patients are enrolled. Alternatively, a homogeneous patient population comprising patients likely to have an outcome and to respond to the intervention, should be included; such an approach is called enrichment and has been proposed to increase the effectiveness of trials and meta-analysis.[132]

Furthermore, it has been argued that all-cause mortality may not be the optimal primary outcome in trials conducted in the ICU, as it may be difficult to show a difference in treatment effect;[133] reasons include decreased mortality rate in the ICU in recent years, which may be hard to decrease even more. Also, mortality in the ICU setting may be influenced by other factors, e.g. effect of other applied treatment modalities, such as core interventions or adjuvant interventions, than the tested intervention.

In our systematic reviews, we found that only a few trials reported patient-centred outcomes, such as quality of life and cognitive function. In addition, other outcomes such as serious adverse events, including individually reported serious adverse events, were often not reported, which may reflect the administrative challenges in registering all serious adverse events in the ICU. This highlights inconsistencies in the outcome selection, definition, and measurement, which hampers the progress towards improvement in care.

To increase the information size in meta-analyses and thereby better informing the evidence base, core outcome sets, which are an agreed, standardised collection of outcomes measured and reported in all trials for a specific clinical trial, have been or are being developed for critical care research.[89, 100, 101, 134] The following core outcome measurement set has been developed for clinical trials evaluating interventions intended to modify duration of mechanical ventilation; time from randomisation to first successful extubation, reintubation, duration of mechanical ventilation, duration of stay, mortality, and quality of life.[101] Another research group has proposed the following core outcome measurement for clinical trials evaluating survivors from acute respiratory failure after discharge; survival, quality of life, mental health, pain, cognition, physical function, muscle and/or nerve function and pulmonary function.[100] Furthermore, Rose and colleagues are working on the development of a core outcome set for trials assessing the effects of interventions on the prevention and/or treatment of delirium.[89]

Finally, assessing the effects of one adjuvant intervention may be insufficient to show a difference in intervention effect due to ICU patients being managed with several treatment modalities at the same time, e.g. oxygen supplementation, stress ulcer prophylaxis and fluid therapy – which may even differ between patients, clinical sites and countries. Future research in the ICU setting may therefore consider heterogeneity between the use of adjuvant interventions and consider randomising patients to intervention packages where e.g. lower levels of oxygen supplementation together with lower levels of fluids are used.

9.6. Improving the evidence for interventions used in the ICU

Systematic reviews of overall low risk of bias, including randomised clinical trials of overall low risk of bias, are the best evidence upon which to base changes in clinical practice. Caution should be taken when only a single randomised trial forms the basis for the evidence, as it may be difficult to predict or determine how clinical heterogeneity influence the results - meta-analysis of multiple less large independent trials may be a better way to evaluate the effect of an intervention.[135, 136] Thus, a single large mega trial may not be a suitable way to evaluate a treatment compared to meta-analysis of less large trials, as the latter has been shown to better lower the rates of false positive findings.[136] Even after inclusion of only five trials, meta-analyses produce results consistent with long-run findings, although firm conclusions should not be drawn as it may not be possible to predict precisely which findings will change with more data.[137]

As previously mentioned, the evidence for interventions used in the ICU setting has characteristically been approved or adapted for specific indications in the non-ICU setting and their use has subsequently been implemented in the ICU setting. The majority of meta-analyses assessing the effects of interventions used in the intensive care are of overall high risk of bias.[116] Furthermore, of all conventional meta-analysed statistically significant outcomes, 87% of meta-analyses have been shown to be inconclusive.[23] Therefore, there is a need to map out the evidence for treatment modalities used in the ICU derived from data retrieved in this setting.

An overview of published systematic reviews within groups of interventions may help this mapping. Network meta-analyses may also contribute to the mapping, such as the assessment of different pharmacological interventions for delirium.[94] Thus, overviews and network meta-analyses may identify knowledge gaps, which can be beneficial to the planning of future research. Risk of bias assessment of the systematic reviews will decide whether a new systematic review is needed; if no systematic review with overall low risk of bias exist, then there is a need to conduct one assessing the evidence for a specific intervention. If a systematic review with overall low risk of bias is identified but is out of date and new trials have been published, then a review update is needed. When a systematic review with overall low risk of bias forms the basis for a low certainty of evidence intervention effect, then the following research model may be applied: 1) perform a cohort study investigating the incidence of the specific scenario, 2) conduct a randomised trial with the lowest possible risk of bias designed and conducted based on the results of the observational study, 3) update the systematic review by including the results of the randomised trial. TSA may reveal whether the required information size and the required number of trials has been reached or not, and whether additional trials are needed.

10. CONCLUSION

Overall, the evidence was sparse or even absent, of low quality, with insufficient information sizes within the adjuvant interventions; haloperidol for the management of delirium, stress ulcer prophylaxis, and oxygen supplementation in patients admitted to the ICU.

The evidence for the use of haloperidol to manage delirium in ICU patients and in critically ill patients was sparse, of low quality and inconclusive due to very sparse data. The certainty of evidence was very low on mortality, serious adverse events, cognitive function, delirium severity and QTc prolongation. Whether the use of haloperidol should be avoided in the ICU setting is a dilemma, as no other pharmacological management option has proven its effect beyond reasonable doubt; thus, haloperidol may still be used when prevention and non-pharmacological interventions have failed.

There is high certainty of evidence that stress ulcer prophylaxis does not reduce mortality, although there is high certainty of evidence that any gastrointestinal bleeding and clinically important bleeding is reduced with stress ulcer prophylaxis in ICU patients. It is worth noting, however, that approximately 95% of patients in the ICU will not experience a bleeding period, when stress ulcer prophylaxis is not administered. This may indicate overutilization of medication. In relation to this, one trial indicated excess mortality when using pantoprazole for the most severely ill patients, with SAPS II score greater than 53. Thus, clinical practice may avoid the general use of stress ulcer prophylaxis and instead target the use for patients with clinical signs of bleeding and high risk of developing clinically significant bleeding. It is inconclusive whether stress ulcer prophylaxis influences the occurrence of serious adverse events, quality of life, pneumonia, myocardial ischemia and C. difficile infection.

With low to very low certainty of evidence, it is inconclusive whether higher versus lower fraction of inspired oxygen, or targets of arterial oxygenation, have beneficial or harmful effects in ICU patients and in critically ill patients. We found no beneficial effects of the use of higher levels of oxygen supplementation; consequently, higher levels may be avoided in routine clinical practice, but may still be used in protocolised randomised trials.

Patients admitted to the ICU are clinically heterogeneous. This heterogeneity, together with a general high clinical heterogeneity between the trials conducted in this setting, may challenge interpretation of randomised clinical trials and meta-analyses. Concluding on subgroup analysis in trials and systematic reviews reflects such challenges and Bayesian analysis may help avoid premature conclusions on sparse data in subgroup analysis.[15, 98, 138]

We developed CHIMS that covers four domains (setting, population, intervention, and outcome). Results of the reliability tests of CHIMS found it to be a reliable tool for assessing and quantifying clinical heterogeneity in meta-analyses. Clinical heterogeneity assessed with CHIMS seems to be higher in ICU meta-analyses compared to non-ICU meta-analyses. Moreover, we found no association between clinical heterogeneity, CHIMS, and statistical heterogeneity, I^2.

Our results, in addition to meta-epidemiological studies evaluating the evidence for interventions used in the intensive care setting, reveal that there is a need to map out the evidence for treatment modalities used in the intensive care, to help identify knowledge gaps and plan future research.

11. REFERENCES

1. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ 1996;312(7023):71-2.
2. Rosenberg WDonald A. Evidence based medicine: an approach to clinical problem-solving. BMJ 1995;310(6987):1122-6.
3. Evidence-based medicine working group. Evidence-based medicine. A new approach to teaching the practice of medicine. JAMA 1992;268(17):2420-5.
4. Garattini S, Jakobsen JC, Wetterslev J, Bertele V, Banzi R, Rath A, Neugebauer EA, Laville M, Masson Y, Hivert V, Eikermann M, Aydin B, Ngwabyt S, Martinho C, Gerardi C, Szmigielski CA, Demotes-Mainard J, Gluud C. Evidence-based clinical practice: Overview of threats to the validity of evidence and how to minimise them. Eur J Intern Med 2016;32:13-21.
5. Savovic J, Turner RM, Mawdsley D, Jones HE, Beynon R, Higgins JPT, Sterne JAC. Association between risk-of-bias assessments and results of randomized trials in Cochrane Reviews: The ROBES meta-epidemiologic study. Am J Epidemiol 2018;187(5):1113-1122.
6. Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. Ann Intern Med 2001;135(11):982-9.
7. Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, Gluud C, Martin RM, Wood AJ, Sterne JA. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. BMJ 2008;336(7644):601-5.
8. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, Elbourne D, Egger M, Altman DG. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c869.
9. Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.handbook.cochrane.org.
10. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, Savovic J, Schulz KF, Weeks L, Sterne JA. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928.
11. Higgins JPT, Savović J, Page MJ, Elbers RG, Sterne JAC. Chapter 8: Assessing risk of bias in a randomized trial. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook.
12. Gagnier JJ, Morgenstern H, Altman DG, Berlin J, Chang S, McCulloch P, Sun X, Moher D. Consensus-based recommendations for investigating clinical heterogeneity in systematic reviews. BMC Med Res Methodol 2013;13:106.
13. Gagnier JJ, Moher D, Boon H, Beyene J, Bombardier C. Investigating clinical heterogeneity in systematic reviews: a methodologic review of guidance in the literature. BMC Med Res Methodol 2012;12:111.
14. Deeks JJ, Higgins JPT, Altman DG (editors). Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook.
15. Borenstein MHiggins JP. Meta-analysis and subgroups. Prev Sci 2013;14(2):134-43.
16. Chess LEGagnier JJ. Applicable or non-applicable: investigations of clinical heterogeneity in systematic reviews. BMC Med Res Methodol 2016;16:19.
17. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook.
18. Wetterslev J, Thorlund K, Brok J, Gluud C. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. J Clin Epidemiol 2008;61(1):64-75.
19. Pogue JMYusuf S. Cumulating evidence from randomized trials: utilizing sequential monitoring boundaries for cumulative meta-analysis. Control Clin Trials 1997;18(6):580-93.

20. Thorlund K, Devereaux PJ, Wetterslev J, Guyatt G, Ioannidis JP, Thabane L, Gluud LL, Als-Nielsen B, Gluud C. Can trial sequential monitoring boundaries reduce spurious inferences from meta-analyses? Int J Epidemiol 2009;38(1):276-86.
21. Pogue JYusuf S. Overcoming the limitations of current meta-analysis of randomised controlled trials. Lancet 1998;351(9095):47-52.
22. Turner RM, Bird SM, Higgins JP. The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews. PLoS One 2013;8(3):e59202.
23. Koster TM, Wetterslev J, Gluud C, Jakobsen JC, Kaufmann T, Eck RJ, Koster G, Hiemstra B, van der Horst ICC, Keus E. Apparently conclusive meta-analyses on interventions in critical care may be inconclusive-a meta-epidemiological study. J Clin Epidemiol 2019;114:1-10.
24. Ioannidis JLau J. Evolution of treatment effects over time: empirical insight from recursive cumulative metaanalyses. Proc Natl Acad Sci U S A 2001;98(3):831-6.
25. Thorlund K, Engstrøm J, Wetterslev J, Brok J, Imberger I, Gluud C. User manual for Trial Sewquential Analysis (TSA). Copenhagen Trial Unit, Rigshospitalet, Denmark. Available from www.ctu.dk/tsa.
26. Wetterslev J, Jakobsen JC, Gluud C. Trial Sequential Analysis in systematic reviews with meta-analysis. BMC Med Res Methodol 2017;17(1):39.
27. Wetterslev J, Thorlund K, Brok J, Gluud C. Estimating required information size by quantifying diversity in random-effects model meta-analyses. BMC Med Res Methodol 2009;9:86.
28. Imberger G, Gluud C, Boylan J, Wetterslev J. Systematic reviews of anesthesiologic interventions reported as statistically significant: problems with power, precision, and type 1 error protection. Anesth Analg 2015;121(6):1611-22.
29. Mascha EJ. Alpha, beta, meta: guidelines for assessing power and type I error in meta-analyses. Anesth Analg 2015;121(6):1430-3.
30. Terkawi AS, Mavridis D, Flood P, Wetterslev J, Terkawi RS, Bin Abdulhak AA, Nunemaker MS, Tiouririne M. Does ondansetron modify sympathectomy due to subarachnoid anesthesia? Meta-analysis, meta-regression, and Trial Sequential Analysis. Anesthesiology 2016;124(4):846-69.
31. Kulinskaya EWood J. Trial sequential methods for meta-analysis. Res Synth Methods 2014;5(3):212-20.
32. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, Clarke M, Gargon E, Gorst S, Harman N, Kirkham JJ, McNair A, Prinsen CAC, Schmitt J, Terwee CB, Young B. The COMET Handbook: version 1.0. Trials 2017;18(Suppl 3):280.
33. Gaudry S, Messika J, Ricard JD, Guillo S, Pasquet B, Dubief E, Boukertouta T, Dreyfuss D, Tubach F. Patient-important outcomes in randomized controlled trials in critically ill patients: a systematic review. Ann Intensive Care 2017;7(1):28.
34. Desai SV, Law TJ, Needham DM. Long-term complications of critical care. Crit Care Med 2011;39(2):371-9.
35. Chelluri L, Im KA, Belle SH, Schulz R, Rotondi AJ, Donahoe MP, Sirio CA, Mendelsohn AB, Pinsky MR. Long-term mortality and quality of life after prolonged mechanical ventilation. Crit Care Med 2004;32(1):61-9.
36. van der Schaaf M, Beelen A, Dongelmans DA, Vroom MB, Nollet F. Poor functional recovery after a critical illness: a longitudinal study. J Rehabil Med 2009;41(13):1041-8.
37. Golder S, Loke YK, Wright K, Norman G. Reporting of adverse events in published and unpublished studies of health care interventions: a systematic review. PLoS Med 2016;13(9):e1002127.
38. Zorzela L, Loke YK, Ioannidis JP, Golder S, Santaguida P, Altman DG, Moher D, Vohra S. PRISMA harms checklist: improving harms reporting in systematic reviews. BMJ 2016;352:i157.
39. Parsons R, Golder S, Watt I. More than one-third of systematic reviews did not fully report the adverse events outcome. J Clin Epidemiol 2019;108:95-101.
40. Johnston BC, Patrick DL, Devji T, Maxwell LJ, Bingham III CO, Beaton D, Boers M, Briel M, Busse JW, Carrasco-Labra A, Christensen R, da Costa BR, El Dib R, Lyddiatt A, Ostelo RW, Shea B, Singh J, Terwee CB, Williamson PR, Gagnier JJ, Tugwell P, Guyatt GH. Chapter 18: Patient-reported outcomes. In: Higgins JPT, Thomas J, Chandler J, Cumpston
M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook.
41. Gifford JM, Husain N, Dinglas VD, Colantuoni E, Needham DM. Baseline quality of life before intensive care: a comparison of patient versus proxy responses. Crit Care Med 2010;38(3):855-60.

42. Scales DC, Tansey CM, Matte A, Herridge MS. Difference in reported pre-morbid health-related quality of life between ARDS survivors and their substitute decision makers. Intensive Care Med 2006;32(11):1826-31.
43. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schunemann HJ. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336(7650):924-6.
44. Page MJ, Shamseer L, Altman DG, Tetzlaff J, Sampson M, Tricco AC, Catala-Lopez F, Li L, Reid EK, Sarkis-Onofre R, Moher D. Epidemiology and reporting characteristics of systematic reviews of biomedical research: a cross-sectional study. PLoS Med 2016;13(5):e1002028.
45. Hunt H, Pollock A, Campbell P, Estcourt L, Brunton G. An introduction to overviews of reviews: planning a relevant research question and objective for an overview. Syst Rev 2018;7(1):39.
46. Pollock M, Fernandes RM, Becker LA, Pieper D, Hartling L. Chapter V: Overviews of reviews. Draft version (8 October 2018) for inclusion in : Higgins JPT, Thomas J, Chandler J, Cumpston MS, Li T, Page MJ, Welch V (editors). Cochrane Handbook for systematic reviews of interventions. London: Cochrane.
47. Chaimani A, Caldwell DM, Li T, Higgins JPT, Salanti G. Chapter 11: Undertaking network meta-analyses. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook.
48. Faltinsen EG, Storebo OJ, Jakobsen JC, Boesen K, Lange T, Gluud C. Network meta-analysis: the highest level of medical evidence? BMJ Evid Based Med 2018;23(2):56-59.
49. Marshall JC, Bosco L, Adhikari NK, Connolly B, Diaz JV, Dorman T, Fowler RA, Meyfroidt G, Nakagawa S, Pelosi P, Vincent JL, Vollman K, Zimmerman J. What is an intensive care unit? A report of the task force of the World Federation of Societies of Intensive and Critical Care Medicine. J Crit Care 2017;37:270-276.
50. Adhikari NK, Fowler RA, Bhagwanjee S, Rubenfeld GD. Critical care and the global burden of critical illness in adults. Lancet 2010;376(9749):1339-46.
51. Murthy S, Leligdowicz A, Adhikari NK. Intensive care unit capacity in low-income countries: a systematic review. PLoS One 2015;10(1):e0116949.
52. Mundy CJ, Bates I, Nkhoma W, Floyd K, Kadewele G, Ngwira M, Khuwi A, Squire SB, Gilks CF. The operation, quality and costs of a district hospital laboratory service in Malawi. Trans R Soc Trop Med Hyg 2003;97(4):403-8.
53. Gray IP, Carter JY. An evaluation of clinical laboratory services in sub-Saharan Africa. Ex africa semper aliquid novi? Clin Chim Acta 1997; 267(1): 103-28.
54. Arora N, Laha SK. The beginners guide to intensive care. Boca Raton: CRC Press; 2018.
55. Smith G, Nielsen M. ABC of intensive care. Criteria for admission. BMJ 1999; 318(7197):1544-7.
56. Smithburger PL, Buckley MS, Culver MA, Sokol S, Lat I, Handler SM, Kirisci L, Kane-Gill SL. A multicenter evaluation of off-label medication use and associated adverse drug reactions in adult medical ICUs. Crit Care Med 2015;43(8):1612-21.
57. Lat I, Micek S, Janzen J, Cohen H, Olsen K, Haas C. Off-label medication use in adult critical care patients. J Crit Care 2011;26(1):89-94.
58. Collet MO, Caballero J, Sonneville R, Bozza FA, Nydahl P, Schandl A, Woien H, Citerio G, van den Boogaard M, Hastbacka J, Haenggi M, Colpaert K, Rose L, Barbateskovic M, Lange T, Jensen A, Krog MB, Egerod I, Nibro HL, Wetterslev J, Perner A. Prevalence and risk factors related to haloperidol use for delirium in adult intensive care patients: the multinational AID-ICU inception cohort study. Intensive Care Med 2018;44(7):1081-1089.
59. Krag M, Perner A, Wetterslev J, Wise MP, Borthwick M, Bendel S, McArthur C, Cook D, Nielsen N, Pelosi P, Keus F, Guttormsen AB, Moller AD, Moller MH. Prevalence and outcome of gastrointestinal bleeding and use of acid suppressants in acutely ill adult intensive care patients. Intensive Care Med 2015;41(5):833-45.
60. European Medicines Agency - Committee for medicinal products for human use. Guideline on medicinal gases. Available at https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-medicinal-gases-pharmaceutical-documentation-including-recommendation-non-clinical-safety_en.pdf [accessed 30 December 2019].
61. Damiani E, Adrario E, Girardis M, Romano R, Pelaia P, Singer M, Donati A. Arterial hyperoxia and mortality in critically ill patients: a systematic review and meta-analysis. Crit Care 2014;18(6):711.

62. Helmerhorst HJ, Roos-Blom MJ, van Westerloo DJ, de Jonge E. Association between arterial hyperoxia and outcome in subsets of critical illness: a systematic review, meta-analysis, and meta-regression of cohort Studies. Crit Care Med 2015;43(7):1508-19.
63. Gomersall CD, Joynt GM, Freebairn RC, Lai CK, Oh TE. Oxygen therapy for hypercapnic patients with chronic obstructive pulmonary disease and acute respiratory failure: a randomized, controlled pilot study. Crit Care Med 2002;30(1):113-6.
64. Panwar R, Hardie M, Bellomo R, Barrot L, Eastwood GM, Young PJ, Capellier G, Harrigan PW, Bailey M. Conservative versus liberal oxygenation targets for mechanically ventilated patients. A pilot multicenter randomized controlled trial. Am J Respir Crit Care Med 2016;193(1):43-51.
65. Girardis M, Busani S, Damiani E, Donati A, Rinaldi L, Marudi A, Morelli A, Antonelli M, Singer M. Effect of conservative vs conventional oxygen therapy on mortality among patients in an intensive care unit: the oxygen-ICU randomized clinical trial. JAMA 2016;316(15):1583-1589.
66. Asfar P, Schortgen F, Boisrame-Helms J, Charpentier J, Guerot E, Megarbane B, Grimaldi D, Grelon F, Anguel N, Lasocki S, Henry-Lagarrigue M, Gonzalez F, Legay F, Guitton C, Schenck M, Doise JM, Devaquet J, Van Der Linden T, Chatellier D, Rigaud JP, Dellamonica J, Tamion F, Meziani F, Mercat A, Dreyfuss D, Seegers V, Radermacher P. Hyperoxia and hypertonic saline in patients with septic shock (HYPERS2S): a two-by-two factorial, multicentre, randomised, clinical trial. Lancet Respir Med 2017;5(3):180-190.
67. Suzuki S, Eastwood GM, Glassford NJ, Peck L, Young H, Garcia-Alvarez M, Schneider AG, Bellomo R. Conservative oxygen therapy in mechanically ventilated patients: a pilot before-and-after trial. Crit Care Med 2014;42(6):1414-22.
68. Helmerhorst HJ, Schultz MJ, van der Voort PH, Bosman RJ, Juffermans NP, de Wilde RB, van den Akker-van Marle ME, van Bodegom-Vos L, de Vries M, Eslami S, de Keizer NF, Abu-Hanna A, van Westerloo DJ, de Jonge E. Effectiveness and clinical outcomes of a two-step implementation of conservative oxygenation targets in critically ill patients: a before and after trial. Crit Care Med 2016;44(3):554-63.
69. Nguyen TL, Collins GS, Lamy A, Devereaux PJ, Daures JP, Landais P, Le Manach Y. Simple randomization did not protect against bias in smaller trials. J Clin Epidemiol 2017;84:105-113.
70. Vincent JL. Improved survival in critically ill patients: are large RCTs more useful than personalized medicine? No. Intensive Care Med 2016;42(11):1778-1780.
71. Vincent JL. We should abandon randomized controlled trials in the intensive care unit. Crit Care Med 2010;38(10 Suppl):S534-8.
72. Vincent JL, Marini JJ, Pesenti A. Do trials that report a neutral or negative treatment effect improve the care of critically ill patients? No. Intensive Care Med 2018;44(11):1989-1991.
73. Barbateskovic M, Larsen LK, Oxenboll-Collet M, Jakobsen JC, Perner A, Wetterslev J. Pharmacological interventions for delirium in intensive care patients: a protocol for an overview of reviews. Syst Rev 2016;5(1):211.
74. Barbateskovic M, Kraus SR, Collet MO, Mathiesen O, Jakobsen JC, Perner A, Wetterslev J. Haloperidol for delirium in critically ill patients - protocol for a systematic review. Acta Anaesthesiol Scand 2018;62(5):712-723.
75. Barbateskovic M, Schjørring OL, Jakobsen JC, Meyhoff CS, Dahl RM, Rasmussen BS, Perner A, Wetterslev J. Higher versus lower inspiratory oxygen fraction or targets of arterial oxygenation for adult intensive care patients. Cochrane Database of Systematic Reviews 2017, Issue 4. Art. No.: CD012631. DOI: 10.1002/14651858.CD012631.
76. Barbateskovic M, Schjorring OL, Jakobsen JC, Meyhoff CS, Rasmussen BS, Perner A, Wetterslev J. Oxygen supplementation for critically ill patients-A protocol for a systematic review. Acta Anaesthesiol Scand 2018;62(7):1020-1030.
77. Barbateskovic M, Marker S, Jakobsen JC, Krag M, Granholm A, Anthon CT, Perner A, Wetterslev J, Moller MH. Stress ulcer prophylaxis in adult intensive care unit patients - a protocol for a systematic review. Acta Anaesthesiol Scand 2018;62(6):744-755.
78. Barbateskovic M, Koster TM, Keus F, Gluud C, Møller MH, van der Horst ICC, Perner A, Wetterslev J. A protocol for constructing a tool to assess clinical heterogeneity in meta-analyses, assessment of interrater variability, and a pilot study of the association between clinical and statistical heterogeneity, Copenhagen Trial Unit 2019, http://ctu.dk/media/13724/2019-protocol-chims-protocol-manual-ver-11.0_11-03-2019.pdf

79. Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. BMC Med Res Methodol 2014;14:120.
80. International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use (ICH) adopts consolidated guideline on good clinical practice in the conduct of clinical trials on medicinal products for human use. International digest of health legislation 1997; 48(2):231–4. [PUBMED: 11656783].
81. Review Manager 2014 [Computer program]. Nordic Cochrane Cenre, The Cochrane Collaboration. Review Manager 5 (RevMan 5). Version 5.3. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2014.
82. TSA 2011 [Computer program]. Copenhagen Trial Unit. TSA - Trial Sequential Analysis. Version 0.9 Beta. Copenhagen: Copenhagen Trial Unit, 2011.
83. GRADE Pro GDT [Computer program]. McMaster University. GRADEpro GDT. Hamilton (ON): McMaster University (developed by Evidence Prime), 2015.
84. Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hrobjartsson A, Roberts C, Shoukri M, Streiner DL. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. J Clin Epidemiol 2011;64(1):96-106.
85. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557-60.
86. Landis JRKoch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159-74.
87. Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, Cates CJ, Cheng HY, Corbett MS, Eldridge SM, Emberson JR, Hernan MA, Hopewell S, Hrobjartsson A, Junqueira DR, Juni P, Kirkham JJ, Lasserson T, Li T, McAleenan A, Reeves BC, Shepperd S, Shrier I, Stewart LA, Tilling K, White IR, Whiting PF, Higgins JPT. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019;366:l4898.
88. Thomas J, Askie LM, Berlin JA, Elliott JH, Ghersi D, Simmonds M, Takwoingi Y, Tierney JF, Higgins HPT. Chapter 22: Prospective approaches to accumulating evidence. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook.
89. Rose L, Agar M, Burry LD, Campbell N, Clarke M, Lee J, Siddiqi N, Page VJ. Development of core outcome sets for effectiveness trials of interventions to prevent and/or treat delirium (Del-COrS): study protocol. BMJ Open 2017;7(9):e016371.
90. NCT03392376, Agents Intervening Against Delirium in Intensive Care Unit (AID-ICU), https://clinicaltrials.gov/ct2/show/NCT03392376. Accessed 17 December 2019.
91. NCT03628391, Efficacy of haloperidol to decrease the burden of delirium in adult critically ill patients (EuRIDICE): a prospective randomised multi-center double-blind placebo-controlled clinicaltTrial, https://clinicaltrials.gov/ct2/show/NCT03628391. Accessed 17 December 2019.
92. Andersen-Ranberg NC, Poulsen LM, Perner A, Wetterslev J, Estrup S, Lange T, Ebdrup BH, Hastbacka J, Morgan MPG, Citerio G, Zafrani L, Caballero J, Oxenboll-Collet M, Weber SO, Andreasen AS, Bestle M, Pedersen HBS, Hildebrandt T, Thee C, Jensen TB, Dey N, Nielsen LG, Mathiesen O. Agents intervening against delirium in the intensive care unit (AID-ICU) - Protocol for a randomised placebo-controlled trial of haloperidol in patients with delirium in the ICU. Acta Anaesthesiol Scand 2019;63(10):1426-1433.
93. Bannon L, McGaughey J, Verghis R, Clarke M, McAuley DF, Blackwood B. The effectiveness of non-pharmacological interventions in reducing the incidence and duration of delirium in critically ill patients: a systematic review and meta-analysis. Intensive Care Med 2019;45(1):1-12.
94. Burry L, Hutton B, Williamson DR, Mehta S, Adhikari NKJ, Cheng W, Ely EW, Egerod I, Fergusson DA, Rose L. Pharmacological interventions for the treatment of delirium in critically ill adults. Cochrane Database of Systematic Reviews 2019, Issue 9. Art. No.: CD011749. DOI: 10.1002/14651858.CD011749.pub2.
95. Marker S, Barbateskovic M, Perner A, Wetterslev J, Jakobsen JC, Krag M, Granholm A, Anthon CT, Møller MH. Prophylactic use of acid suppressants in adult acutely ill hospitalised patients: a systematic review with meta-analysis and Trial Sequential Analysis. [Submitted manuscript].
96. Krag M, Marker S, Perner A, Wetterslev J, Wise MP, Schefold JC, Keus F, Guttormsen AB, Bendel S, Borthwick M, Lange T, Rasmussen BS, Siegemund M, Bundgaard H, Elkmann T, Jensen JV, Nielsen

RD, Liboriussen L, Bestle MH, Elkjaer JM, Palmqvist DF, Backlund M, Laake JH, Badstolokken PM, Gronlund J, Breum O, Walli A, Winding R, Iversen S, Jarnvig IL, White JO, Brand B, Madsen MB, Quist L, Thornberg KJ, Moller A, Wiis J, Granholm A, Anthon CT, Meyhoff TS, Hjortrup PB, Aagaard SR, Andreasen JB, Sorensen CA, Haure P, Hauge J, Hollinger A, Scheuzger J, Tuchscherer D, Vuilliomenet T, Takala J, Jakob SM, Vang ML, Paelestik KB, Andersen KLD, van der Horst ICC, Dieperink W, Fjolner J, Kjer CKW, Solling C, Solling CG, Karttunen J, Morgan MPG, Sjobo B, Engstrom J, Agerholm-Larsen B, Moller MH. Pantoprazole in Patients at Risk for Gastrointestinal Bleeding in the ICU. N Engl J Med 2018;379(23):2199-2208.

97. Marker S, Perner A, Wetterslev J, Krag M, Lange T, Wise MP, Borthwick M, Bendel S, Keus F, Guttormsen AB, Schefold JC, Moller MH. Pantoprazole prophylaxis in ICU patients with high severity of disease: a post hoc analysis of the placebo-controlled SUP-ICU trial. Intensive Care Med 2019;45(5):609-618.
98. Granholm A, Marker S, Krag M, Zampieri FG, Thorsen-Meyer H, Kaas-Hansen BS, van der Horst ICC, Lange T, Wetterslev J, Perner A, Møller MH. Heterogeneity of treatment effect of stress ulcer prophylaxis in adult ICU patients: a secondary analysis of the SUP-ICU trial. Intensive Care Med [accepted for publication].
99. Granholm A, Zeng L, Dionne JC, Perner A, Marker S, Krag M, MacLaren R, Ye Z, Moller MH, Alhazzani W. Predictors of gastrointestinal bleeding in adult ICU patients: a systematic review and meta-analysis. Intensive Care Med 2019;45(10):1347-1359.
100. Needham DM, Sepulveda KA, Dinglas VD, Chessare CM, Friedman LA, Bingham CO, Turnbull AE. Core outcome measures for clinical research in acute respiratory failure survivors. An international modified delphi consensus study. Am J Respir Crit Care Med 2017;196(9):1122-1130.
101. Blackwood B, Ringrow S, Clarke M, Marshall JC, Connolly B, Rose L, McAuley DF. A core outcome set for critical care ventilation trials. Crit Care Med 2019;47(10):1324-1331.
102. NCT02321072. Optimal oxygenation in the intensive care unit (O2-ICU). clinicaltrials.gov/ct2/show/NCT02321072 (first received 9 December 2014).
103. NCT02713451. Liberal oxygenation versus conservative oxygenation in ARDS (LOCO2). clinicaltrials.gov/ct2/show/NCT02713451 (first received 18 March 2016).
104. NCT03141099. Blood pressure and OXygenation targets after OHCA (BOX). clinicaltrials.gov/ct2/show/NCT03141099 (first received 4 May 2017).
105. NCT03174002. Handling oxygenation targets in the intensive care unit (HOT-ICU). clinicaltrials.gov/ct2/show/NCT03174002 (first received 2 June 2017).
106. NCT03287466. Targeted OXYgen therapy in critical illness (TOXYC). clinicaltrials.gov/ct2/show/NCT03287466 (first received 19 September 2017).
107. Barbateskovic M, Schjørring OL, Russo Krauss S, Meyhoff CS, Jakobsen JC, Rasmussen BS, Perner A, Wetterslev J. Higher versus lower levels of oxygen supplementation in critically ill patients. A systematic review wth meta-analysis and Trial Sequential Analysis. [Submitted manuscript]
108. Rhodes A, Evans LE, Alhazzani W, Levy MM, Antonelli M, Ferrer R, Kumar A, Sevransky JE, Sprung CL, Nunnally ME, Rochwerg B, Rubenfeld GD, Angus DC, Annane D, Beale RJ, Bellinghan GJ, Bernard GR, Chiche JD, Coopersmith C, De Backer DP, French CJ, Fujishima S, Gerlach H, Hidalgo JL, Hollenberg SM, Jones AE, Karnad DR, Kleinpell RM, Koh Y, Lisboa TC, Machado FR, Marini JJ, Marshall JC, Mazuski JE, McIntyre LA, McLean AS, Mehta S, Moreno RP, Myburgh J, Navalesi P, Nishida O, Osborn TM, Perner A, Plunkett CM, Ranieri M, Schorr CA, Seckel MA, Seymour CW, Shieh L, Shukri KA, Simpson SQ, Singer M, Thompson BT, Townsend SR, Van der Poll T, Vincent JL, Wiersinga WJ, Zimmerman JL, Dellinger RP. Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock: 2016. Intensive Care Med 2017;43(3):304-377.
109. Meyhoff TS, Møller MH, Hjortrup PB, Cronhjort M, Perner A, Wetterslev J. Lower versus higher fluid volumes during initial management of sepsis – a systematic review with meta-analysis and trial sequential analysis. Chest [accepted for publication]
110. Berthelsen RE, Perner A, Jensen AK, Rasmussen BS, Jensen JU, Wiis J, Behzadi MT, Bestle MH. Forced fluid removal in intensive care patients with acute kidney injury: The randomised FFAKI feasibility trial. Acta Anaesthesiol Scand 2018;62(7):936-944.
111. Wetterslev M, Haase N, Hassager C, Belley-Cote EP, McIntyre WF, An Y, Shen J, Cavalcanti AB, Zampieri FG, Guimaraes HP, Granholm A, Perner A, Moller MH. New-onset atrial fibrillation in adult critically ill patients: a scoping review. Intensive Care Med 2019;45(7):928-938.
112. Feinberg J, Nielsen EE, Korang SK, Halberg Engell K, Nielsen MS, Zhang K, Didriksen M, Lund L, Lindahl N, Hallum S, Liang N, Xiong W, Yang X, Brunsgaard P, Garioud A, Safi S, Lindschou J,

Kondrup J, Gluud C, Jakobsen JC. Nutrition support in hospitalised adults at nutritional risk. Cochrane Database of Systematic Reviews 2017, Issue 5. Art. No.: CD011598. DOI: 10.1002/14651858.CD011598.pub2.

113. Allingstrup MJ, Kondrup J, Wiis J, Claudius C, Pedersen UG, Hein-Rasmussen R, Bjerregaard MR, Steensen M, Jensen TH, Lange T, Madsen MB, Moller MH, Perner A. Early goal-directed nutrition versus standard of care in adult intensive care patients: the single-centre, randomised, outcome assessor-blinded EAT-ICU trial. Intensive Care Med 2017;43(11):1637-1647.
114. Eck RJ, Bult W, Wetterslev J, Gans ROB, Meijer K, van der Horst ICC, Keus F. Low dose low-molecular-weight heparin for thrombosis prophylaxis: systematic review with meta-analysis and Trial Sequential Analysis. J Clin Med 2019;8(12)
115. Sjovall F, Perner A, Hylander Moller M. Empirical mono- versus combination antibiotic therapy in adult intensive care patients with severe sepsis - A systematic review with meta-analysis and trial sequential analysis. J Infect 2017;74(4):331-344.
116. Koster TM, Wetterslev J, Gluud C, Keus F, van der Horst ICC. Systematic overview and critical appraisal of meta-analyses of interventions in intensive care medicine. Acta Anaesthesiol Scand 2018;62:1041-49.
117. West SL, Gartlehner G, Mansfield AJ, Poole C, Tant E, Lenfestey N, Lux LJ, Amoozegar J, Morton SC, Carey TC, Viswanathan M, Lohr KN. Comparative effectiveness review methods: clinical heterogeneity. Agency for Healthcare Research and Quality 2010. Available from http://effectivehealthcare.ahrq.gov/.
118. Barbateskovic M, Krauss SR, Collet MO, Andersen-Ranberg NC, Mathiesen O, Jakobsen JC, Perner A, Wetterslev J. Haloperidol for the treatment of delirium in critically ill patients: A systematic review with meta-analysis and Trial Sequential Analysis. Acta Anaesthesiol Scand 2019 doi: 10.1111/aas.13501. [Epub ahead of print];
119. Barbateskovic M, Marker S, Granholm A, Anthon CT, Krag M, Jakobsen JC, Perner A, Wetterslev J, Moller MH. Stress ulcer prophylaxis with proton pump inhibitors or histamin-2 receptor antagonists in adult intensive care patients: a systematic review with meta-analysis and trial sequential analysis. Intensive Care Med 2019;45(2):143-158.
120. Barbateskovic M, Schjørring OL, Russo Krauss S, Jakobsen JC, Meyhoff CS, Dahl RM, Rasmussen BS, Perner A, Wetterslev J. Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the intensive care unit. Cochrane Database of Systematic Reviews 2019, Issue 11. Art. No.: CD012631. DOI: 10.1002/14651858.CD012631.pub2.
121. von Hippel PT. The heterogeneity statistic I(2) can be biased in small meta-analyses. BMC Med Res Methodol 2015;15:35.
122. Ioannidis JP, Patsopoulos NA, Evangelou E. Uncertainty in heterogeneity estimates in meta-analyses. BMJ 2007;335(7626):914-6.
123. Rucker G, Schwarzer G, Carpenter JR, Schumacher M. Undue reliance on I(2) in assessing heterogeneity may mislead. BMC Med Res Methodol 2008;8:79.
124. Langan D, Higgins JP, Simmonds M. An empirical comparison of heterogeneity variance estimators in 12 894 meta-analyses. Res Synth Methods 2015;6(2):195-205.
125. Langan D, Higgins JPT, Jackson D, Bowden J, Veroniki AA, Kontopantelis E, Viechtbauer W, Simmonds M. A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. Res Synth Methods 2019;10(1):83-98.
126. IntHout J, Ioannidis JP, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Med Res Methodol 2014;14:25.
127. Cornell JE, Mulrow CD, Localio R, Stack CB, Meibohm AR, Guallar E, Goodman SN. Random-effects meta-analysis of inconsistent effects: a time for change. Ann Intern Med 2014;160(4):267-70.
128. Thorlund K, Wetterslev J, Awad T, Thabane L, Gluud C. Comparison of statistical inferences from the DerSimonian-Laird and alternative random-effects model meta-analyses - an empirical assessment of 920 Cochrane primary outcome meta-analyses. Res Synth Methods 2011;2(4):238-53.
129. Djurisic S, Rath A, Gaber S, Garattini S, Bertele V, Ngwabyt SN, Hivert V, Neugebauer EAM, Laville M, Hiesmayr M, Demotes-Mainard J, Kubiak C, Jakobsen JC, Gluud C. Barriers to the conduct of randomised clinical trials within all disease areas. Trials 2017;18(1):360.
130. Ecarnot F, Quenot JP, Besch G, Piton G. Ethical challenges involved in obtaining consent for research from patients hospitalized in the intensive care unit. Ann Transl Med 2017;5(Suppl 4):S41.

131. Neugebauer EAM, Rath A, Antoine SL, Eikermann M, Seidel D, Koenen C, Jacobs E, Pieper D, Laville M, Pitel S, Martinho C, Djurisic S, Demotes-Mainard J, Kubiak C, Bertele V, Jakobsen JC, Garattini S, Gluud C. Specific barriers to the conduct of randomised clinical trials on medical devices. Trials 2017;18(1):427.
132. Temple R. Enrichment of clinical study populations. Clin Pharmacol Ther 2010;88(6):774-8.
133. Santacruz CA, Pereira AJ, Celis E, Vincent JL. Which Multicenter Randomized Controlled Trials in Critical Care Medicine Have Shown Reduced Mortality? A Systematic Review. Crit Care Med 2019;47(12):1680-1691.
134. Blackwood B, Marshall J, Rose L. Progress on core outcome sets for critical care research. Curr Opin Crit Care 2015;21(5):439-44.
135. Borm GF, Lemmers O, Fransen J, Donders R. The evidence provided by a single trial is less reliable than its statistical analysis suggests. J Clin Epidemiol 2009;62(7):711-715.e1.
136. IntHout J, Ioannidis JP, Borm GF. Obtaining evidence by a single well-powered trial or several modestly powered trials. Stat Methods Med Res 2016;25(2):538-52.
137. Herbison P, Hay-Smith J, Gillespie WJ. Meta-analyses of small numbers of trials often agree with longer-term results. J Clin Epidemiol 2011;64(2):145-53.
138. Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR. Bayesian methods in health technology assessment: a review. Health Technol Assess 2000;4(38):1-130.

12. SUPPLEMENTARY MATERIALS

PAPER I: Pharmacological interventions for prevention and management of delirium in intensive care patients: A systematic overview of reviews and meta-analyses

PAPER II: Haloperidol for the treatment of delirium in critically ill patients: A systematic review with meta-analysis and Trial Sequential Analysis

PAPER III: Stress ulcer prophylaxis with proton pump inhibitors or histamin-2 receptor antagonists in adult intensive care patients: A systematic review with meta-analysis and Trial Sequential Analysis

PAPER V: Higher vs lower levels of oxygen supplementation in critically ill patients: A systematic review with meta-analysis and Trial Sequential Analysis

PAPER VI: Clinical Heterogeneity In a Meta-analysis Score - Appendix A

Clinical Heterogeneity In a Meta-analysis Score - Appendix B

Clinical Heterogeneity In a Meta-analysis Score - Appendix C

Supplementary materials for paper I, II, III, V and VI can be accessed from:

http://www.ctu.dk/publications/supplementary-material/marija-barbateskovic-phd-thesis-supplementary-material.aspx

13. PAPERS

I. **Pharmacological interventions for prevention and management of delirium in intensive care patients: a systematic overview of reviews and meta-analyses.** Barbateskovic M, Krauss SR, Collet MO, Larsen LK, Jakobsen JC, Perner A, Wetterslev J. BMJ Open 2019;9(2):e024562.

II. **Haloperidol for the treatment of delirium in critically ill patients: a systematic review with meta-analysis and Trial Sequential Analysis.** Barbateskovic M, Krauss SR, Collet MO, Andersen-Ranberg NC, Mathiesen O, Jakobsen JC, Perner A, Wetterslev J. Acta Anaesthesiol Scand 2019. doi: 10.1111/aas.13501. [Epub ahead of print].

III. **Stress ulcer prophylaxis with proton pump inhibitors or histamin-2 receptor antagonists in adult intensive care patients: a systematic review with meta-analysis and Trial Sequential Analysis.** Barbateskovic M, Marker S, Granholm A, Anthon CT, Krag M, Jakobsen JC, Perner A, Wetterslev J, Møller MH. Intensive Care Med 2019; 45(2):143-158.

IV. **Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the intensive care unit.** Barbateskovic M, Schjørring OL, Krauss SR, Jakobsen JC, Meyhoff CS, Dahl RM, Rasmussen BS, Perner A, Wetterslev J. Cochrane Database of Systematic Reviews 2019, Issue 11. Art. No.: CD012631. DOI: 10.1002/14651858.CD012631.pub2.

V. **Higher versus lower levels of oxygen supplementation in critically ill adults. A systematic review with meta-analysis and Trial Sequential Analysis**. Barbateskovic M, Schjørring OL, Krauss SR, Jakobsen JC, Meyhoff CS, Rasmussen BS, Perner A, Wetterslev J. [In review]

VI. **CHIMS: Clinical Heterogeneity In Meta-Analysis Score - a new tool to assess and quantify clinical heterogeneity in meta-analyses of interventions.** Barbateskovic M, Koster TM, Eck RJ, Maagaard M, Afshari A, Blokzijl F, Cronhjort M, Dieperink W, Fabritius ML, Feinberg J, French C, Gareb B, Geisler A, Granholm A, Hiemstra B, Hu R, Imberger G, Jensen BT, Jonsson AB, Karam O, Kong DZ, Korang SK, Koster G, Lai B, Liang N, Lundstrøm LH, Marker S, Meyhoff T, Nielsen EE, Nørskov AK, Petersen MW, Risom EC, Rygård SL, Safi S, Sethi N, Sjövall F, Lauridsen SV, van Bakelen N, Volbeda M, van der Horst ICC, Gluud C, Perner A, Møller MH, Keus F, Wetterslev J. [In review]

PAPER I

BMJ Open

Pharmacological interventions for prevention and management of delirium in intensive care patients: a systematic overview of reviews and meta-analyses

Marija Barbateskovic,[1,2] Sara Russo Krauss,[1] Marie Oxenboell Collet,[2,3] Laura Krone Larsen,[4] Janus Christian Jakobsen,[1,2,5] Anders Perner,[2,3] Jørn Wetterslev[1,2]

To cite: Barbateskovic M, Krauss SR, Collet MO, *et al*. Pharmacological interventions for prevention and management of delirium in intensive care patients: a systematic overview of reviews and meta-analyses. *BMJ Open* 2019;**9**:e024562. doi:10.1136/bmjopen-2018-024562

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2018-024562).

Received 1 June 2018
Revised 7 September 2018
Accepted 7 December 2018

For numbered affiliations see end of article.

Correspondence to
Marija Barbateskovic;
marija.barbateskovic@ctu.dk

ABSTRACT

Objectives We assessed the evidence from reviews and meta-analyses of randomised clinical trials on the effects of pharmacological prevention and management of delirium in intensive care unit (ICU) patients.

Methods We searched for reviews in July 2017 in: Cochrane Library, MEDLINE, Embase, Science Citation Index, BIOSIS Previews, CINAHL and LILACS. We assessed whether reviews were systematic according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and assessed the methodological quality using ROBIS.

Outcome measures Primary outcomes: all-cause mortality, serious adverse events, prevention of delirium and management of delirium. Secondary outcomes: quality of life; non-serious adverse events and cognitive function.

Results We included 378 reviews: 369 narrative reviews, eight semisystematic reviews which failed on a maximum of two arbitrary PRISMA criteria and one systematic review fulfilling all 27 PRISMA criteria. For the prevention of delirium, we identified the one systematic review and eight semisystematic reviews all assessing the effects of alpha-2-agonists. None found evidence of a reduction of mortality (systematic review RR 0.99, 95% CI 0.79 to 1.24). The systematic review and three semisystematic reviews found no evidence of an effect for the prevention of delirium (systematic review RR 0.85, 0.63 to 1.14). Conversely, four semisystematic reviews found a beneficial effect. Serious adverse events, quality of life, non-serious adverse events and cognitive function were not assessed. We did not identify any systematic or semisystematic reviews addressing other pharmacological interventions for the prevention of delirium. For the management of manifest delirium, we did not identify any systematic or semisystematic review assessing any pharmacological agents.

Conclusion Based on systematic reviews, the evidence for the use of pharmacological interventions for prevention or management of delirium is poor or sparse. A systematic review with low risk of bias assessing the effects of pharmacological prevention of delirium and management of manifest delirium in ICU patients is urgently needed.

PROSPERO registration number CRD42016046628.

Strengths and limitations of this study

- We used a transparent and systematic method which followed widely accepted methodological standards.
- We conducted a thorough and comprehensive literature search.
- Preferred Reporting Items for Systematic Reviews and Meta-Analyses was chosen as the gold standard for defining a systematic review.
- We did not search for individual trials or performed meta-analyses and Trial Sequential Analysis within each of the groups of pharmacological agents.

INTRODUCTION

Delirium is a complex acute organic syndrome characterised by a reduced ability to focus, sustain or shift attention, and either a change in cognition or the development of perceptual disturbances.[1] Delirium is classified in motoric subtypes: (1) hypoactive delirium; (2) hyperactive delirium and (3) a mixed form delirium. Hypoactive and mixed delirium are most common in intensive care unit (ICU) patients,[2 3] and hypoactive delirium has been suggested to have worse outcomes.[4] In ICU patients, 25% to 89% are reported to be affected by delirium, which is associated with increased mortality in these patients.[5–9] Furthermore, delirium is associated with increased morbidity, including increased duration of mechanical ventilation, and ICU and hospital length of stay.[6 10–16] Patients with delirium may experience functional decline after ICU discharge and long-term cognitive impairment.[11 12 15]

Up-to-date critical care guidelines recommend non-pharmacological strategies in both the prevention and management of manifest delirium.[17] These strategies may include early mobilisation and reorientation

of the patient, risk factor assessment and normalisation of the sleep–wake cycle.[18] When delirium is suspected or identified, guidelines suggest that patients should be evaluated to identify potential underlying causes, allowing for deficiencies to be corrected, or exposures to be removed. Only when non-drug methods have failed to control symptoms should pharmacological interventions be used.[19 20] Nonetheless, a recently performed inception cohort study found that haloperidol was used as management option in 46% of ICU patients diagnosed with delirium, and dexmedetomidine in 21%.[16]

Pharmacological interventions for delirium have focused on alterations in neurotransmitter pathways, in particular dopaminergic and cholinergic pathways. Several pharmacological strategies have been used against delirium in the ICU patients: antipsychotics; sedatives; cholinesterase inhibitors; opioids; and melatonin and melatonin antagonists. Haloperidol is considered the drug of choice when managing manifest delirium in ICU settings[21–25] and some international guidelines recommend haloperidol in the management of manifest ICU delirium.[19 26 27] However, the two latest iterations of the guideline by the American College of Critical Care Medicine and the Society of Critical Care Medicine no longer recommend managing delirium with haloperidol due to lack of evidence.[17 28] In general, pharmacological interventions are not recommended for the prevention of delirium in ICU patients.[19 26–28]

Systematic reviews and meta-analyses have become one of the most widely used methods to quantify the effects of medical interventions and are frequently being recognised as the best available evidence for decisions about healthcare management and policy.[29 30] A preliminary search identified several reviews investigating the effects of pharmacological interventions for the prevention and management of delirium. However, uncertainty regarding the benefits and harms of pharmacological interventions appeared to be considerable, and trials have shown either positive,[31 32] equipoise[33 34] or negative results.[35]

The objective of this overview of reviews was to systematically and critically assess the quantity and the quality of the available reviews and meta-analyses of randomised clinical trials on the effects of pharmacological prevention and management of delirium in ICU patients.

METHODS

We conducted this systematic overview of reviews with a registered (PROSPERO CRD42016046628) and published protocol,[36] in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Electronic supplementary material (ESM), table 1). We used the systematic review methods principles outlined in the Cochrane Handbook[37] and the recommendations given by Robinson *et al*.[38]

Criteria for considering reviews for inclusion

We included all reviews and meta-analyses of pharmacological interventions for the prevention of delirium or management of manifest delirium (defined as diagnosed delirium) in adult ICU patients. We predefined a systematic review as a review positively fulfilling the PRISMA reporting guidelines.[39]

We defined adult ICU patients as those treated in an ICU (or similar terms defined by the review authors) of any specialty, for example, medical, surgical, trauma, cardiac. We included reviews of ICU patients aged 18 years or older and included both acute surgery patients and elective cardiac surgery patients.

Table 1 Summary of risk of bias assessment of the single systematic review and the eight semisystematic reviews using ROBIS

		ROBIS Phase 2				ROBIS Phase 3
Review	Violated PRISMA criteria	Study eligibility criteria	Identification and selection of studies	Data collection and study appraisal	Synthesis and findings	Overall risk of bias in the review
Tan *et al*[53]	#4; #5	☹	☹	☺	☹	☹
Lin *et al*[51]	#5; #27	☹	☹	☹	☹	☹
Fraser[52]	#5; #8	☹	☹	☺	☹	☹
Xia *et al*[47]	#5	☹	☹	☺	☹	☹
Zhang *et al*[48]	#5	☹	☹	☺	☹	☹
Pasin *et al*[50]	#5; #27	☹	☹	☺	☹	☹
Chen *et al*[46]	0	☺	☺	☺	☺	☺
Tran *et al*[54]	#15; #22	☺	☺	☹	☹	☹
Liu *et al*[49]	#5	☹	☹	☹	☹	☹

#4, objectives; #5, protocol and registration; #8, search; 15, risk of bias across studies (methods); #22, risk of bias across studies (results); #27, funding; ☹, low risk; ☺, high risk.

We excluded reviews on ICU patients with delirium caused by alcohol withdrawal, terminally ill patients, patients admitted to emergency departments and elective surgery patients, except cardiac surgery.

Results on all primary and secondary outcomes of the included systematic reviews were a priori planned to be reported.[36] However, we defined the primary and secondary outcomes in this overview of reviews as follows[36]:

Primary outcomes

1. All-cause mortality
2. Proportion of participants with a serious adverse event, defined as an event (experience) or reaction in any untoward medical occurrence that at any dose results in death, is life-threatening, requires prolongation of hospitalisation or results in persistent or significant disability/incapacity[40]
3. Proportion of participants with resolution of delirium symptom at end of treatment (management of delirium) and proportion of participants with delirium despite the administration of a pharmacological agent before being diagnosed with delirium (prevention of delirium)

Secondary outcomes

1. Quality of life as defined by review authors (eg, measured with SF36)[41]
2. Proportion of participants with non-serious adverse events defined as adverse events which are not serious
3. Cognitive function as defined by review authors (eg, measured with Repeatable Battery for the Assessment of Neuropsychological Status)[42] (continuous score)

Search methods for identification of reviews

We searched the Cochrane Library, MEDLINE (OvidSP), Embase (OvidSP), Science Citation Index-Expanded (Web of Science), BIOSIS Previews (Web of Science), Cumulative Index to Nursing & Allied Health Literature (CINAHL), Latin American Caribbean Health Sciences Literature (LILACS) and Allied and Complementary Medicine Database (AMED) in July 2017, in order to identify reviews eligible for inclusion. Full search strategies and time spans of the searches are provided in electronic supplementary material—ESM.

Data collection and analysis

Four authors (MB, SRK, MOC, LKL) independently screened the titles and abstracts of all reports identified in the searches using Covidence and comparison was made within pairs.[43] Reports deemed potentially relevant by any of the review authors were obtained in full text, and the full-text papers were assessed for eligibility by two review authors independently before being assessed for inclusion and compared within pairs. Disagreements were resolved by consensus. Reviews containing a methods section and/or a literature search were hereafter checked against the PRISMA criteria.[39] Initially, it was our intention to only include systematic reviews fulfilling all 27 PRISMA criteria, but we decided pragmatically to define a group of reviews which failed on a maximum of two arbitrary PRISMA criteria as semisystematic reviews.

Four authors (MB, SRK, MOC, LKL) independently extracted predefined data of the included reviews using a data extraction form (supplementary material), which was specifically designed and piloted by the review team, and comparisons were made in pairs.

We extracted the following review characteristics:

1. Review identification: authors, year, title
2. From the systematic review(s), we extracted data on the number of trials included, the number of participants included, ICU population (eg, medical or surgical), diagnostic criteria of delirium, type of pharmacological agent(s) included, primary and secondary outcomes, results on primary and secondary outcomes, type of meta-analytic and sequential analysis used and the authors' conclusion

In addition, for all included reviews and meta-analyses, we extracted information on whether haloperidol was recommended for the management of delirium registered as either 'Yes/No/Not stated'. Disagreements concerning the extracted data were discussed and decision reached between the authors.

Assessment of methodological quality of included reviews

The methodological quality of the reviews failing on a maximum of two arbitrary PRISMA criteria were hereafter assessed with the ROBIS tool.[44]

Data synthesis

We a priori[36] planned to perform meta-analysis and trial sequential analysis[45] of the trials with overall low risk of bias. However, as we solely identified trials with overall high risk of bias, we did not perform the analyses.

We categorised reviews into:

1. Systematic reviews (a review positively fulfilling all 27 PRISMA criteria)[39]
2. Semisystematic reviews being in overall agreement with the PRISMA statement except failing on a maximum of two arbitrary PRISMA criteria
3. Narrative reviews (any review not fulfilling the criteria for a systematic review or the criteria for a semisystematic review)

For the systematic reviews assessed to be of low risk of bias, two authors (MB, MOC) independently assessed the methodological quality of each included trial with the Cochrane risk of bias tool.[37] Disagreements were discussed, and agreement was reached between the authors. Results are presented narratively by the indication for use (prevention or/and management), followed by the type of pharmacological agent and the type of outcome.

PATIENTS AND PUBLIC INVOLVEMENT

Patients and the public were not involved in this research.

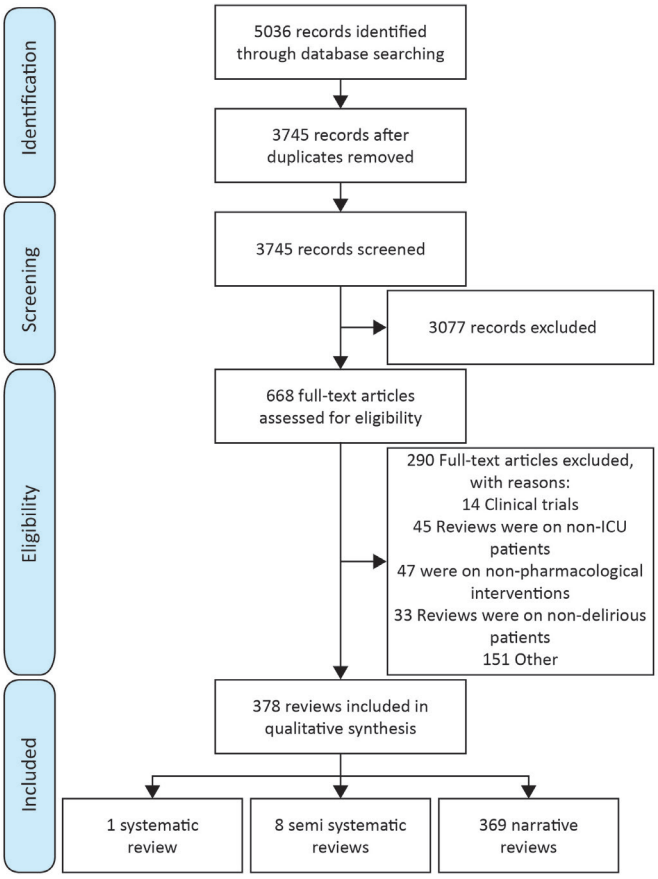


Figure 1 Preferred reporting items for systematic reviews and meta-analyses flowchart.

RESULTS

We identified 5036 potentially relevant references and finally included 378 reviews (figure 1).

Description of included reviews

We only identified one systematic review[46] fulfilling all 27 PRISMA criteria (ESM table 2), eight semisystematic reviews[47–54] failing on a maximum of two PRISMA criteria and 369 narrative reviews.

The systematic review

- Chen *et al*[46] assessed the safety and efficacy of alpha-2 agonists for sedation, compared with traditional sedatives, in mechanically ventilated critically ill patients. This review included seven trials randomising 1624 participants. All included trials investigated adults and compared dexmedetomidine with traditional sedatives (propofol, midazolam or lorazepam).

Semisystematic reviews

1. Tan *et al*[53] assessed the effects of using dexmedetomidine as a sedative and analgesic agent compared with placebo or alternative sedative agents, such as propofol and benzodiazepines, in critically ill patients; 24 randomised trials, involving 2419 patients, were included.
2. Lin *et al*[51] assessed the effects of using dexmedetomidine compared with alternative sedative agents following cardiac surgery; five randomised trials and six observational studies were included. We report on a subgroup analysis including five randomised trials and a prospective descriptive study.
3. Fraser *et al*[52] reviewed benzodiazepine compared with non-benzodiazepine (four randomised trials with dexmedetomidine and two with propofol) regimens in mechanically ventilated ICU patients. Six randomised trials, involving 1225 patients, were included.
4. Xia *et al*[47] assessed the influence of dexmedetomidine and propofol on adult ICU sedation. Ten randomised trials, involving 1202 participants, were included.
5. Zhang *et al*[48] included all postoperative trials reporting on delirium risk. We report on only one comparison, alpha-2-adrenoreceptor agonists compared with other sedatives for the risk of postoperative delirium, where only cardiac surgical trials have been included, as the other outcomes included patient groups we excluded (two randomised trials on dexmedetomidine and one on clonidine, involving 445 patients).
6. Pasin *et al*[50] compared dexmedetomidine with any comparator in the ICU setting (nine randomised trials in ICU, four in cardiac surgery and one in cervical spine surgery, including a total of 3029 patients).
7. Tran *et al*[54] assessed alpha-2 agonists (all trials reported on dexmedetomidine) for non-procedural sedation in critically ill brain-injured patients on mechanical ventilation. Both randomised trials and observational studies were included. Six randomised trials including a total of 318 patients were included. However, due to lack of clinical homogeneity of the randomised trials and studies, pooling was deemed inappropriate. We only report on outcomes which were defined a priori.
8. Liu *et al*[49] compared the effects of dexmedetomidine and propofol sedation in adult patients after cardiac surgery; eight randomised trials involving 969 patients were included.

Risk of bias in the systematic review and the eight semisystematic reviews

We assessed the systematic review by Chen *et al*[46] as overall low risk of bias (table 1).

However, the seven included trials[55–60] were all overall high risk of bias (figure 2). The eight semisystematic reviews failing on a maximum of two arbitrary PRISMA criteria, by Tan *et al*,[53] Lin *et al*,[51] Fraser *et al*,[52] Xia *et al*,[47] Zhang *et al*,[48] Pasin,[50] Tran[54] and Liu *et al*,[49] were all overall high risk of bias. All 46 trials included in these eight semisystematic reviews were overall high risk of bias.

Effects of pharmacological interventions for delirium in ICU patients

Prevention of delirium

Antipsychotics

We did not identify any systematic review or semisystematic review assessing the effects of antipsychotics (eg, haloperidol) for the prevention of delirium.

Table 2 Pooled effect estimates reported by the systematic review and semisystematic reviews by outcome and type of pharmacological agent

	Antipsychotics	Sedatives (dexmedetomidine)	Cholinesterase inhibitors	Opioids	Melatonine
Primary outcome					
All-cause mortality	-*		–*	–*	–*
Chen	–*	RR 0.99, 0.79 to 1.24; 6 randomised trials including 1584 patients	–*	–*	–*
Tan	–*	RR 0.85, 0.64 to 1.13; 16 randomised trials including 1839 patients	–*	–*	–*
Lin	–*	RR 1.00, 0.28 to 3.60, 3 randomised trials including 444 patients	–*	–*	–*
Xia	–*	RR 0.83, 0.32 to 2.12; 5 randomised trials including 267 patients	–*	–*	–*
Fraser	–*	RR 1.01, 0.78 to 1.30; 4 randomised trials including 1101 patients	–*	–*	–*
Serious adverse events	–*	–*	–*	–*	–*
Delirium prevention	–*		–*	–*	–*
Chen	–*	RR 0.85; 0.63 to 1.14; 7 randomised trials including 1624 patients	–*	–*	–*
Tan	–*	RR 0.79, 0.56 to 1.11; 8 randomised trials including 1754 patients	–*	–*	–*
Fraser	–*	RR 0.82, 0.61 to 1.11; 2 randomised trials including 469 patients	–*	–*	–*
Zhang	–*	RR 0.55, 0.23 to 1.28; 3 randomised trials including 445 patients†	–*	–*	–*
Lin	–*	RR 0.35, 0.19 to 0.63; 3 randomised trials including 478 patients	–*	–*	–*
Xia	–*	RR 0.40, 0.22 to 0.74; 3 randomised trials including 658 patients	–*	–*	–*
Liu	–*	RR 0.40, 0.24 to 0.64; 4 randomised trials including 393 patients	–*	–*	–*
Pasin	–*	RR 0.68, 0.49 to 0.96; 14 randomised trials including 3029 patients	–*	–*	–*
Tran	–*	Meta-analysis not performed, 0 trials included on this outcome	–*	–*	–*
Delirium management	-*	-*	-*	-*	-*
Secondary outcomes					
Quality of life	–*	–*	–*	–*	–*
Non-serious adverse events	–*		–*	–*	–*
Tran	–*	Meta-analysis not performed, 3 included trials was described narratively	–*	–*	–*
Cognitive function	–*		–*	–*	–*

*No systematic review or semisystematic review identified or assessed this outcome.
†Clonidine and dexmedetomidine.

Sedatives

All-cause mortality

When assessing mortality (table 2), Chen *et al*[46] did not find evidence for a difference when comparing dexmedetomidine with traditional sedatives (midazolam, lorazepam or propofol).

Neither did Tan *et al*[53] and Lin *et al*[51] when comparing dexmedetomidine with traditional sedatives. Additionally,

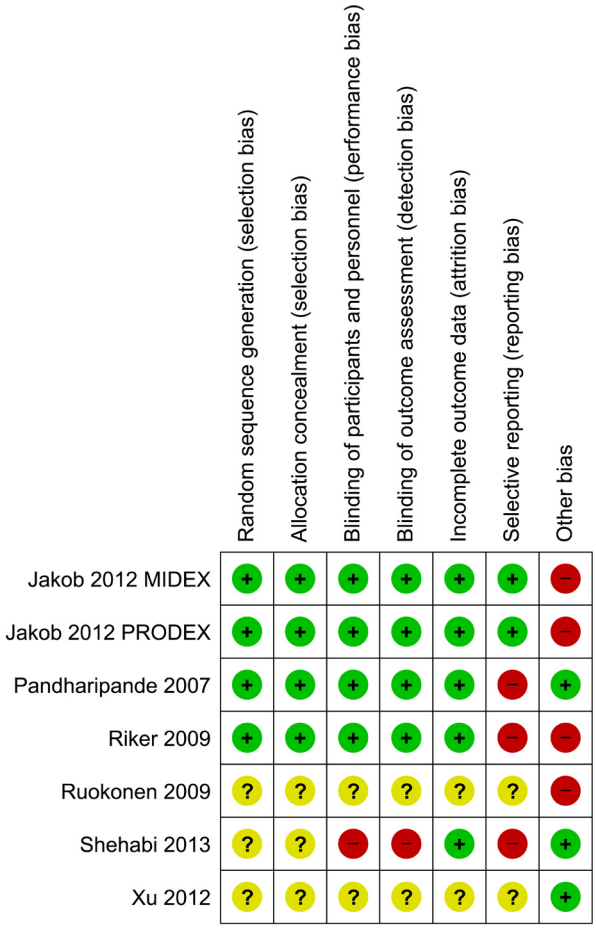


Figure 2 Risk of bias summary: review authors' judgements about each risk of bias item for each included trial in the only included systematic review (Chen 2015).

Xia *et al*[47] compared dexmedetomidine with propofol and also found no difference in mortality. Fraser *et al*[52] compared benzodiazepines with non-benzodiazepines (dexmedetomidine or propofol) and found no difference in mortality.

Serious adverse events

We did not identify any systematic review or semisystematic review assessing the effects of sedatives on risk of serious adverse events.

Risk of delirium

When assessing the effect of prophylactic use of alpha-2-agonists compared with alternative sedatives on the subsequent risk of delirium (table 2), the systematic review (on dexmedetomidine)[46] and three semisystematic reviews (two assessing dexmedetomidine[52 53] and one overall alpha-2-agonists[48]) did not find evidence of an effect.

Conversely, four semisystematic reviews[47–51] and a subgroup analysis (including two trials and a total of 415 patients) in a semisystematic review, which assessed alpha-2-agonists in the primary analysis,[48] found evidence of a beneficial effect of dexmedetomidine compared with different alternative sedatives.[47–51] In various subgroup analyses (on patients undergoing invasive ventilation, compared with midazolam only, restricted to general ICU), without any adjustment for statistical multiplicity, dexmedetomidine was found to have a beneficial effect for the prevention of delirium.[50]

Quality of life

We did not identify any systematic review or semisystematic review assessing quality of life.

Proportion of participants with non-serious adverse events

When assessing adverse events, Tran *et al*[54] narratively reported on three trials. Two trials found no evidence of a difference in adverse events comparing dexmedetomidine with propofol, or between dexmedetomidine and midazolam.[61 62] The third trial comparing dexmedetomidine with normal saline found that dexmedetomidine was associated with higher rates of bradycardia, but with lower rates of tachycardia.[63]

Cognitive function

We did not identify any systematic review or semisystematic review assessing cognitive function.

Additional outcomes reported by the systematic review and the semisystematic reviews

Twenty-three additional outcomes (mainly) on the effect of dexmedetomidine versus other sedatives were reported by the systematic review and semisystematic reviews (supplementary material table 3).

Cholinesterase inhibitors

We did not identify any systematic review or semisystematic review assessing the effects of cholinesterase inhibitors for the prevention of delirium.

Opioids

We did not identify any systematic review or semisystematic review assessing the effects of opioids for the prevention of delirium.

Melatonine and melatonine inhibitors

We did not identify any systematic review or semisystematic review assessing the effects of melatonine or melatonine inhibitors for the prevention of delirium.

Management of delirium

Antipsychotics

We did not identify any systematic review or semisystematic review assessing the effects of antipsychotics (eg, haloperidol) for the management of manifest delirium (table 2).

Of all 378 included reviews, 227 (60%) stated that haloperidol was indicated for the management of delirium, 43 (11%) stated that haloperidol was contraindicated and 108 (29%) did not state whether haloperidol was indicated or not.

Table 3 Summary of findings

Pharmacological intervention	No. of systematic reviews according to PRISMA with low risk of bias	No. of systematic reviews according to PRISMA with high risk of bias	No. of semisystematic reviews according to PRISMA*	Quality of the evidence	Comments
Delirium prevention	1	0	8	low	Seven trials with overall high risk of bias included in the systematic review with low risk of bias. The eight semisystematic reviews were all high risk of bias and included solely trials with overall high risk of bias.
Delirium management	0	0	0	No evidence	No systematic reviews according to PRISMA were identified. Neither was a semisystematic review identified.

Presence and quality of evidence by type of pharmacological intervention.
*In agreement with the PRISMA statement except two arbitrary PRISMA criteria.
PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Sedatives

We did not identify any systematic review or semisystematic review assessing the effects of sedatives for the management of manifest delirium.

Cholinesterase inhibitors

We did not identify any systematic review or semisystematic review assessing the effects of cholinesterase inhibitors for the management of manifest delirium.

Opioids

We did not identify any systematic review or semisystematic review assessing the effects of opioids for the management of manifest delirium.

Melatonine and melatonine inhibitors

We did not identify any systematic review or semisystematic review assessing the effects of melatonine or melatonine inhibitors for the management of manifest delirium.

DISCUSSION

Summary of main results

This overview addresses the evidence for the prevention of delirium and management of manifest delirium with pharmacological agents in ICU patients. We identified only one systematic review[46] out of a total of 378 reviews which addressed this topic. We classified eight as semisystematic reviews[47–53] and 369 as narrative reviews. We only found the systematic review to have overall low risk of bias; all eight semisystematic reviews had overall high risk of bias. The identified systematic review with low risk of bias included seven randomised clinical trials[55–60]; which all had overall high risk of bias. Our main results are summarised in the Summary of findings table (table 3).

Strengths and limitations of this study

This overview of reviews has several methodological strengths. We conducted a comprehensive literature search to identify reviews and meta-analyses in six major electronic databases, with specifically designed search strategies with no limits to publication year, type of publication or language. We used a transparent and systematic method, which was registered and published before the initiation of this project. Each phase of the screening, data extraction, data collection and methodological evaluations were performed by independent review authors working in pairs.

This overview of reviews also has methodological limitations. First, we chose PRISMA as the gold standard for defining a systematic review. One may argue that it is difficult for older reviews to adhere to the PRISMA statement, as this was published in 2009. One may also argue that there may be PRISMA criteria that might not be as important as others, for example, a structured abstract. In contrast, risk of bias evaluation in individual trials is of huge importance for the conclusion of the review.[64] Therefore, we chose pragmatically to classify all reviews, failing on a maximum of two PRISMA criteria, as semisystematic reviews. Second, we did not search for individual trials to perform a systematic review with meta-analyses and trial sequential analysis within each of the groups of pharmacological agents. Unfortunately, our results revealed that no systematic review on delirium management with any pharmacological agent has been published. Thus, we cannot discuss the evidence on pharmacological prevention or management strategies based on published trials, but merely according to the published reviews.

Current research within delirium is challenged by methodological and clinical limitations. The main limitations revealed by this overview of reviews is the overall high risk of bias found both in all the semisystematic reviews and all the included trials. It is therefore likely that we purport results that are also biased, that is, beneficial results may be overestimated, and harms may be underestimated.[64–66] In addition, we found a significant limitation to the research in the ICU delirium field, as systematic reviews adhering (or largely) to the PRISMA criteria all examined dexmedetomidine, which therefore dominates the current literature on pharmacological agents for delirium. Furthermore, the mechanisms of delirium are still not fully established and the underlying cause of delirium in medical ICU patients may be different from those in postoperative ICU patients, suggesting different optimal prevention and management strategies in the mixed ICU population. Certain subgroups of patients with delirium and risk factors at baseline (eg, age, severity of illness, exposure to a surgical procedure, cognitive dysfunction) may influence patient-centred outcomes differently. Current published trials have not stratified according to these factors but may in future research add new knowledge to the ICU field. Another important consideration is that many so-called placebo-controlled trials are not truly placebo-controlled, as some trials include rescue medications like haloperidol 'as needed'.

No study has previously attempted to systematically collect and evaluate all published reviews within pharmacological interventions for delirium. We found that narrative and non-systematic reviews dominate the literature on pharmacological interventions for delirium. Our findings confirm the observations by Siontis *et al*[67] that publications of erratic quality are produced in massive scales, in publications on the same topic, making it difficult to quickly get an evidence-based insight and overview. Our results reveal that many reviews cite trial results uncritically, leaving readers with the impression that, for example, haloperidol is a proven suitable pharmacological agent for the management of manifest delirium. Rapid access to current research to ensure evidence-based decision making and practice is increasingly demanded by the healthcare system, but guideline developers and decision makers are likely to be overwhelmed by the high numbers of published reviews of erratic quality.

Delirium prevention

Using a pharmacological delirium prevention protocol in adult ICU patients is not currently recommended.[17] The identified systematic review and eight semisystematic reviews considered prevention of delirium with dexmedetomidine, when used as a sedative, and found conflicting results, five in favour of dexmedetomidine[47–51] and three showing equipoise[46 52 53] results. However, trials with overall high risk of bias and small sample sizes not reaching the required information size in a meta-analysis,[68] as well as demonstrating huge heterogeneity of unexplained origin, prevent us from presenting any recommendations for the use of dexmedetomidine for the prevention of ICU delirium. We did not find a systematic review or semisystematic review addressing delirium prevention with haloperidol. To our knowledge, 10 randomised trials on haloperidol including a total of 3772 ICU patients or patients having major surgeries have been published.[32–34 69–75]

Sedation trials for the prevention of delirium overshadows research in preventive strategies. However, today, sedation is generally lessened, and light sedation and daily sedative interruption are recommended (low-quality evidence).[17] Sedation with dexmedetomidine and propofol are recommended over benzodiazepines in mechanically ventilated adults (low quality of evidence)[17]; however, no pharmacological agent is recommended for the prevention of delirium.[17] Patients may presumably benefit from being sedated with an agent which may lower the incidence of delirium, but using an agent to prevent delirium may then compete with the trend of minimising sedation.[76]

Delirium management

We did not find a systematic review according to the PRISMA criteria addressing pharmacological agents for the management of manifest delirium in ICU patients. To our knowledge, seven randomised trials investigating the effect of haloperidol for the management of manifest delirium in critically ill patients have been published[35 77–82] including only a total of 394 critically ill patients. Our overview of reviews demonstrates that the majority of reviews (60%), discussing the effect or use of haloperidol for delirium management, cite that haloperidol is indicated, and only 11% states that haloperidol is contraindicated. For whatever reason, the widespread use and endorsement of haloperidol contradicts the frequent serious adverse reactions shown in other settings,[28] and the fact that the Food and Drug Administration warns against the use of haloperidol in patients with dementia-related psychosis, because of a 1.6-times increased mortality.[83]

Unanswered questions and future research

In evidence-based medicine, systematic reviews of randomised trials rank highest. However, systematic reviews must be performed based on methods aiming to minimise systematic and random errors; otherwise, the results will be questionable. In addition to a thorough and systematic bias risk assessment, meta-analysis needs to reach a required information size (meta-analytic sample size) based on a minimal important clinical difference to conclude whether an intervention is better than another. Otherwise, a conclusion based on meta-analyses with high risk of random error[45 65 84] may be communicated. The lack of evidence and poor quality of the present evidence on the use of pharmacological agents for delirium leave clinicians to decide which pharmacological intervention to use. Research on how to deal with the management of manifest delirium, when all non-pharmacological

options have been used, is highly warranted. Although multicomponent, non-pharmacological intervention focusing on reducing modifiable risk factors for delirium, improving cognition and optimising sleep, mobility, hearing, and vision in critically ill adults, as well as early mobilisation, is recommended to reduce the incidence and duration in the ICU, this is only supported by low quality of evidence.[17] In settings outside the ICU, non-pharmacological multicomponent protocols have shown promising results (moderate level of quality).[85 86] However, such multifaceted interventions have not been adequately studied in the ICU setting. Based on the available evidence, one might get the idea that there is some evidence for the effect of dexmedetomidine to prevent delirium. However, as our overview underlines, there is really no valid evidence to support the use of dexmedetomidine and none at all that dexmedetomidine is better than haloperidol (or vice versa), which seems to be the preferred agent so far.[16 19]

CONCLUSION

Our overview of reviews demonstrated that systematic reviews and semisystematic reviews currently available in the delirium literature are heterogeneous in quality with high risk of bias. The results were conflicting regarding the effect of dexmedetomidine for the prevention of delirium based on the high-quality systematic review and the semisystemtic reviews. There is no evidence for the use of any pharmacological agent for the management of manifest delirium based on systematic or semisystematic reviews.

There is an urgent need for a systematic review with low risk of bias assessing the effects of pharmacological prevention of delirium and management of manifest delirium in ICU patients. Especially the effects of haloperidol need to be assessed, because haloperidol is the most recommended drug for the management of delirium. Future systematic reviews should aim to adhere to the PRISMA statement, so risk of systematic errors is minimised, and the best available evidence is presented. Furthermore, future trials on any antidelirious agent should report on patient-centred outcomes.

Identifying the most effective intervention for both the prevention of delirium and management of manifest delirium in ICU patients will benefit patients, relatives and healthcare systems around the world.

Difference between protocol and review

In our published protocol which was written a priori initiation of the overview, we stated that we would categorise reviews into the following groups: (1) systematic reviews according to PRISMA with low risk of bias assessed with ROBIS; (2) systematic reviews according to PRISMA with high risk of bias assessed with ROBIS; and (3) non-systematic reviews according to PRISMA.

Because we only found one systematic review fulfilling all the PRISMA criteria, we decided post protocol publication to acknowledge reviews almost fulfilling the PRISMA criteria by adding the category semisystematic reviews.

Author affiliations
[1]Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark
[2]Centre for Research in Intensive Care, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark
[3]Department of Intensive Care, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark
[4]Department of of Neuroanaesthesiology, Rigshospitalet, Copenhagen University hospital, Copenhagen, Denmark
[5]Department of Cardiology, Holbaek Hospital, Holbaek, Denmark

Acknowledgements We thank Dezhao Kong, Ning Lianhg, Di Wang and Sainan Fang for helping with the translation of the Chinese reviews.

Contributors Coordinating the review: MB. Conception and design of the review: MB, JCJ, AP, JW. Performing search strategies and literature searches: MB. Performing the literature screening: MB, SRK, LKL, MOC. Performing the data extraction: MB, SRK, LKL, MOC. Writing the review: MB, JW, JCJ, AP, MOC, LKL, SRK. All authors critically reviewed and revised the manuscript and approved the final version for publication. MB is the guarantor of the review.

Funding MB, MOC, JCJ, AP and JW were supported by the public Innovation Fund Denmark (4108-00011B), which did not have any role in study design, data collection, data analysis, data interpretation or writing of the report. No other sources of financial support were obtained for this review.

Competing interests MB, MOC, JCJ, AP and JW had financial support from the public Innovation Fund Denmark; AP is head of research at the Dept. of Intensive Care unit at Rigshospitalet. The department receives support for research from CSL Behring, Fresenius Kabi and Ferring Pharmaceuticals. AP and JW are members of the AID-ICU steering group (NCT03392376); no other relationships or activities that could appear to have influenced the submitted work.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

REFERENCES

1. American Psychiatric Association. *Practice Guidelines for the treatment of psychiatric disorders*, 2006.
2. Peterson JF, Pun BT, Dittus RS, *et al*. Delirium and its motoric subtypes: a study of 614 critically ill patients. *J Am Geriatr Soc* 2006;54:479–84.
3. Pandharipande P, Cotton BA, Shintani A, *et al*. Motoric subtypes of delirium in mechanically ventilated surgical and trauma intensive care unit patients. *Intensive Care Med* 2007;33:1726–31.
4. Meagher DJ, Trzepacz PT. Motoric subtypes of delirium. *Semin Clin Neuropsychiatry* 2000;5:75–85.
5. Ely EW, Inouye SK, Bernard GR, *et al*. Delirium in mechanically ventilated patients: validity and reliability of the confusion assessment method for the intensive care unit (CAM-ICU). *JAMA* 2001;286:2703–10.
6. Ely EW, Gautam S, Margolin R, *et al*. The impact of delirium in the intensive care unit on hospital length of stay. *Intensive Care Med* 2001;27:1892–900.
7. Ely EW, Girard TD, Shintani AK, *et al*. Apolipoprotein E4 polymorphism as a genetic predisposition to delirium in critically ill patients. *Crit Care Med* 2007;35:112–7.
8. Hipp DM, Ely EW. Pharmacological and nonpharmacological management of delirium in critically ill patients. *Neurotherapeutics* 2012;9:158–75.

9. Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. *Lancet* 2014;383:911–22.
10. Milbrandt EB, Deppen S, Harrison PL, *et al*. Costs associated with delirium in mechanically ventilated patients. *Crit Care Med* 2004;32:955–62.
11. Girard TD, Jackson JC, Pandharipande PP, *et al*. Delirium as a predictor of long-term cognitive impairment in survivors of critical illness. *Crit Care Med* 2010;38:1513–20.
12. Jackson JC, Gordon SM, Hart RP, *et al*. The association between delirium and cognitive decline: a review of the empirical literature. *Neuropsychol Rev* 2004;14:87–98.
13. Lin SM, Liu CY, Wang CH, *et al*. The impact of delirium on the survival of mechanically ventilated patients. *Crit Care Med* 2004;32:2254–9.
14. Pisani MA, Kong SY, Kasl SV, *et al*. Days of delirium are associated with 1-year mortality in an older intensive care unit population. *Am J Respir Crit Care Med* 2009;180:1092–7.
15. van den Boogaard M, Schoonhoven L, Evers AW, *et al*. Delirium in critically ill patients: impact on long-term health-related quality of life and cognitive functioning. *Crit Care Med* 2012;40:112–8.
16. Collet MO, Caballero J, Sonneville R, *et al*. Prevalence and risk factors related to haloperidol use for delirium in adult intensive care patients: the multinational AID-ICU inception cohort study. *Intensive Care Med* 2018;44:1081–9.
17. Devlin JW, Skrobik Y, Gélinas C, *et al*. Clinical practice guidelines for the prevention and management of pain, agitation/sedation, delirium, immobility, and sleep disruption in adult patients in the ICU. *Crit Care Med* 2018;46:e825–e873.
18. Bannon L, McGaughey J, Clarke M, *et al*. Impact of non-pharmacological interventions on prevention and treatment of delirium in critically ill patients: protocol for a systematic review of quantitative and qualitative research. *Syst Rev* 2016;5:75.
19. Borthwick MBR, Craig M, Egan A, *et al*. Detection, prevention and treatment of delirium in critically ill patients. *Intensive Care Society* 2006 http://members.ics.ac.uk/AsiCommon/Controls/BSA/Downloader.aspx?iDocumentStorageKey=d616cece-070a-48a4-bfaa-18da5a546634&iFileTypeCode=PDF&iFileName=Detection,%20Prevention%20and%20Treatment%20of%20Delirium (Accessed Mar 2018).
20. British Geriatrics Society. The prevention, diagnosis and management of delirium in older people - National guideline. 2006 https://www.rcplondon.ac.uk/guidelines-policy/prevention-diagnosis-referral-and-management-delirium-older-people (Accessed Mar 2018).
21. Mac Sweeney R, Barber V, Page V, *et al*. A national survey of the management of delirium in UK intensive care units. *QJM* 2010;103:243–51.
22. Devlin JW, Bhat S, Roberts RJ, *et al*. Current perceptions and practices surrounding the recognition and treatment of delirium in the intensive care unit: a survey of 250 critical care pharmacists from eight states. *Ann Pharmacother* 2011;45:1217–29.
23. George SPC, Patteril MV. Managing delirium in Intensive care. *Intensive Care Medicine* 2009;35(1 supplement):s258.
24. Meyer AMA, Dornblaser EK, Bellamy CJ, *et al*. A multi-center characterization of antipsychotic use for the treatment of delirium in medical ICU patients. *Pharmacotherapy* 2011;31:421e.
25. Trogrlic Z, Ista E, Slooter A, Bakker J, *et al*. Correction: Current practices in ICU delirium management: a prospective multicenter study in the Netherlands. *Critical Care* 2013;17:P548.
26. Martin J, Heymann A, Bäsell K, *et al*. Evidence and consensus-based German guidelines for the management of analgesia, sedation and delirium in intensive care-short version. *Ger Med Sci* 2010;8:Doc02.
27. Dansk Selskab for Anæstesi og Intensiv Medicin (DASAIM). Sedationsstrategi - Målrettet behandling af gener forbundet med kritisk sygdom. 2014 https://www.google.dk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwjkz_yS4cXOAhWF1iwKHQySCpMQFggcMAA&url=http%3A%2F%2Fwww.dasaim.dk%2Fwp-content%2Fuploads%2F2015%2F09%2FSedationsstrategi-sept15.pdf&usg=AFQjCNF9xGfr3aoeEBMTTPHrDpBm2BqMSw&cad=rja (Accessed Mar 2018).
28. Barr J, Fraser GL, Puntillo K, *et al*. Clinical practice guidelines for the management of pain, agitation, and delirium in adult patients in the intensive care unit. *Crit Care Med* 2013;41:263–306.
29. Grimshaw JM, Santesso N, Cumpston M, *et al*. Knowledge for knowledge translation: the role of the Cochrane Collaboration. *J Contin Educ Health Prof* 2006;26:55–62.
30. Laupacis A, Straus S. Systematic reviews: time to address clinical and policy relevance as well as methodological rigor. *Ann Intern Med* 2007;147:273–4.
31. Devlin JW, Roberts RJ, Fong JJ, *et al*. Efficacy and safety of quetiapine in critically ill patients with delirium: a prospective, multicenter, randomized, double-blind, placebo-controlled pilot study. *Crit Care Med* 2010;38:419–27.
32. Reade MC, O'Sullivan K, Bates S, *et al*. Dexmedetomidine vs. haloperidol in delirious, agitated, intubated patients: a randomised open-label trial. *Crit Care* 2009;13:R75.
33. Girard TD, Pandharipande PP, Carson SS, *et al*. Feasibility, efficacy, and safety of antipsychotics for intensive care unit delirium: the MIND randomized, placebo-controlled trial. *Crit Care Med* 2010;38:428–37.
34. Page VJ, Ely EW, Gates S, *et al*. Effect of intravenous haloperidol on the duration of delirium and coma in critically ill patients (Hope-ICU): a randomised, double-blind, placebo-controlled trial. *Lancet Respir Med* 2013;1:515–23.
35. van Eijk MM, Roes KC, Honing ML, *et al*. Effect of rivastigmine as an adjunct to usual care with haloperidol on duration of delirium and mortality in critically ill patients: a multicentre, double-blind, placebo-controlled randomised trial. *Lancet* 2010;376:1829–37.
36. Barbateskovic M, Larsen LK, Oxenbøll-Collet M, *et al*. Pharmacological interventions for delirium in intensive care patients: a protocol for an overview of reviews. *Syst Rev* 2016;5:211.
37. Higgins J, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*: The Cochrane Collaboration, 2011.
38. Robinson KA, Chou R, Berkman ND, *et al*. Twelve recommendations for integrating existing systematic reviews into new reviews: EPC guidance. *J Clin Epidemiol* 2016;70:38–44.
39. Liberati A, Altman DG, Tetzlaff J, *et al*. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009;339:b2700.
40. International Council for Harmonisation of Technical Requirements for pharmaceuticals for human use (ICH). *Guideline for Good Clinical Practice E6(R2)*, 2016.
41. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473–83.
42. Randolph C, Tierney MC, Mohr E, *et al*. The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): preliminary clinical validity. *J Clin Exp Neuropsychol* 1998;20:310–9.
43. Veritas Health Innovation. *Covidence systematic review software*. Melbourne, Australia, 2018. https://www.covidence.org/. (Accessed Mar 2018).
44. Whiting P, Savović J, Higgins JP, *et al*. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol* 2016;69:225–34.
45. Wetterslev J, Thorlund K, Brok J, *et al*. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. *J Clin Epidemiol* 2008;61:64–75.
46. Chen K, Lu Z, Xin YC, *et al*. Alpha-2 agonists for long-term sedation during mechanical ventilation in critically ill patients. *Cochrane Database Syst Rev* 2015;1:Cd010269.
47. Xia ZQ, Chen SQ, Yao X, *et al*. Clinical benefits of dexmedetomidine versus propofol in adult intensive care unit patients: a meta-analysis of randomized clinical trials. *J Surg Res* 2013;185:833–43.
48. Zhang H, Lu Y, Liu M, *et al*. Strategies for prevention of postoperative delirium: a systematic review and meta-analysis of randomized trials. *Crit Care* 2013;17:R47.
49. Liu X, Xie G, Zhang K, *et al*. Dexmedetomidine vs propofol sedation reduces delirium in patients after cardiac surgery: A meta-analysis with trial sequential analysis of randomized controlled trials. *J Crit Care* 2017;38:190–6.
50. Pasin L, Landoni G, Nardelli P, *et al*. Dexmedetomidine reduces the risk of delirium, agitation and confusion in critically Ill patients: a meta-analysis of randomized controlled trials. *J Cardiothorac Vasc Anesth* 2014;28:1459–66.
51. Lin YY, He B, Chen J, *et al*. Can dexmedetomidine be a safe and efficacious sedative agent in post-cardiac surgery patients? a meta-analysis. *Crit Care* 2012;16:R169.
52. Fraser GL, Devlin JW, Worby CP, *et al*. Benzodiazepine versus nonbenzodiazepine-based sedation for mechanically ventilated, critically ill adults: a systematic review and meta-analysis of randomized trials. *Crit Care Med* 2013;41(9 Suppl 1):S30–8.
53. Tan JA, Ho KM, Km H. Use of dexmedetomidine as a sedative and analgesic agent in critically ill adult patients: a meta-analysis. *Intensive Care Med* 2010;36:926–39.
54. Tran A, Blinder H, Hutton B, *et al*. A systematic review of alpha-2 agonists for sedation in mechanically ventilated neurocritical care patients. *Neurocrit Care* 2018;28:12–25.
55. Jakob SM, Ruokonen E, Grounds RM, *et al*. Dexmedetomidine vs midazolam or propofol for sedation during prolonged

mechanical ventilation: two randomized controlled trials. *JAMA* 2012;307:1151–60.
56. Pandharipande PP, Pun BT, Herr DL, *et al*. Effect of sedation with dexmedetomidine vs lorazepam on acute brain dysfunction in mechanically ventilated patients: the MENDS randomized controlled trial. *JAMA* 2007;298:2644–53.
57. Riker RR, Shehabi Y, Bokesch PM, *et al*. Dexmedetomidine vs midazolam for sedation of critically ill patients: a randomized trial. *JAMA* 2009;301:489–99.
58. Ruokonen E, Parviainen I, Jakob SM, *et al*. Dexmedetomidine versus propofol/midazolam for long-term sedation during mechanical ventilation. *Intensive Care Med* 2009;35:282–90.
59. Shehabi Y, Bellomo R, Reade MC, *et al*. Early goal-directed sedation versus standard sedation in mechanically ventilated critically ill patients: a pilot study*. *Crit Care Med* 2013;41:1983–91.
60. Xu JB WY, Shi QS. A combined protocol for dexmedetomidine used in prolonged sedation in intensive care unit. *Modern Medicine Journal of China* 2012;14:20–2.
61. Srivastava VK, Agrawal S, Kumar S, *et al*. Comparison of dexmedetomidine, propofol and midazolam for short-term sedation in postoperatively mechanically ventilated neurosurgical patients. *J Clin Diagn Res* 2014;8:Gc04–7.
62. Mirski MA, Lewin JJ, Ledroux S, *et al*. Cognitive improvement during continuous sedation in critically ill, awake and responsive patients: the Acute Neurological ICU Sedation Trial (ANIST). *Intensive Care Med* 2010;36:1505–13.
63. Zhao LH, Shi ZH, Chen GQ, *et al*. USe of dexmedetomidine for prophylactic analgesia and sedation in patients with delayed extubation after craniotomy: A randomized controlled trial. *J Neurosurg Anesthesiol* 2017;29:132–9.
64. Savović J, Jones HE, Altman DG, *et al*. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. *Ann Intern Med* 2012;157:429–38.
65. Jakobsen JC, Wetterslev J, Winkel P, *et al*. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC Med Res Methodol* 2014;14:120.
66. Lundh A, Lexchin J, Mintzes B, *et al*. Industry sponsorship and research outcome. *Cochrane Database Syst Rev* 2017;2:Mr000033.
67. Siontis KC, Hernandez-Boussard T, Ioannidis JP. Overlapping meta-analyses on the same topic: survey of published studies. *BMJ* 2013;347:f4501.
68. Wetterslev J, Thorlund K, Brok J, *et al*. Estimating required information size by quantifying diversity in random-effects model meta-analyses. *BMC Med Res Methodol* 2009;9:86.
69. Wang W, Li HL, Wang DX, *et al*. Haloperidol prophylaxis decreases delirium incidence in elderly patients after noncardiac surgery: a randomized controlled trial*. *Crit Care Med* 2012;40:731–9.
70. Kalisvaart KJ, de Jonghe JF, Bogaards MJ, *et al*. Haloperidol prophylaxis for elderly hip-surgery patients at risk for delirium: a randomized placebo-controlled study. *J Am Geriatr Soc* 2005;53:1658–66.
71. van den Boogaard M, Schoonhoven L, van Achterberg T, *et al*. Haloperidol prophylaxis in critically ill patients with a high risk for delirium. *Crit Care* 2013;17:R9.
72. Kaneko TCJ, Ishikura T, Kobayashi M, *et al*. Prophylactic consecutive administration of haloperidol can reduce the occurrence of postoperative delirium in gastrointestinal surgery. *Yonago Acta medica* 1999;42:179–84.
73. Fukata S, Kawabata Y, Fujisiro K, *et al*. Haloperidol prophylaxis does not prevent postoperative delirium in elderly patients: a randomized, open-label prospective trial. *Surg Today* 2014;44:2305–13.
74. Al-Qadheeb NS, Skrobik Y, Schumaker G, *et al*. Preventing ICU subsyndromal delirium conversion to delirium with low-dose iv haloperidol: A double-blind, placebo-controlled pilot study. *Crit Care Med* 2016;44:583–91.
75. van den Boogaard M, Slooter AJC, Brüggemann RJM, *et al*. Effect of Haloperidol on Survival Among Critically Ill Adults With a High Risk of Delirium. *JAMA* 2018;319:680–90.
76. Strøm T, Martinussen T, Toft P. A protocol of no sedation for critically ill patients receiving mechanical ventilation: a randomised trial. *Lancet* 2010;375:475–80.
77. Skrobik YK, Bergeron N, Dumont M, *et al*. Olanzapine vs haloperidol: treating delirium in a critical care setting. *Intensive Care Med* 2004;30:444–9.
78. Atalan N, Efe Sevim M, Akgün S, *et al*. Morphine is a reasonable alternative to haloperidol in the treatment of postoperative hyperactive-type delirium after cardiac surgery. *J Cardiothorac Vasc Anesth* 2013;27:933–8.
79. Tagarakis GI, Voucharas C, Tsolaki F, *et al*. Ondasetron versus haloperidol for the treatment of postcardiotomy delirium: a prospective, randomized, double-blinded study. *J Cardiothorac Surg* 2012;7:25.
80. Breitbart W, Marotta R, Platt MM, *et al*. A double-blind trial of haloperidol, chlorpromazine, and lorazepam in the treatment of delirium in hospitalized AIDS patients. *Am J Psychiatry* 1996;153:231–7.
81. Lin CJ SF, Fang CK, Chen HW, *et al*. An open trial comparing haloperidol with olanzapine for the treatment of delirium in palliative and hospice center cancer patients. *Journal of Internal Medicine of Taiwan* 2008;19:346–54.
82. Han CS, Kim YK. A double-blind trial of risperidone and haloperidol for the treatment of delirium. *Psychosomatics* 2004;45:297–301.
83. Food and Drug Administration (FDA), 2005. https://www.google.dk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwjJ-6eB07nQAhVF1xoKHXJyCP8QFggcMAA&url=http%3A%2F%2Fwww.accessdata.fda.gov%2Fdrugsatfda_docs%2Flabel%2F2008%2F015923s082%2C018701s057lbl.pdf&usg=AFQjCNGbdzQFtVi2X5DcDmVDGRFmnujVYQ&cad=rja (Accessed Mar 2018).
84. Turner RM, Bird SM, Higgins JP. The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews. *PLoS One* 2013;8:e59202.
85. Abraha I, Trotta F, Rimland JM, *et al*. Efficacy of Non-Pharmacological Interventions to Prevent and Treat Delirium in Older Patients: A Systematic Overview. The SENATOR project ONTOP Series. *PLoS One* 2015;10:e0123090.
86. Abraha I, Rimland JM, Trotta F, *et al*. Non-pharmacological interventions to prevent or treat delirium in older patients: Clinical practice recommendations the SENATOR-ONTOP series. *J Nutr Health Aging* 2016;20:927–36.

PAPER II

Received: 13 August 2019 | Revised: 2 October 2019 | Accepted: 4 October 2019

DOI: 10.1111/aas.13501

REVIEW

acta Anaesthesiologica Scandinavica

Haloperidol for the treatment of delirium in critically ill patients: A systematic review with meta-analysis and Trial Sequential Analysis

Marija Barbateskovic[1,2] | Sara R. Krauss[1] | Marie O. Collet[2,3] | Nina C. Andersen-Ranberg[2,4] | Ole Mathiesen[2,4,5] | Janus C. Jakobsen[1,2,6,7] | Anders Perner[2,3] | Jørn Wetterslev[1,2]

[1]Copenhagen Trial Unit, Centre for Clinical Intervention Research, Copenhagen, Denmark

[2]Centre for Research in Intensive Care, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

[3]Department of Intensive Care, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

[4]Centre for Anaesthesiological Research, Department of Anaesthesiology and Intensive Care, Zealand University Hospital, Koege, Denmark

[5]Department of Clinical Medicine, Faculty of Health and Medical Sciences, Copenhagen University, Copenhagen, Denmark

[6]Department of Cardiology, Holbaek Hospital, Holbaek, Denmark

[7]Department of Regional Health Research, The Faculty of Heath Sciences, University of Southern Denmark, Odense, Denmark

Correspondence
Marija Barbateskovic, PhD candidate, Copenhagen Trial Unit, Centre for Clinical Intervention Research, DK2100 Copenhagen, Denmark.
Email: marija.barbateskovic@ctu.dk

Funding information
This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors. M Barbateskovic' contribution to this work, was supported by Innovation Fund Denmark [grant number 4108-00011B].

Background: Haloperidol is the most frequently used drug to treat delirium in the critically ill patients. Yet, no systematic review has focussed on the effects of haloperidol in critically ill patients with delirium.

Methods: We conducted a systematic review with meta-analysis and Trial Sequential Analysis of randomized clinical trials (RCTs) assessing the effects of haloperidol vs any intervention on all-cause mortality, serious adverse reactions/events, days alive without delirium, health-related quality of life (HRQoL), cognitive function and delirium severity in critically ill patients with delirium. We also report on QTc prolongation, delirium resolution and extrapyramidal symptoms.

Results: We included 8 RCTs with 11 comparisons (n = 951). We adjudicated one trial as having overall low risk of bias. Three trials used rescue haloperidol; excluding these, we did not find an effect of haloperidol vs control on all-cause mortality (RR 1.01; 95% CI 0.33-3.06; I^2 = 0%; 112 participants; 3 trials; 4 comparisons; very low certainty) or delirium severity (SMD −0.15; 95% CI −0.61-0.30; I^2 = 27%; 134 participants; 3 trials; 4 comparisons; very low certainty). No trials reported adequately on serious adverse reactions/events. Only one trial reported on days alive without delirium, cognitive function and QTc prolongation, and no trials reported on HRQoL. Sensitivity analyses, including trials using rescue haloperidol, did not change the results.

Conclusions: The evidence for the use of haloperidol to treat critically ill patients with delirium is sparse, of low quality and inconclusive. We therefore have no certainty regarding any beneficial, harmful or neutral effects of haloperidol in these patients.

1 | INTRODUCTION

Delirium has been reported to affect up to 89% of the critically ill patients and has been associated with poor clinical outcomes including lengthened mechanical ventilation and hospital stay and increased mortality.[1-6] Furthermore, surviving patients may experience functional decline and long-term cognitive impairment as a consequence of delirium.[6,7]

Haloperidol is the most frequently used pharmacological intervention for delirium treatment in Intensive Care Unit (ICU) settings.[8-11] The 2002 recommendations of the Society of Critical Care Medicine for clinical practice guidelines for the sustained use of sedatives and analgesics in the critically ill adults recommended haloperidol as the pharmacological agent for the treatment of delirium (Grade C recommendation, based on case series).[12] However, in the 2013 update of the guideline, this recommendation was changed; haloperidol was no longer recommended due to lack of evidence on the duration of delirium.[13] The latest 2018 update of the same guideline suggests that haloperidol may be used in some delirious cases but not systematically and again the recommendation was graded with low evidence.[14]

We have recently demonstrated that current available reviews on delirium management in ICU are of heterogeneous quality with high risk of bias; and we found no systematic reviews as per the PRISMA definitions assessing the effects of haloperidol for the treatment of delirium in ICU.[15] A newly published Cochrane review investigating the effect of pharmacological interventions in critically ill patients with delirium allowed the inclusion of trials with non-delirious patients, however, the trials included patients at risk of developing delirium.[16]

As no former systematic review has been conducted on haloperidol for delirium in critically ill patients, fulfilling the PRISMA criteria,[15,17] with meta-analysis and Trial Sequential Analysis (TSA)[18] our objective was to assess the benefits and harms of haloperidol vs placebo or any intervention for the treatment of delirium in critically ill patients. Our primary comparison was that of haloperidol with placebo. We hypothesized an increase in mortality, serious adverse reactions/events and QTc prolongation; a reduction in delirium duration and severity; and a beneficial effect on health-related quality of life (HRQoL) and cognitive status of haloperidol.

Editorial Comment

There is a need for effective treatments for delirium among critically ill patients. Haloperidol may be one of the more commonly used drugs for this purpose in clinical practice. This trustworthy systematic review presents an analysis of the pooled evidence for the use of haloperidol to treat delirium in patient in the intensive care unit.

2 | METHODS

This systematic review was conducted according to the pre-planned statistical analysis plan of the published protocol.[19] We registered the protocol in the international prospective register of systematic reviews database (PROSPERO) (CRD42017081133), used the methodology of the Cochrane Collaboration [20] and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [17] (Electronic Supplementary Material (ESM)).

2.1 | Eligibility criteria

We included randomized clinical trials (RCTs), irrespective of publication status, reported outcomes, publication date and language. Only RCTs with critically ill patients with delirium at trial enrolment were included. Critical illness included any clinical setting where patients are at high risk of dying or who have actual or potential life-threatening health problems and who are admitted to a high-dependency facility in the hospital, ie an ICU, a coronary care unit or similar facility. We did also include trials on acutely operated patients and elective cardiac surgical patients.

We included any trial comparing haloperidol with placebo, any other pharmacological agent, or combinations of pharmacological and non-pharmacological interventions (single or bundle).

RCTs were excluded if haloperidol was administered in both groups per protocol or if it was administered as a combination therapy with another pharmacological agent.

Our focus was to assess the association between haloperidol and the treatment of delirium (rather than prevention), thus, patients were required to be delirious prior to being randomized to trial drug. We did not accept agitation alone as an inclusion criterion.

2.2 | Outcomes

Our predefined co-primary outcomes were all-cause mortality and proportion of participants with one or more serious adverse reaction (SAR). We used serious as defined by ICH-GCP[21] either as reported by triallists or according to the SAR in the Summary Product Characteristics of haloperidol. Co-secondary outcomes were days alive without delirium within 28 days; HRQoL; cognitive function and delirium severity. We report on QTc prolongation as an exploratory outcome and post hoc analyses on delirium resolution and extrapyramidal symptoms. For all outcomes, we used the trial results reported at the time point closest to 3 months.

2.3 | Search methods

We searched the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, Embase, Science Citation Index, Biosis Previews, Cumulative Index to Nursing & Allied Health Literature (CINAHL) and Latin American Caribbean Health Sciences Literature (LILACS) from inception to 5 March 2019 (ESM).

In addition, we searched for ongoing and unpublished trials in the following registers: ClinicalTrials.gov; WHO International Clinical Trials Registry Platform (ICTRP); EU clinical trial register and Australian New Zealand Clinical Trials Registry (ANZCTR). US Food and Drug Administration (FDA); European Medicines Agency (EMA) and websites of medical companies were searched for unpublished trials. Ultimately, we searched the reference lists of the included trials and previous meta-analyses to identify further relevant trials.

2.4 | Trial selection and data extraction

Two review authors (MB, SRK) independently screened titles and abstracts. Reports deemed potentially relevant were obtained in full-text and assessed for inclusion. Disagreements were resolved by consensus and JW were consulted when agreement could not be met.

Two review authors (MB and SRK) independently extracted predefined data of the included trials using a predefined data collection form (ESM). The following data were collected: (a) Trial: country, duration of the trial, date of publication; (b) Participants: numbers randomized, numbers analysed, numbers lost to follow-up/withdrawn, type of population, age, gender, disease severity, setting, delirium assessment, inclusion criteria and exclusion criteria; (c) Interventions: intervention, comparator, duration and co-interventions; (d) Outcomes: predefined primary, secondary outcomes and timing of outcome measurement.[19]

2.5 | Risk of bias assessment

MB and SRK independently assessed the risk of systematic errors (bias) of the included trials using the Cochrane Collaboration's risk of bias tool.[20] We specifically assessed the following domains: (a) Random sequence generation; (b) Allocation concealment; (c) Blinding of participants and personnel; (d) Blinding of outcome assessment; (e) Incomplete outcome data; (f) Selective outcome reporting; and (g) Other bias, including early stopping and bias due to vested financial interest or academic bias. The included trials were adjudicated as 'overall low risk of bias' when all bias domains were adjudicated as low risk of bias. Conversely, trials were adjudicated as 'overall high risk of bias' when unclear or high risk of bias was adjudicated in one or more domains.

We planned to assess publication bias, by inspecting funnel plots for signs of asymmetry when ten or more trials were included in an analysis [20,22] and planned to test for asymmetry with the Harbord test.[23]

2.6 | Data synthesis

2.6.1 | 2.6.1 Summary measures

Risk ratios (RRs) with 95% confidence intervals (CIs) and CIs adjusted for sparse data, multiple outcomes and testing (TSA adjusted CIs) were calculated for dichotomous outcomes. For continuous outcomes, end-scores were used, and mean difference (MD) and standardized mean difference (SMD) with CIs and TSA adjusted CIs were planned to be calculated.

2.6.2 | Meta-analysis

We considered the comparison of haloperidol with placebo or with other pharmacological agents in trials not using rescue haloperidol (escape medication) as our primary comparison. We calculated pooled effect estimates using Review Manager.[24] We used a family-wise error rate of 5% [22] and considered a p-value of 0.05/[(2 + 1)/2] = 0.033 or less as statistical significant in the analyses of each co-primary outcome, and we considered a p-value of 0.05/[(3 + 1)/2.5] = 0.025 or less as statistical significant in the analyses of each co-secondary outcome to account for statistical multiplicity due to multiple outcomes. We calculated Bayes factor to assess if the summary effect estimates fitted better with the null hypothesis than alternative hypotheses of the anticipated intervention effects.[22]

2.6.3 | Dealing with missing data

Corresponding authors were contacted to clarify important missing data related to the methods, data reporting or if further trial details were needed (ESM).

We conducted a predefined sensitivity analysis by imputing missing outcome data in a best-worst case scenario and a worst-best case scenario to assess the potential impact of loss to follow-up.[19,22]

2.6.4 | Assessment of heterogeneity

We assessed heterogeneity by visual inspection of the forest plots, and calculated the inconsistency statistics (I^2) and the diversity statistics (D^2).[25] We assessed intervention effects with both random-effects model meta-analyses and fixed-effect model meta-analyses. We used the more conservative point estimate of the 2, which is the point estimate closest to no effect. If the estimates from the 2 models were approximately equal, we used the estimate with the widest CI.[19,22]

2.6.5 | Sensitivity analyses and subgroup analyses

We planned to conduct the following predefined subgroup analyses: trials with overall high risk of bias compared to trials with overall low risk of bias and grouping according to patient population, used control intervention in the trials and delirium diagnosis. We conducted a post hoc sensitivity analysis where we included trials using haloperidol as rescue medication.

2.6.6 | Trial Sequential Analysis

We used TSA to assess the risk of random errors due to sparse data, multiple outcomes and multiple testing of accumulating data,[18,26-34] and we calculated the required information size.[25]

We used a power of 90% (beta 10%) and a diversity [25] as suggested by the trials in the meta-analysis [22] or a diversity of

20% if the measured heterogeneity was zero.[34] As anticipated intervention effects for the primary and secondary outcomes in the TSA, we used a realistic a priori RRR or RRI of 20%. Furthermore, in a secondary TSA we used a RRR or RRI based on the 95% confidence limit closest to null effect in the traditional meta-analysis.[19]

We planned to present 95% CI and TSA adjusted CI. For a more detailed description of the statistical analysis plan and TSA, we refer to the published review protocol.[19]

2.6.7 | Grading certainty of evidence

We used The Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach[35] to assess the overall certainty of evidence for all pre-defined outcomes. We appraised the certainty of evidence and our confidence in the effect estimates based on risk of bias, inconsistency, indirectness, imprecision and publication bias. Thus, we rated the overall certainty of evidence as high, moderate, low or very low.

3 | RESULTS

3.1 | Study selection

We identified 5392 references and included 8 RCTs[36-43] with 11 comparisons (Figure 1) and a total of 951 participants. We listed reasons for exclusion of key excluded trials, which included 33 RCTs of haloperidol for the treatment of delirium in patients not being critically ill and 5 RCTs due to wrong indications (ESM). In addition, we identified 4 ongoing trials[44-47] and 8 terminated trials[48-55] with no results (ESM).

3.2 | Characteristics of included trials

The included trials were published between 1996 and 2018 (Table 1). Seven trials were published as full trial reports and one trial published its results on clinicaltrials.gov. The 8 included trials covered 11 comparisons, of which the control group was placebo in 2,[39,41] dexmedetomidine in 1,[37] morphine in 1,[36] benzodiazepine (lorazepam) in 1,[38] ondansetron in 2[37,43] and antipsychotics (chlorpromazine, ziprasidone, risperidone, olanzapine) in 4.[38-40,42] Three trials used haloperidol as rescue medication.[37,39,42] All trials included adult critically ill patients. Five trials included adults admitted to an ICU,[37,39-42] 2 trials included cardiac surgical patients[36,43] and 1 trial included medical patients.[38] Details and additional information of the included trials are presented in the ESM.

The number of participants in the trials ranged from 24 to 566. Mean age of participants ranged from 31 years to 71 years and proportion of men ranged between 54% and 91% in the included trials.

3.3 | Risk of bias

We adjudicated 1 trial as having overall low risk of bias; the remaining 7 had overall high risk of bias (Figure 2).

3.4 | Effect of interventions

3.4.1 | All-cause mortality

Four of 8 trials (6 comparisons)[36,38,39,41] with a total of 678 participants and a mean follow-up of 34 days (range 8 to 90 days) reported on all-cause mortality. One trial was overall low risk of bias and included 566 participants. Two trials were placebo-controlled trials. One trial used haloperidol as rescue drug.

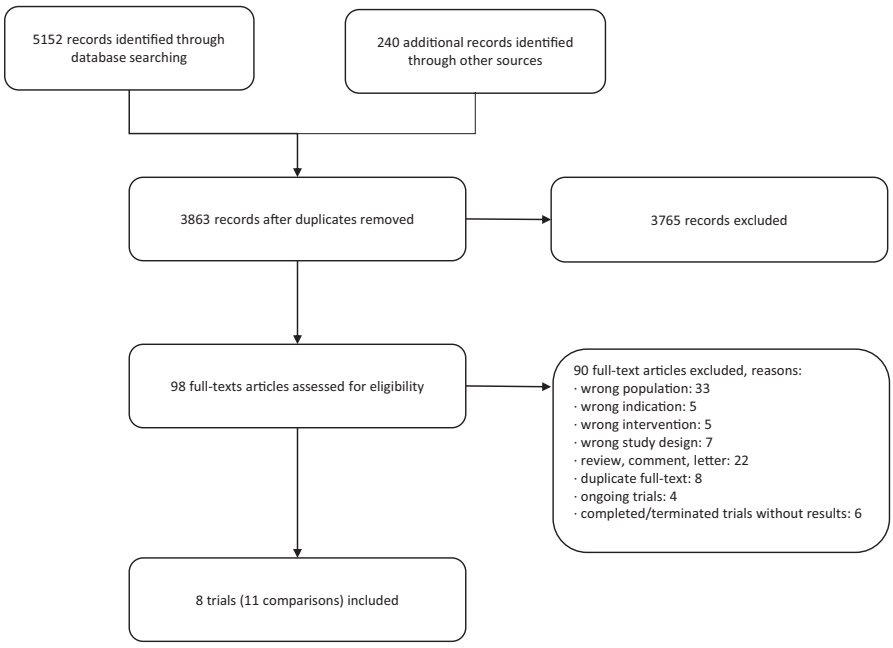


FIGURE 1 PRISMA flow diagram

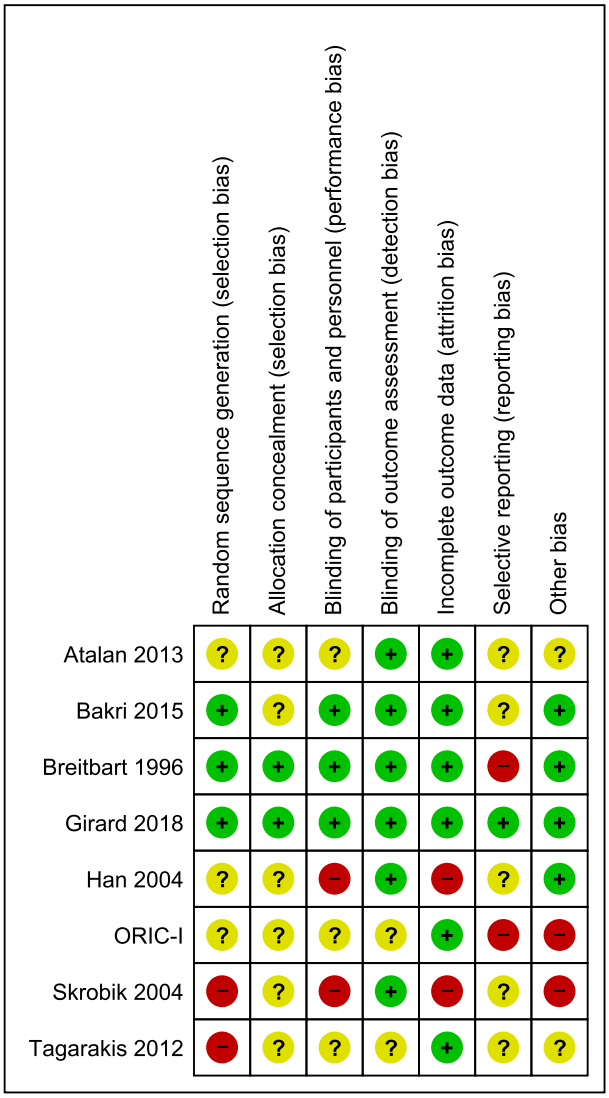


FIGURE 2 Risk of bias summary. Risk of bias summary: review authors' judgements about each risk of bias item for each included study. Green represents a low risk of bias, yellow an unclear risk of bias and red a high risk of bias

Meta-analysis, regardless of risk of bias, showed no evidence of a difference in haloperidol vs control for the treatment of delirium when assessing mortality (fixed effect model RR 1.01; 95% CI 0.33-3.06; I^2 = 0%; 112 participants; 3 trials; 4 comparisons; Figure 3). The certainty of evidence, using the GRADE approach, was very low due to serious risk of bias, indirectness and imprecision (Table 2).

As only 1% of the required information size had been reached, TSA adjusted CI could not be calculated. Bayes Factors are presented in the ESM.

The sensitivity analyses on missing data indicated that incomplete outcome data alone had the potential to influence the results: best-worst case scenario RR 0.85, 95% CI 0.29-2.48 and worst-best case scenario RR 1.03, 95% CI 0.34-3.15 (ESM).

The subgroup analysis excluding trials at overall high risk of bias could not be performed as no trials were overall low risk of bias. We found no interaction between intervention effect and use of control intervention, including patient population, and type of delirium in subgroup analyses (ESM).

Sensitivity analysis including the trial using rescue haloperidol showed similar results (fixed effect model RR 1.10; 95% CI 0.88-1.37; I^2 = 0%; TSA adjusted CI 0.65-1.89; 678 participants; 4 trials; 6 comparisons; ESM).

3.4.2 | Serious adverse reactions

Four trials (5 comparisons) reported on the proportion of patients with serious adverse reactions/events,[36,37,40,41] although none defined the adverse reactions/events according to ICH-GCP. All 4 trials reported zero events in each group despite reporting on mortality. Only one trial reported on individual SAEs.[41] The certainty of evidence was judged to be very low due to serious risk of bias, inconsistency, indirectness and imprecision (Table 2).

3.4.3 | Days alive without delirium within 28 days

One trial with overall low risk of bias[39] with 2 comparisons and 566 participants reported on days alive without delirium or coma during the 14-day intervention period. The trial used rescue haloperidol. A total of 8 days (0-11) in the haloperidol group, 8 days (0-11) in the placebo group and 8 days (2-11) in the ziprasidone group were reported.

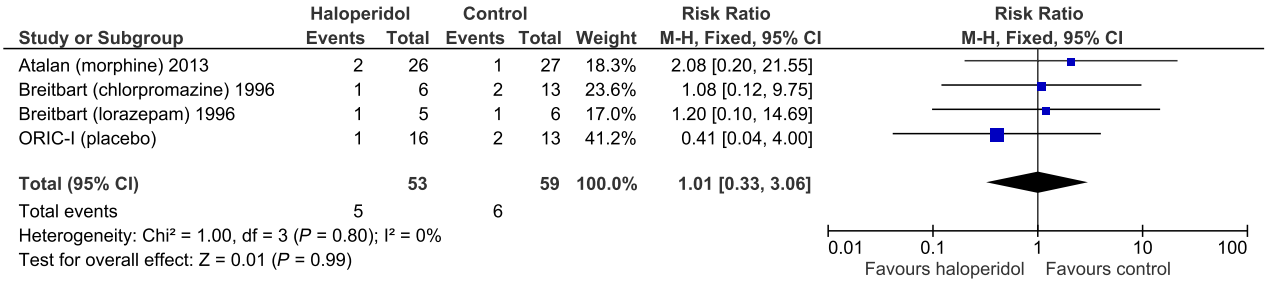


FIGURE 3 Forest plot of all-cause mortality, excluding trials using rescue haloperidol. No trials were overall low risk of bias. Parenthesis following author name show used control intervention. Size of squares for risk ratio reflects weight of trial in pooled analysis. Horizontal bars represent 95% confidence intervals

TABLE 1 Characteristics of included trials

Trial	N*	Setting	Intervention	Comparator	Duration of intervention	Outcomes*
Atalan 2013 [36]	53	Patients with hyperactive delirium after cardiac surgery admitted to ICU	5 mg haloperidol IM every hour until the adequate sedation and target RASS scores (between −1 and + 1) were achieved	5 mg morphine IM every hour until the adequate sedation and target RASS scores (between −1 and + 1) were achieved	Maximum 10 days	All-cause mortality Serious adverse reactions
Bakri (dexmedetomidine) 2015[37]	48	Post-operative trauma patients with delirium admitted to ICU	5 mg haloperidol twice daily (infusion) Rescue haloperidol was used	1 µg/kg dexmedetomidine or (infusion). Rescue haloperidol was used	3 days	Serious adverse reactions Delirium severity QTc prolongation Delirium resolution
Bakri (ondansetron) 2015[37]	48	Post-operative trauma patients with delirium admitted to ICU	5 mg haloperidol twice daily (infusion). Rescue haloperidol was used	4 mg ondansetron twice daily (infusion). Rescue haloperidol was used	3 days	Serious adverse reactions Delirium severity QTc prolongation Delirium resolution
Breitbart (chlorpromazine) 1996[38]	19	AIDS patients with delirium admitted to a high dependency AIDS unit	Haloperidol (oral or IM) dose according to delirium symptoms. Mean haloperidol dose the first 24 hours was 2.8 mg. Average maintenance dose was 1.4 mg.	Mean chlorpromazine dose the first 24 hours was 50 mg. Average maintenance dose was 36 mg	Maximum 6 days	All-cause mortality Cognitive function Delirium severity Extrapyramidal symptoms
Breitbart (lorazepam) 1996[38]	11	AIDS patients with delirium admitted to a high dependency AIDS unit	Haloperidol (oral or IM) dose according to delirium symptoms. Mean haloperidol dose the first 24 hours was 2.8 mg. Average maintenance dose was 1.4 mg.	Mean lorazepam dose the first 24 hours was 3 mg. Average maintenance dose was 4.6 mg.	Maximum 6 days	All-cause mortality Cognitive function Delirium severity Extrapyramidal symptoms
Girard (placebo) 2018[39]	280	Patients with delirium admitted to ICU	IV haloperidol. Mean daily doses of haloperidol administered were 11.0 mg Rescue haloperidol was used	Placebo Rescue haloperidol was used	Maximum 14 days	All-cause mortality Days alive without delirium QTc prolongation Extrapyramidal symptoms
Girard (ziprazidone) 2018[39]	286	Patients with delirium admitted to ICU	IV haloperidol. Mean daily doses of haloperidol administered were 11.0 mg Rescue haloperidol was used	IV ziprasidone. Mean daily doses of ziprasidone administered were 20.0 mg Rescue haloperidol was used	Maximum 14 days	All-cause mortality Days alive without delirium QTc prolongation Extrapyramidal symptoms
Han 2004[40]	24	Patients with delirium admitted to ICU**	Oral flexible dose haloperidol. Mean dose of haloperidol was 1.71 mg	Oral flexible dose risperidone. Mean dose of risperidone 1.02	7 days	Serious adverse reactions Delirium severity Delirium resolution
ORIC-I 2017[41]	29	Mechanically ventilated patients with delirium	5 mg IV haloperidol every 12 hours	Placebo	Until liberation from mechanical ventilation or 28 days, whichever came first	All-cause mortality Serious adverse reactions QTc prolongation

(Continues)

TABLE 1 (Continued)

Trial	N*	Setting	Intervention	Comparator	Duration of intervention	Outcomes*
Skrobik 2004[42]	73	Patients with delirium admitted to a medical-surgical ICU	Enteral or oral haloperidol. Initially 2.5-5 mg every 8 hours (patients over 60 received a lower initial dose haloperidol 0.5-1 mg) Rescue haloperidol was used	Enteral or oral olanzapine. Initially 5 mg daily (patients over 60 received a lower initial dose olanzapine 2.5 mg) Rescue haloperidol was used	5 days	Delirium severity Extrapyramidal symptoms
Tagarakis 2012[43]	80	Patients with delirium after on-pump cardiac surgery	5 mg IV haloperidol	8 mg IV ondansetron	Unclear	Delirium severity Delirium resolution

*Analysed

**One patient in each group was admitted to an oncology ward

3.4.4 | Quality of life

None of the included trials reported any data on quality of life.

3.4.5 | Cognitive function

One overall high risk of bias trial[38] with 1 comparisons and 11 participants reported on cognitive function measured with Mini-Mental State. Mean end scores at end of intervention were: haloperidol group 17.18 (SD 12.12), chlorpromazine group 15.05 (SD 10.43) and lorazepam 11.50 (SD 8.69). The certainty of evidence was judged to be very low due to serious risk of bias, indirectness and imprecision (Table 2).

3.4.6 | Severity of delirium

Five overall high risk of bias trials[37,38,40,42,43] (7 comparisons; comparing haloperidol with dexmedetomidine in 1, ondansetron in 2, antipsychotics in 3 and benzodiazepine in 1) reported on delirium severity. Two trials used ICDSC,[37,42] 1 trial used delirium rating scale,[38] 1 trial used Memorial Delirium Assessment Scale[40] and 1 trial used a 4 point mental scoring scale.[43] No trials were placebo-controlled and 2 trials used haloperidol as rescue drug.[37,42]

Meta-analysis, regardless of risk of bias, showed no evidence of a difference in haloperidol vs control for the treatment of delirium when assessing delirium severity (random effects model SMD −0.15; 95% CI −0.61-0.30; I^2 = 27%; 134 participants; 3 trials; 4 comparisons; Figure 4). The certainty of evidence was judged to be very low due to serious risk of bias, inconsistency, indirectness and impression (Table 2).

The TSA program does not facilitate meta-analysis of SMDs. SMDs was used because the mean response was not measured on the same scale. We decided not to convert scores into the frequently used scale as 3 different scales (in 3 trials) were used. For the same reason, analyses were not conducted within trials using the same scale. Bayes factor is not possible to calculate from SMD.

The sensitivity analyses on missing data indicated that incomplete outcome data did not have the potential to influence the results (best-worst case scenario and worst-best case scenario (ESM).

The subgroup analysis excluding trials at overall high risk of bias could not be performed as no trial was overall low risk of bias. Subgroup analysis on delirium type could not be performed as none of the 4 trials specified the type of delirium. We found no interaction between intervention effect and use of control intervention, including used control intervention and patient population in subgroup analyses (ESM).

Sensitivity analysis including the trial using rescue haloperidol showed similar results (fixed effect model RR −0.05; 95% CI −0.28-0.19; I^2 = 0%; 303 participants; 5 trials; 7 comparisons; ESM).

TABLE 2 GRADE – Summary of findings of predefined outcomes regardless of overall risk of bias – based on trials not using rescue haloperidol

Certainty assessment							No of patients		Effect			
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Haloperidol	Control	Relative (95% CI)	Absolute (95% CI)	Certainty	importance
All-cause mortality (follow up: range 8 days to 30 days)												
4	Randomized trials	Serious[a]	Not serious[b]	Serious[c]	Serious[d]	Publication bias strongly suspected[e]	5/53 (9.4%)	6/59 (10.2%)	RR 1.01 (0.33 to 3.06)	1 more per 1,000 (from 68 fewer to 209 more)	⨁◯◯◯ VERY LOW	CRITICAL
Serious adverse reactions/events												
4	Randomized trials	Serious[f]	Serious[g]	Serious[h]	Serious[i]	Publication bias strongly suspected[e]	4 trials (5 comparisons) reported zero events in any group. Meta-analysis not performed.				⨁◯◯◯ VERY LOW	CRITICAL
Days alive without delirium												
0	Randomized trials										—	CRITICAL
Quality of life												
0	Randomized trials										—	CRITICAL
Cognitive function												
1	Randomized trials	Serious[l]	Not serious	Serious[m]	Serious[n]	Publication bias strongly suspected[e]	1 trial with 2 comparisons reported on cognitive function. Meta-analysis not performed.				⨁◯◯◯ VERY LOW	CRITICAL
Delirium severity												
3	Randomized trials	Serious [o]	Serious[p,r]	Serious[q]	Serious[e]	Publication bias strongly suspected[e]	63	71	—	SMD 0.15 SD lower (0.61 lower to 0.3 higher)	⨁◯◯◯ VERY LOW	IMPORTANT

(Continued)

TABLE 2 (Continued)

Certainty assessment							No of patients		Effect			
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Haloperidol	Control	Relative (95% CI)	Absolute (95% CI)	Certainty	importance
QTc prolongation												
1	Randomized trials	Serious[p,r]	Not serious[s]	Serious[t]	Serious[u]	Publication bias strongly suspected[e]	1 trial reported on QTc prolongation. Meta-analysis not performed.				⨁◯◯◯ VERY LOW	IMPORTANT

Abbreviations: CI, Confidence interval; RR, Risk ratio; SMD, Standardized mean difference.

[a]3/4 trials had overall high risk of bias.

[b]I^2 = 0%, *P* = .96, overlap of confidence intervals

[c]Trials used different control interventions.

[d]TSA-adjusted confidence interval 0.67-1.83 with the cumulative Z-curve not reaching the trial sequential monitoring boundary and not reaching the futility area.

[e]8 trials identified in trials registers which were either terminated, completed or status unknown and trial results were not available. Serious adverse reactions/events.

[f]4/4 trials had overall high risk of bias.

[g]Trials did not adhere to ICH-GCP.

[h]4/4 trials compared haloperidol to an active drug.

[i]Meta-analysis not performed. Optimal information size could not be calculated

[j]1 comparison compared haloperidol with placebo

[k]Only one trial included.

[l]1/1 trial had overall high risk of bias.

[m]Haloperidol was compared with chlorpromazine and lorazepam.

[n]Only one very small trial included.

[o]3/3 trials had overall high risk of bias.

[p]I^2 = 27%; *P* = .25; overlap of confidence intervals.

[q]Different delirium scales were used.

[r]1/2 trials had overall high risk of bias.

[s]I^2 = 16%; *P* = .31; overlap of confidence intervals.

[t]1/3 comparisons compared haloperidol with an active drug.

[u]TSA was not possible due to too little information. Optimal information size is therefore not met.

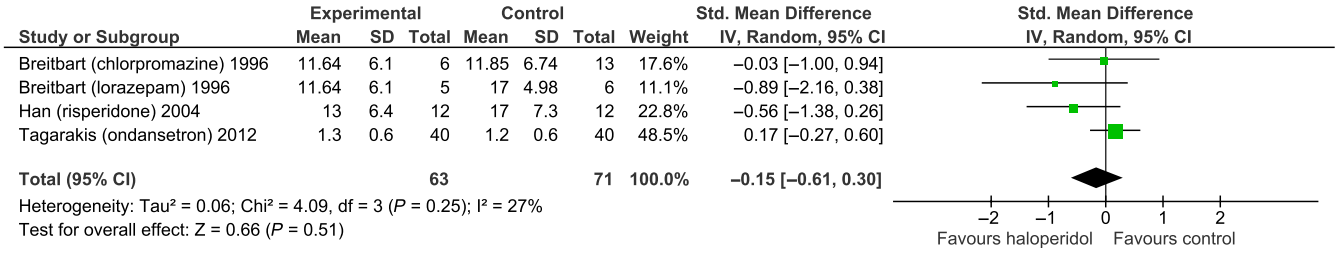


FIGURE 4 Forest plot of delirium severity, excluding trials using rescue haloperidol. No trials were overall low risk of bias. Parenthesis following author name show used control intervention. Size of squares for standardized mean difference reflects weight of trial in pooled analysis. Horizontal bars represent 95% confidence intervals

3.4.7 | QTc prolongation

Three trials,[37,39,41] of which 1 was overall low risk of bias (5 comparisons; comparing haloperidol with placebo in 2, antipsychotics in 2 and dexmedetomidine in 1), reported on QTc prolongation. Two trials used rescue haloperidol.[37,39]

In the trial not using rescue haloperidol, a total of 18.8% of the participants in the haloperidol group vs 7.8% of the participants in the control group had QTc prolongation. The certainty of evidence was very low due to serious risk of bias, indirectness and imprecision (Table 2).

Sensitivity analysis including the trial using rescue haloperidol showed similar results (random effects model RR 0.97; 95% CI 0.48-1.94; I^2 = 16%; 691 participants; 3 trials; 5 comparisons; ESM).

3.4.8 | Post hoc analyses

Post hoc analyses on delirium resolution and extrapyramidal symptoms showed no evidence of a difference of haloperidol vs control for the treatment of delirium when assessing delirium resolution and extrapyramidal symptoms (ESM).

4 | DISCUSSION

The 8 included trials covered 11 comparisons of which the control group was placebo in 2; of which one trial used rescue haloperidol and the other trial only analysed 29 patients. Active comparators were used in the other trials/comparisons and a total of 3 trials used haloperidol as rescue drug. Our primary comparison excluding trials using haloperidol as rescue medication provided very low certainty of evidence to support or refute the use of haloperidol for the treatment of delirium in critically ill patients. The TSA showed that only 1% of the required information size to detect or reject a 20% RRR or RRI in mortality was accrued and 11.237 patients probably need to be randomized before firm conclusion can be drawn for the effect on mortality. The effects on serious adverse reactions/events, days alive without delirium, quality of life, cognitive function, delirium severity and QTc prolongation were also inconclusive due to sparse or no data. Thus, the use of haloperidol as the preferred drug to treat delirium in critically ill patients lacks evidence from RCTs.

4.1 | Strengths and limitation

Strengths of this review include the systematic, transparent and robust methodology used, including a pre-published protocol,[19] the use of Cochrane methodology,[20] reporting as per the PRISMA statement,[17] an up-to-date comprehensive literature search, and the independent study selection, data extraction and risk of bias assessment by 2 authors. Also, we used TSA to assess the overall risk of random error to increase the reliability of the results of the meta-analysis, and to identify the required information size. Finally, we assessed the certainty of evidence using GRADE.

Limitations of our review results include a high risk of clinical heterogeneity between trials. The most obvious reasons are active comparators as only 2 placebo-controlled comparisons were included and the inclusion of trials using rescue haloperidol and diverse patient populations. Furthermore, the use of different delirium screening tools complicate the comparability of the trials included as a participant in one trial may have delirium when assessed with one tool but not when assessed with another tool. Publication bias was detected as we identified 8 trials without results. None of the included trials reported detailed data on serious adverse reactions/events according to the ICH-GCP recommendation;[21] however, 4 trials reported zero serious adverse reactions/events in both groups, although mortality was reported. Accordingly, serious adverse reactions/events are likely to be considerably underreported. Finally, sparse data on all reported outcomes resulted in no firm evidence on the balance between the benefits and harms for these outcomes.

4.2 | Our results in relation to previous reviews

Previous reviews on the treatment of delirium in critically ill patients have been shown not to be systematic according to PRISMA guideline.[15] Besides methodological weaknesses, a common problem with the previous reviews are the inclusion of trials of both prevention (including trials of patients being enrolled regardless of delirium status at enrolment) and treatment of delirium. Furthermore, trials may have been missed and not included and a clear-cut definition of the patient population has often not been adequately described or discussed; for example we decided to exclude the trial by Reade et al,[56] which included patients with delirium or agitation, as only 30/40%

of the participants had delirium at enrolment. Several reviews on either delirium prevention or treatment in all hospitalized patients have been published. However, only a few reviews focusing on delirium treatment in the critically ill patients have been published, and these also found no evidence of effect of haloperidol on the studied outcomes.[16,57-59] Other reviews report on length of ICU and hospital stay, and apart from being biased and not patient centred outcomes such data are not normally distributed and, thus, should not be meta-analysed.

A Cochrane review on antipsychotics for the treatment of delirium in hospitalized patients, however, with the exclusion of ICU patients, did not find evidence for a difference on any of the studied outcomes.[60]

4.3 | Clinical implications and perspectives

Many critically ill patients develop delirium and haloperidol is still the most commonly used pharmacological intervention.[8] In this systematic review, we did not find evidence of neither a beneficial nor a harmful effect of the use of haloperidol and the uncertainty of its effects remains high.

Currently 4 randomized clinical trials are recruiting patients, but especially the AID-ICU trial[46] and the EuRIDICE trial[47] comparing haloperidol with placebo aiming to reach a combined total of 1742 participants will contribute a higher certainty of evidence. Nevertheless, true placebo-controlled trials, using other rescue drugs than haloperidol, reporting on patient-centred outcomes such as all-cause mortality, days alive without delirium, serious adverse reactions/events, HRQoL and cognitive status on delirium treatment are urgently needed.

The lack of evidence on the use of haloperidol for the treatment of delirium challenges the clinicians managing these patients. In spite of the low certainty, we still need to systematically screen and identify critically ill patients with delirium and haloperidol may still be included in the treatment when prevention and non-pharmacological interventions have failed as suggested in the updated guidance.[14,61]

5 | CONCLUSIONS

The evidence for the use of haloperidol to treat critically ill patients with delirium is sparse, of low quality and inconclusive. We therefore have no certainty regarding any beneficial, harmful or neutral effects of haloperidol in these patients. We therefore need many more patients randomized into trials with overall low risk of bias not using haloperidol as rescue drug, to ensure the safety of critically ill patients with delirium.

ACKNOWLEDGEMENTS

We thank De Zhao Kong and Ning Liang for translation of Chinese papers and Dimitrinka Nikolova for translation of Russian papers.

CONFLICTS OF INTEREST

Marija Barbateskovic, Sara Russo Kraus, Janus Christian Jakobsen: Nothing to declare. Nina Christine Andersen-Ranberg: Coordinating investigator of the AID-ICU trial. Ole Mathiesen: Initiator of the AID-ICU trial. Anders Perner: Head of Research in the ICU at Rigshospitalet, which receives support for research from Ferring Pharmaceuticals and the Novo Nordisk Foundation. Dr Perner is initiator of the AID-ICU trial. Jørn Wetterslev: Member of the Copenhagen Trial Unit task force for developing TSA theory, manual and software.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

This information is summarized in the Electronic Supplementary Material.

ORCID

Marija Barbateskovic https://orcid.org/0000-0001-8566-3660
Marie O. Collet https://orcid.org/0000-0002-8387-3960
Nina C. Andersen-Ranberg https://orcid.org/0000-0002-0804-1064
Ole Mathiesen https://orcid.org/0000-0003-4544-0619
Janus C. Jakobsen https://orcid.org/0000-0002-3642-2120
Anders Perner https://orcid.org/0000-0002-4668-0123
Jørn Wetterslev https://orcid.org/0000-0001-7778-1771

REFERENCES

1. Pisani MA, Kong SY, Kasl SV, Murphy TE, Araujo KL, Van Ness PH. Days of delirium are associated with 1-year mortality in an older intensive care unit population. *Am J Respir Crit Care Med*. 2009;180(11):1092-1097.
2. Lin SM, Liu CY, Wang CH, et al. The impact of delirium on the survival of mechanically ventilated patients. *Crit Care Med*. 2004;32(11):2254-2259.
3. Ely EW, Gautam S, Margolin R, et al. The impact of delirium in the intensive care unit on hospital length of stay. *Intensive Care Med*. 2001;27(12):1892-1900.
4. Salluh JI, Wang H, Schneider EB, et al. Outcome of delirium in critically ill patients: systematic review and meta-analysis. *BMJ*. 2015;350:h2538.
5. Ely EW, Shintani A, Truman B, et al. Delirium as a predictor of mortality in mechanically ventilated patients in the intensive care unit. *JAMA*. 2004;291(14):1753-1762.
6. Girard TD, Jackson JC, Pandharipande PP, et al. Delirium as a predictor of long-term cognitive impairment in survivors of critical illness. *Crit Care Med*. 2010;38(7):1513-1520.
7. van den Boogaard M, Schoonhoven L, Evers AW, van der Hoeven JG, van Achterberg T, Pickkers P. Delirium in critically ill patients: impact on long-term health-related quality of life and cognitive functioning. *Crit Care Med*. 2012;40(1):112-118.
8. Collet MO, Caballero J, Sonneville R, et al. Prevalence and risk factors related to haloperidol use for delirium in adult intensive care patients: the multinational AID-ICU inception cohort study. *Intensive Care Med*. 2018;44(7):1081-1089.
9. Mo Y, Zimmermann AE, Thomas MC. Practice patterns and opinions on current clinical practice guidelines regarding the management of delirium in the intensive care unit. *J Pharm Pract*. 2017;30(2):162-171.
10. Devlin JW, Bhat S, Roberts RJ, Skrobik Y. Current perceptions and practices surrounding the recognition and treatment of delirium in

the intensive care unit: a survey of 250 critical care pharmacists from eight states. *Ann Pharmacother*. 2011;45(10):1217-1229.
11. Sweeney RM, Barber V, Page V, et al. A national survey of the management of delirium in UK intensive care units. *QJM*. 2010;103(4):243-251.
12. Jacobi J, Fraser GL, Coursin DB, et al. Clinical practice guidelines for the sustained use of sedatives and analgesics in the critically ill adult. *Crit Care Med*. 2002;30(1):119-141.
13. Barr J, Fraser GL, Puntillo K, et al. Clinical practice guidelines for the management of pain, agitation, and delirium in adult patients in the intensive care unit. *Crit Care Med*. 2013;41(1):263-306.
14. Devlin JW, Skrobik Y, Gelinas C, et al. Clinical practice guidelines for the prevention and management of pain, agitation/sedation, delirium, immobility, and sleep disruption in adult patients in the ICU. *Crit Care Med*. 2018;46(9):e825-e873.
15. Barbateskovic M, Krauss SR, Collet MO, et al. Pharmacological interventions for prevention and management of delirium in intensive care patients: a systematic overview of reviews and meta-analyses. *BMJ Open*. 2019;9(2):e024562.
16. Burry L, Hutton B, Williamson DR, et al. Rose L. Pharmacological interventions for the treatment of delirium in critically ill adults. *Cochrane Database Syst Rev*. 2019;9:Cd011749.
17. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*. 2009;339:b2700.
18. Wetterslev J, Jakobsen JC, Gluud C. Trial Sequential Analysis in systematic reviews with meta-analysis. *BMC Med Res Methodol*. 2017;17(1):39.
19. Barbateskovic M, Kraus SR, Collet MO, et al. Haloperidol for delirium in critically ill patients - protocol for a systematic review. *Acta Anaesthesiol Scand*. 2018;62(5):712-723.
20. Higgins JPT, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboratio*. 2011. Available from www.handbook.cochrane.org. Accessed June 4, 2019.
21. International Conference on Harmonisation Expert Working GroupInternational Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH Harmonised Tripartite Guideline. Guideline for Good Clinical Practice CFR & ICH Guidelines. Vol. 1, Philadelphia (PA): Barnett International/PAREXEL, 1997.
22. Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC Med Res Methodol*. 2014;14:120.
23. Harbord RM, Egger M, Sterne JA. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Stat Med*. 2006;25(20):3443-3457.
24. RevMan, Review Manager (RevMan) (Computer program), Version 5.3, Copenhagen: The Nordic Cochrane Center 2014, The Cochrane Collaboration, Accessed 20 June 2019, Available from: http://community.cochrane.org/tools/review-production-tools/revman-5/revman-5-download
25. Wetterslev J, Thorlund K, Brok J, Gluud C. Estimating required information size by quantifying diversity in random-effects model meta-analyses. *BMC Med Res Methodol*. 2009;9:86.
26. Brok J, Thorlund K, Gluud C, Wetterslev J. Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses. *J Clin Epidemiol*. 2008;61(8):763-769.
27. Brok J, Thorlund K, Wetterslev J, Gluud C. Apparently conclusive meta-analyses may be inconclusive - Trial sequential analysis adjustment of random error risk due to repetitive testing of accumulating data in apparently conclusive neonatal meta-analyses. *Int J Epidemiol*. 2009;38(1):287-298.
28. Higgins JP, Whitehead A, Simmonds M. Sequential methods for random-effects meta-analysis. *Stat Med*. 2011;30(9):903-921.
29. Imberger G, Gluud C, Boylan J, Wetterslev J. Systematic Reviews of Anesthesiologic Interventions Reported as Statistically Significant: Problems with Power, Precision, and Type 1 Error Protection. *Anesth Analg*. 2015;121(6):1611-1622.
30. Mascha EJ. Alpha, Beta, Meta: Guidelines for Assessing Power and Type I Error in Meta-Analyses. *Anesth Analg*. 2015;121(6):1430-1433.
31. Pogue JM, Yusuf S. Cumulating evidence from randomized trials: utilizing sequential monitoring boundaries for cumulative meta-analysis. *Control Clin Trials*. 1997;18(6):580-593; discussion 661–6.
32. Terkawi AS, Mavridis D, Flood P, et al. Does ondansetron modify sympathectomy due to subarachnoid anesthesia?: meta-analysis, meta-regression, and Trial Sequential Analysis. *Anesthesiology*. 2016;124(4):846-869.
33. Thorlund K, Devereaux PJ, Wetterslev J, et al. Can trial sequential monitoring boundaries reduce spurious inferences from meta-analyses? *Int J Epidemiol*. 2009;38(1):276-286.
34. Thorlund K, Imberger G, Johnston BC, et al. Evolution of heterogeneity (I2) estimates and their 95% confidence intervals in large meta-analyses. *PLoS ONE*. 2012;7(7):e39471.
35. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. *Bmj*. 2004;328(7454):1490.
36. Atalan N, Efe Sevim M, Akgun S, Fazliogullari O, Basaran C. Morphine is a reasonable alternative to haloperidol in the treatment of postoperative hyperactive-type delirium after cardiac surgery. *J Cardiothorac Vasc Anesth*. 2013;27(5):933–938.
37. Bakri MH, Ismail EA, Ibrahim A. Comparison of dexmedetomidine or ondansetron with haloperidol for treatment of postoperative delirium in trauma patients admitted to intensive care unit: randomized controlled trial. *Anaesth, pain & intensive care*. 2015;19(2):118–123.
38. Breitbart W, Marotta R, Platt MM, et al. A double-blind trial of haloperidol, chlorpromazine, and lorazepam in the treatment of delirium in hospitalized AIDS patients. *Am J Psychiatry*. 1996;153(2):231–237.
39. Girard TD, Exline MC, Carson SS, et al. Haloperidol and ziprasidone for treatment of delirium in critical illness. *N Engl J Med*. 2018;379(26):2506–2516.
40. Han CS, Kim YK. A double-blind trial of risperidone and haloperidol for the treatment of delirium. *Psychosomatics*. 2004;45(4):297–301.
41. NCT00300391 ORIC-I: Optimizing Recovery From Intensive Care: Mechanical Ventilation and Delirium, https://clinicaltrials.gov/ct2/show/NCT00300391?term=oric-I&rank=1 (accessed 4 June 2019).
42. Skrobik YK, Bergeron N, Dumont M, Gottfried SB. Olanzapine vs haloperidol: treating delirium in a critical care setting. *Intensive Care Med*. 2004;30(3):444–449.
43. Tagarakis GI, Voucharas C, Tsolaki F, et al. Ondasetron versus haloperidol for the treatment of postcardiotomy delirium: a prospective, randomized, double-blinded study. *J Cardiothorac Surg*. 2012;7:25.
44. IRCT20180911040998N1, Comparison the Effect of Quetiapine and Haloperidol on the Treatment of delirium in ICU, http://en.irct.ir/trial/33804. Accessed 4 June 2019.
45. IRCT20121231011956N10, Comparison the Effectiveness of Haloperidol and quetiapine for delirium in the Emergency department and intensive Care Unit, http://en.irct.ir/trial/29718. Accessed 4 June 2019.
46. NCT03392376, Agents Intervening Against Delirium in Intensive Care Unit (AID-ICU), https://clinicaltrials.gov/ct2/show/NCT03392376. Accessed 4 June 2019.
47. NCT03628391, Efficacy of Haloperidol to Decrease the Burden of Delirium in Adult Critically Ill Patients (EuRIDICE): a Prospective Randomised Multi-center Double-blind Placebo-controlled Clinical Trial, https://clinicaltrials.gov/ct2/show/NCT03628391. Accessed 4 June 2019.

48. NCT02343575, Valproic Acid for Treatment of Hyperactive or Mixed Delirium in ICU, https://clinicaltrials.gov/ct2/show/NCT02343575. Accessed 4 June 2019.
49. NCT02345902, Treatment of Hypoactive Delirium and Outcome Measures (THDOM), https://clinicaltrials.gov/ct2/show/NCT02345902. Accesssed 4 June 2019.
50. NCT01811459, Trial Comparing Haloperidol, Quetiapine and Placebo in the Pharmacological Treatment of Delirium (Haloquet), https://clinicaltrials.gov/ct2/show/NCT01811459. Accessed 4 June 2019.
51. NCT01140529, Dexmedetomidine for the Treatment of Delirium After Heart Surgery (DexinDelir), https://clinicaltrials.gov/ct2/show/NCT01140529. Accessed 4 June 2019.
52. NCT00833300, Haloperidol vs Olanzapine for the Management of ICU Delirium, https://clinicaltrials.gov/ct2/show/NCT00833300. Accessed 4 June 2019.
53. NCT00599287, Methylphenidate, Rivastigmine or Haloperidol in Hypoactive Delirium in Intensive Care Patients, https://clinicaltrials.gov/ct2/show/NCT00599287. Accessed 4 June 2019.
54. ACTRN12606000085572, Dexmedetomidine and Haloperidol for the management of emergence delirium in intensive care (DeHedic), https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?xml:id=1140. Accessed 4 June 2019.
55. Trial NL495 (NTR537), Delirium treatment at the surgical ward (DELTa S) Treatment of delirium: rivastigmine or haloperidol as primary treatment for delirium in elderly patients with a fractured hip. A randomized placebo-controlled study, https://www.trialregister.nl/trial/495. Accessed 4 June 2019.
56. Reade MC, O'Sullivan K, Bates S, Goldsmith D, Ainslie WR, Bellomo R. Dexmedetomidine vs. haloperidol in delirious, agitated, intubated patients: a randomised open-label trial. *Crit Care*. 2009;13(3):R75.
57. Zayed Y, Barbarawi M, Kheiri B, et al. Haloperidol for the management of delirium in adult intensive care unit patients: A systematic review and meta-analysis of randomized controlled trials. *J Crit Care*. 2019;50:280–286.
58. Bathula M, Gonzales JP. The pharmacologic treatment of intensive care unit delirium: a systematic review. *Ann Pharmacother*. 2013;47(9):1168–1174.
59. Serafim RB, Bozza FA, Soares M, et al. Pharmacologic prevention and treatment of delirium in intensive care patients: A systematic review. *J Crit Care*. 2015;30(4):799–807.
60. Burry L, Mehta S, Perreault MM, et al. Antipsychotics for treatment of delirium in hospitalised non-ICU patients. *Cochrane Database Syst Rev*. 2018;6:Cd005594.
61. European Medicines Agency. Haloperidol—Assessment report; 2017. https://www.ema.europa.eu/en/documents/referral/haldol-article-30-referral-assessment-report_en.pdf. Accessed 4 June 2019.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Barbateskovic M, Krauss SR, Collet MO, et al. Haloperidol for the treatment of delirium in critically ill patients: A systematic review with meta-analysis and Trial Sequential Analysis. *Acta Anaesthesiol Scand*. 2019;00:1–13. https://doi.org/10.1111/aas.13501

PAPER III

Intensive Care Med (2019) 45:143–158
https://doi.org/10.1007/s00134-019-05526-z

SYSTEMATIC REVIEW

Stress ulcer prophylaxis with proton pump inhibitors or histamin-2 receptor antagonists in adult intensive care patients: a systematic review with meta-analysis and trial sequential analysis

Marija Barbateskovic[1,2*], Søren Marker[2,3], Anders Granholm[3], Carl Thomas Anthon[3], Mette Krag[3], Janus Christian Jakobsen[1,2,4], Anders Perner[2,3], Jørn Wetterslev[1,2] and Morten Hylander Møller[2,3]

Abstract

Purpose: Most intensive care unit (ICU) patients receive stress ulcer prophylaxis. We present updated evidence on the effects of prophylactic proton pump inhibitors (PPIs) or histamine 2 receptor antagonists (H2RAs) versus placebo/no prophylaxis on patient-important outcomes in adult ICU patients.

Methods: We conducted a systematic review with meta-analysis and trial sequential analysis (TSA) of randomised clinical trials assessing the effects of PPI/H2RA versus placebo/no prophylaxis on mortality, gastrointestinal (GI) bleeding, serious adverse events (SAEs), health-related quality of life (HRQoL), myocardial ischemia, pneumonia, and *Clostridium (Cl.) difficile* enteritis in ICU patients.

Results: We identified 42 trials randomising 6899 ICU patients; 3 had overall low risk of bias. We did not find an effect of stress ulcer prophylaxis on mortality [relative risk 1.03, 95% confidence interval (CI) 0.94–1.14; TSA-adjusted CI 0.94–1.14], but the occurrence of any GI bleeding was reduced as compared with placebo/no prophylaxis (0.60, 95% CI 0.47–0.77; TSA-adjusted CI 0.36–1.00). The conventional meta-analysis indicated that clinically important GI bleeding was reduced (RR 0.63, 95% CI 0.48–0.81), but the TSA-adjusted CI 0.35–1.13 indicated lack of firm evidence. The effects of stress ulcer prophylaxis on SAEs, HRQoL, pneumonia, myocardial ischemia and *Cl. difficile* enteritis are uncertain.

Conclusions: In this updated systematic review, we were able to refute a relative change of 20% of mortality. The occurrence of GI bleeding was reduced, but we lack firm evidence for a reduction in clinically important GI bleeding. The effects on SAEs, HRQoL, pneumonia, myocardial ischemia and *Cl. difficile* enteritis remain inconclusive.

Keywords: Critical care, Peptic ulcer, Gastrointestinal haemorrhage, Meta-analysis, Proton pump inhibitors, Histamine-2 receptor antagonists, Stress ulcer prophylaxis

*Correspondence: marija.barbateskovic@ctu.dk
[1] Copenhagen Trial Unit, Centre for Clinical Intervention Research, 2100 Copenhagen, Denmark
Full author information is available at the end of the article

Springer

Introduction

Patients admitted to the intensive care unit (ICU) are at risk of stress-related gastrointestinal (GI) mucosal damage that may evolve to ulceration and bleeding [1]. The reported prevalence of GI bleeding ranges from 5 to 10% in recent reports, and GI bleeding is associated with an increased risk of death and length of stay in the ICU [2–5]. Stress ulcer prophylaxis is routinely used in the ICU, even though recommendations in international guidelines are conflicting [6, 7]. However, the quantity and quality of evidence supporting use of stress ulcer prophylaxis in adult ICU patients is low with no firm evidence for benefit or harm [8, 9]. Importantly, increased rates of myocardial ischaemia, *Clostridium (Cl.) difficile* enteritis and hospital-acquired pneumonia with the use of stress ulcer prophylaxis have been suggested [1, 8, 10, 11]. Several randomised clinical trials (RCT) and systematic reviews have compared the effects of proton pump inhibitors (PPIs) and histamine-2-receptor antagonist (H2RAs), but neither PPIs nor H2RAs have demonstrated superiority as compared with placebo or no prophylaxis [10, 12–15].

Recently, new relevant trials, including the SUP-ICU trial, have been published [3, 5, 16–18]. Consequently, we performed an updated systematic review on stress ulcer prophylaxis with PPI or H2RA versus placebo or no prophylaxis in adult ICU patients. We hypothesised an absence of effect on mortality, a reduction of GI bleeding, and an increase of infectious adverse events and myocardial ischemia.

Take-home message

Stress ulcer prophylaxis with PPI or H2RA did not seem to affect mortality, but likely reduced the occurrence of gastrointestinal bleeding.

Methods

We conducted this systematic review according to the preplanned statistical analysis plan of the published protocol [19]. We registered the protocol in the international prospective register of systematic reviews database (PROSPERO) (CRD42018089151) and used the methodology of the Cochrane Collaboration [20], the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) [S1, Electronic Supplementary Material, (ESM)] [21], Keus et al. [22], Jakobsen et al. [23], and Grading of Recommendations Assessment, Development, and Evaluation (GRADE) [24].

Eligibility criteria

We included any RCT comparing stress ulcer prophylaxis with either PPI or H2RA versus placebo or no prophylaxis in adult ICU patients. We accepted any dose, formulation and duration of intervention [19].

Search methods for identification of studies

We did not restrict the search by language, date, publication status or any other trial characteristics. MB searched the following electronic databases: Cochrane Central Register of Controlled Trials (CENTRAL) in the Cochrane Library; Ovid MEDLINE; Ovid Embase; Science Citation Index Expanded (Web of Science); Biosis Previews (Web of Science); and PubMed. The systematic search included the following keywords: peptic ulcer; gastrointestinal haemorrhage; proton pumps; histamine h2 receptor antagonists; critical illness; critical care; intensive care units; artificial respiration; craniocerebral trauma; heart arrest; myocardial infarction; sepsis; and surgery. The full search is available in the ESM. The literature search was updated on 11 October 2018. We manually identified additional potential eligible trials by screening the reference lists of the included studies, other relevant systematic reviews, and searched trial registries.

Selection of studies

At least two authors (MB, SM, AG or CTA) independently screened each title and abstract. Reports deemed potentially relevant were obtained in full-text and assessed for inclusion in accordance with the inclusion criteria. Disagreements were resolved by consensus and MHM/JW were consulted when agreement could not be met.

Data extraction and management

Two review authors (MB and SM) independently extracted predefined data of the included trials using a predefined data collection form (S2, ESM). The following data were collected: (1) Trial: country, duration of the trial, date of publication, and type of trial (single versus multi centre); (2) Participants: numbers randomised, numbers analysed, numbers lost to follow-up/withdrawn, type of population, mean/median age, sex, inclusion criteria, and exclusion criteria; (3) Interventions: intervention, comparator, and concomitant interventions; (4) Outcomes: predefined primary and secondary outcomes [19].

Outcomes

Predefined co-primary outcomes were all-cause mortality and the proportion of participants with any GI bleeding (overt and clinically important bleeding defined by trialists). Co-secondary outcomes were: the proportion of participants with one or more serious adverse events (SAEs) (as defined by trialists using the term 'serious

adverse event', 'severe adverse event', 'serious adverse reaction', 'serious complication', 'severe complication' or similar terms fulfilling the criteria of the Good Clinical Practice Guideline of the International Conference on Harmonization (ICH-GCP) definition [25]); health-related quality of life (HRQoL) (any valid scale used by trialists); proportion of participants with myocardial ischemia (as defined by trialists); proportion of participants with hospital-acquired pneumonia (as defined by trialists); proportion of participants with *CI. difficile* enteritis (as defined by trialists).

For all outcomes, we used the trial results reported at time-points closest to 90 days.

Risk of bias

MB and SM independently assessed the risk of systematic errors (bias) in the included trials using the Cochrane Collaboration's risk of bias tool [20], with additional prespecified criteria (ESM) [19]. Two review contributors not involved in the SUP-ICU trial [3] assessed risk of bias and extracted data from this trial. We specifically assessed the following domains: (1) random sequence generation; (2) allocation concealment; (3) blinding of participants and personnel; (4) blinding of outcome assessment; (5) incomplete outcome data; (6) selective reporting; and (7) other biases, including baseline imbalance, early stopping and bias due to vested financial interest or academic bias. The included trials were judged as 'overall low risk of bias' when all bias domains were judged as low risk of bias. Conversely, trials were judged as 'overall high risk of bias' when unclear or high risk of bias was judged in one or more domains [26].

We assessed publication bias by inspecting funnel plots for signs of asymmetry when ten or more trials were included in an analysis [20, 23]. We tested asymmetry with the Harbord test [27].

Data synthesis

Summary measures

We calculated relative risks (RRs) with 95% confidence intervals (CIs) and trial sequential analysis (TSA)-adjusted CIs [28] for all outcomes. We hypothesised an absence of effect on mortality, a reduction of GI bleeding, and an increase of infectious adverse events and myocardial ischemia, assuming a required information size corresponding to a relative risk reduction (RRR) or a relative risk increase (RRI) of 20% [19, 29].

Meta-analyses

The primary analysis included trials with overall low risk of bias. We calculated pooled effect estimates using Review Manager [30]. We considered a P value of $0.05/[(2+1)/2] = 0.033$ or less as statistically significant in the analyses of each primary outcome, and we considered a P value of $0.05/[(5+1)/2] = 0.017$ or less as statistically significant in the analyses of each secondary outcome, in order to restrict the family-wise error rates (FWER) to 0.05 [23]. We calculated Bayes factor to assess if the summary effect estimates fitted better with the null hypothesis than alternative hypotheses of the anticipated intervention effects [23].

Dealing with missing data

Corresponding authors were contacted to clarify important missing data related to the methods, data reporting, or if further trial details were needed (S4, ESM).

We conducted a predefined sensitivity analysis by imputing missing outcome data in a best-/worst-case scenario and a worst-/best-case scenario to assess the potential impact of loss to follow-up. In the best-/worst-case scenario analysis, it was assumed that all participants lost to follow-up in the experimental group did not experience the event, and that all those with missing outcomes in the control group did experience the event. In the worst-/best-case scenario analysis, it was assumed that all participants lost to follow-up in the experimental group did not experience the event, and that all those with missing outcomes in the control group did experience the event [19, 23].

Assessment of heterogeneity

We assessed heterogeneity by visual inspection of the forest plots, the inconsistency statistics (I^2) and the estimates of diversity (D^2) [31]. When $I^2 = 0$, we used a fixed effects model [32, 33], and when I^2 was above zero, we used both fixed and random effects models [32, 34, 35], and reported the most conservative estimate being the point estimate closest to no effect or the estimate with the widest CI.

Subgroup analyses

We planned to conduct the following predefined subgroup analyses: high versus low risk of bias; medical versus surgical versus mixed ICU setting; shock versus no shock; renal replacement therapy (RRT) versus no RRT; invasive mechanical ventilation versus no invasive mechanical ventilation versus unknown status; PPI versus H2RA; and placebo versus no prophylaxis [19]. In addition, we conducted post hoc subgroup analyses on the co-primary outcomes: one according to a dose of PPI (max 40 mg daily versus > 40 mg daily) and one according to publication year (median publication date 1993/1994). We accepted the definitions used in the included trials, and only trials defining subgroups on a trial level were included. Presence of statistical heterogeneity was

assessed by the χ^2 test with significance set at $P<0.10$ [19].

Sensitivity analyses

We conducted a sensitivity analysis to assess the potential impact of reporting bias by excluding trials not reporting on clinically important bleeding [19].

In two post hoc sensitivity analyses, we estimated the number of patients with one or more SAEs: (1) highest proportion of reported SAEs in each trial, and (2) all reported SAEs cumulated in each trial (information available in the ESM).

Trial sequential analysis

TSA is a sequential meta-analysis considering how much information (randomised patients) is needed to conclude on a specific a priori anticipated intervention effect in updated, repetitive testing meta-analyses. If information size is smaller than required in the meta-analysis, the TSA-adjusted CI becomes wider than the conventional naïve, meta-analytic 95% CI, and the threshold for statistical significance becomes more restrictive. However, if the required information size is reached, the TSA-adjusted CI and the naïve CI, anticipating a specific intervention effect, becomes identical.

We used TSA to assess the risk of random errors due to sparse data and multiple testing of accumulating data [36–44], and to calculate the required information size [31]. The calculated required information size takes into account the control event proportion, the anticipated heterogeneity variance (D^2) [22] of the meta-analysis, and the assumption of a plausible RRR or RRI.

We used a FWER of 5% [23] leading to a statistical significance level of 3.3% and 96.7% CIs for each of the two co-primary outcomes and 1.7% and 98.3% CIs, respectively, for each of the five co-secondary outcomes [19]. We used a beta of 10%, and a D^2 [31] as suggested by the trials in the meta-analysis [23], or a D^2 of 20% if the measured heterogeneity was zero [45]. As anticipated intervention effects for the primary and secondary outcomes in the TSA, we used a realistic a priori RRR or RRI of 20%. Furthermore, we used an RRR or RRI based on the 95% confidence limit closest to a null effect in the traditional meta-analysis [19]. In addition, we have made a TSA anticipating a 15% RRR of mortality on the meta-analysis of new trials published after our first review [34].

We present 95% CIs and TSA-adjusted CIs, adjusted for multiplicity of outcomes, sparse data, and repetitive testing for all estimates. For a more detailed description of the statistical analysis plan and TSA, we refer to the published protocol [19].

Grading quality of evidence

We used the GRADE approach [24] to assess the overall certainty of evidence for all outcomes. We appraised the certainty of evidence and our confidence in the effect estimates based on risk of bias, inconsistency, indirectness, imprecision and publication bias. Thus, we rated the overall certainty of evidence as high, moderate, low or very low.

Results

Study selection

We identified 10,054 references (Fig. 1) and included 41 RCTs [3–5, 12, 16–18, 46–79] with a total of 6790 participants. Some 37 trials were in English, 2 in German [75, 78], 1 in Portuguese [54], and 1 in French [61].

Characteristics of the included studies

The included trials were published between 1977 and 2018. Some 35 trials were published as full trial reports and 6 as conference abstracts. The 41 included trials covered 44 trial comparisons; 32 trials assessed H2RAs and 12 assessed PPIs. The control group was placebo in 31 trials and no prophylaxis in 13 trials. Details and additional information of the included trials are presented in S3 and S4, ESM. Characteristics of the excluded studies and ongoing trials are summarised in S5, ESM.

Risk of bias assessment

Three trials were judged as having overall low risk of bias [3–5]; the remaining 38 all had overall high risk of bias (Figs. 2 and S4 in the ESM) [12, 16–18, 46–79].

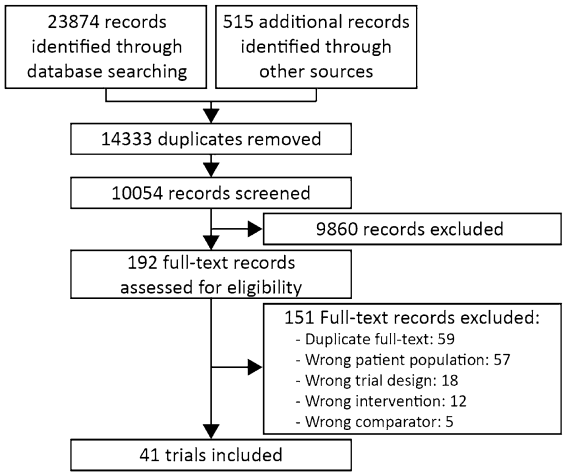


Fig. 1 PRISMA flowchart

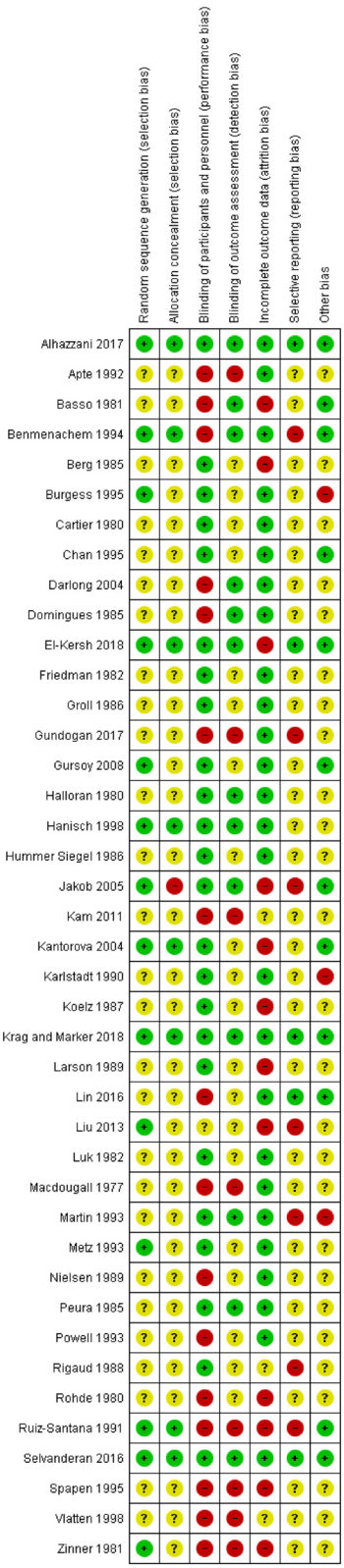


Fig. 2 Risk of bias summary as per the Cochrane Handbook. *Green* represents a low risk of bias, *yellow* an unclear risk of bias, and *red* a high risk of bias

Outcomes

Mortality

A total of 28 trials with 5656 participants reported data on all-cause mortality, including the 3 trials with overall low risk of bias with 3587 participants.

The meta-analysis of the three trials with overall low risk of bias did not show any difference in all-cause mortality between stress ulcer prophylaxis and placebo/no prophylaxis: RR 1.03 (95% CI 0.94, 1.14; $P=0.52$; $I^2=0\%$; TSA-adjusted CI 0.94, 1.14; Bayes factor 239,649) (Fig. 3) (S6–S9, ESM). TSA showed that the boundary for futility was crossed, indicating firm evidence for no difference in mortality between the groups. The certainty of evidence, using the GRADE approach, was high (Table 1).

The corresponding summary estimate of all 28 trials ($n=5656$) regardless of risk of bias was RR 1.01 (95% CI 0.93, 1.10; $P=0.75$; $I^2=0\%$; TSA-adjusted CI was 0.93, 1.10; Bayes factor 941,833) (Fig. 3).

The sensitivity analyses on missing data were consistent with the primary analysis (S10–S11, ESM), and Harbord's test did not indicate asymmetry [$P=0.83$ (S12, ESM)]. The certainty of evidence was moderate due to risk of bias (Table 1).

The subgroup analyses of PPI versus H2RA and placebo versus no prophylaxis showed no interaction (Table S6, ESM). We observed an interaction in the subgroup analysis of ICU setting (test-of-interaction $P=0.08$), suggesting that surgical ICU patients had lower risk of mortality with stress ulcer prophylaxis, compared with medical ICU patients (S6, ESM). Additional subgroup analyses were consistent with the primary analysis (Table S6, ESM). The subgroup analyses of RRT versus no RRT and shock versus no shock could not be performed as no trials (nor stratified subgroups) were eligible for inclusion in these analyses. In the post hoc subgroup analyses of dosing of PPI and publication year, there was no interaction (Table S6, ESM). TSA anticipating a 15% RRR showed that the boundary for futility was crossed, indicating firm evidence for no difference in mortality between the groups (S8, ESM).

GI bleeding

A total of 39 trials with 6627 participants reported on GI bleeding, including the three trials with overall low risk of bias with 3596 participants.

The meta-analysis of the three trials with overall low risk of bias showed a reduction in GI bleeding with stress ulcer prophylaxis versus placebo/no prophylaxis: RR 0.60 (95% CI 0.47, 0.77; $P<0.0001$; $I^2=0\%$; TSA-adjusted CI 0.36, 1.00; Bayes factor 0.004) (Fig. 4), and TSA showed that the required information size to detect a 20% relative

(See figure on next page.)

Fig. 3 **a** Forest plot of mortality in trials with overall low risk of bias versus trials with overall high risk of bias. *Size of squares* for risk ratio reflects weight of trial in pooled analysis. *Horizontal bars* represent 95% confidence intervals. **b** Trial sequential analysis of all 28 trial regardless of risk of bias of the effect of proton pump inhibitors/histamine 2 receptor antagonists versus placebo/no prophylaxis on mortality using a control event proportion of 26.7% (from the included trials), a diversity (D2) of 0%, an alpha of 3.3%, a power of 90%, and a relative risk reduction of 20%. The relative risk was 1.01 with a TSA-adjusted CI 0.93, 1.10. The required information size of 2985 was reached, suggesting that a 20% relative risk increase/reduction can be excluded"

difference had been reached (S13, ESM). The certainty of evidence was high (Table 1).

The corresponding summary effect estimate of all 39 trials ($n=6627$) regardless of risk of bias was RR 0.52 (95% CI 0.45, 0.61; $P<0.00{,}001$; $I^2=43\%$; TSA-adjusted CI 0.39, 0.68; Bayes factor 9×10^{-9}) and TSA showed that the required information size to detect a 20% relative difference had been reached (Fig. 4).

The sensitivity analyses on missing data were consistent with the primary analysis (S10 and S11, ESM), and Harbord's test did not indicate asymmetry [$P=0.33$ (S16, ESM)]. The certainty of evidence was low due to risk of bias and inconsistency (Table 1).

The subgroup analyses of PPI versus H2RA and placebo versus no prophylaxis showed no interaction (Table S6, ESM). Additional subgroup analyses were consistent with the primary analysis (S6, S14 and S15, ESM). In the post hoc subgroup analyses of dosing of PPI and publication year, there was no interaction (Table S6, ESM).

A total of 14 trials ($n=4833$) reported on clinically important GI bleeding. The meta-analysis showed a reduction in clinically important GI bleeding with stress ulcer prophylaxis versus placebo/no prophylaxis: RR 0.63 (95% CI 0.48, 0.81; $P=0.0005$; $I^2=1\%$, Bayes factor 0.017) (S17, ESM). However, this was not confirmed by TSA (TSA-adjusted CI 0.35, 1.13), indicating that the required information size to detect or reject a 20% relative difference had not been reached (S18, ESM).

Serious adverse events

Four trials (three with overall low risk of bias, $n=3587$ participants) reported on SAEs [3, 12, 52, 64], although not defining the adverse events according to ICH-GCP. All four trials reported zero events in each group despite reporting mortality and GI bleeding.

A total of 42 trials reported on outcomes categorised by us as SAEs according to the ICH-GCP definition [25] (S19 and S24, ESM).

The two post hoc analyses estimating the number of patients with one or more SAEs were inconclusive. Details of the analyses are available in S19–S29, ESM . The certainty of evidence was judged to be low/very low due to risk of bias, inconsistency, imprecision, very serious indirectness and strongly suspected publication bias (Table 1).

Health-related quality of life

No trials reported data on HRQoL.

Myocardial ischaemia

We identified one trial (low risk of bias, 3291 participants) which reported on myocardial ischaemia [3]; RR 1.07 (95% CI 0.85, 1.61). TSA highlighted that only 11% of the required information size had been reached. The certainty of evidence was judged to be low due to very serious imprecision (Table 1).

Hospital-acquired pneumonia

A total of 16 trials with 4951 participants reported data on pneumonia, including the three trials with overall low risk of bias with 3596 participants.

The meta-analysis of the three trials with overall low risk of bias showed no difference in hospital-acquired pneumonia between stress ulcer prophylaxis and placebo/no prophylaxis: RR 1.01 (95% CI 0.87, 1.18; $P=0.64$; $I^2=0\%$; TSA-adjusted CI 0.77, 1.33; Bayes factor 82) (S30 and S31, ESM), and TSA showed that only 52% of the required information size had been reached. The certainty of evidence was moderate due to imprecision (Table 1).

The corresponding summary estimate of all 16 trials ($n=4951$) regardless of risk of bias was RR 1.07 (95% CI 0.94, 1.21; $P=0.34$; $I^2=0\%$; TSA-adjusted CI 0.89, 1.27; Bayes factor 7465) (S32 and S33, ESM), and TSA showed that only 70% of the required information size had been reached. The sensitivity analyses of missing data were consistent with the primary analysis (S34 and S35, ESM). Harbord's test did not indicate asymmetry [$P=0.17$ (S36, ESM)]. The certainty of evidence was low due to risk of bias and imprecision (Table 1).

The subgroup analyses of PPI versus H2RA and placebo versus no prophylaxis showed no interaction (Table S6, ESM). Additional subgroup analyses were consistent with the primary analysis; however, there was interaction in the analysis of ICU setting (test-of-interaction $P=0.06$), suggesting that medical ICU patients had higher risk of hospital-acquired pneumonia, compared with surgical or mixed ICU patients (S6, ESM).

a

Study or Subgroup	PPI/H2RA Events	PPI/H2RA Total	Control Events	Control Total	Weight	Risk Ratio M-H, Fixed, 95% CI
1.2.1 Trials with overall low risk of bias						
Alhazzani 2017	17	49	13	42	1.9%	1.12 [0.62, 2.03]
Krag and Marker 2018	510	1642	499	1640	67.8%	1.02 [0.92, 1.13]
Selvanderan 2016	30	106	25	108	3.4%	1.22 [0.77, 1.93]
Subtotal (95% CI)		**1797**		**1790**	**73.0%**	**1.03 [0.94, 1.14]**
Total events	557		537			
Heterogeneity: Chi² = 0.64, df = 2 (P = 0.72); I² = 0%						
Test for overall effect: Z = 0.64 (P = 0.52)						
1.2.2 Trials with overall high risk of bias						
Apte 1992	11	16	7	18	0.9%	1.77 [0.91, 3.44]
Benmenachem 1994	28	100	19	100	2.6%	1.47 [0.88, 2.46]
Burgess 1995	1	16	0	18	0.1%	3.35 [0.15, 76.93]
El-Kersh 2018	7	55	8	47	1.2%	0.75 [0.29, 1.91]
Groll 1986	13	114	13	107	1.8%	0.94 [0.46, 1.93]
Gursoy 2008	6	60	2	15	0.4%	0.75 [0.17, 3.35]
Halloran 1980	8	26	10	24	1.4%	0.74 [0.35, 1.56]
Hanisch 1998	7	57	12	57	1.6%	0.58 [0.25, 1.37]
Jakob 2005	5	20	5	20	0.7%	1.00 [0.34, 2.93]
Kantorova (H2RA) 2004	11	71	6	37	1.1%	0.96 [0.38, 2.38]
Kantorova (PPI) 2004	14	72	7	38	1.2%	1.06 [0.47, 2.39]
Karlstadt 1990	5	54	2	33	0.3%	1.53 [0.31, 7.43]
Lin 2016	2	60	0	60	0.1%	5.00 [0.25, 102.00]
Liu (H2RA) 2013	14	58	10	26	1.9%	0.63 [0.32, 1.22]
Liu (PPI) 2013	17	58	10	27	1.9%	0.79 [0.42, 1.49]
Macdougall 1977	20	26	31	36	3.5%	0.89 [0.70, 1.14]
Martin 1993	8	65	7	66	0.9%	1.16 [0.45, 3.02]
Nielsen 1989	0	12	0	13		Not estimable
Peura 1985	7	21	7	18	1.0%	0.86 [0.37, 1.98]
Powell (H2RA) 1993	0	11	0	5		Not estimable
Powell (PPI) 1993	1	20	0	5	0.1%	0.86 [0.04, 18.45]
Rohde 1980	7	14	6	14	0.8%	1.17 [0.52, 2.60]
Ruiz-Santana 1991	7	19	7	30	0.7%	1.58 [0.66, 3.79]
Spapen 1995	4	20	2	10	0.4%	1.00 [0.22, 4.56]
Zinner 1981	9	100	17	100	2.3%	0.53 [0.25, 1.13]
Subtotal (95% CI)		**1145**		**924**	**27.0%**	**0.96 [0.82, 1.13]**
Total events	212		188			
Heterogeneity: Chi² = 16.53, df = 22 (P = 0.79); I² = 0%						
Test for overall effect: Z = 0.46 (P = 0.64)						
Total (95% CI)		**2942**		**2714**	**100.0%**	**1.01 [0.93, 1.10]**
Total events	769		725			
Heterogeneity: Chi² = 18.04, df = 25 (P = 0.84); I² = 0%						
Test for overall effect: Z = 0.31 (P = 0.75)						
Test for subgroup differences: Chi² = 0.53, df = 1 (P = 0.47), I² = 0%						

Risk Ratio M-H, Fixed, 95% CI

0.01 0.1 1 10 100

Favours PPI/H2RA Favours control

b

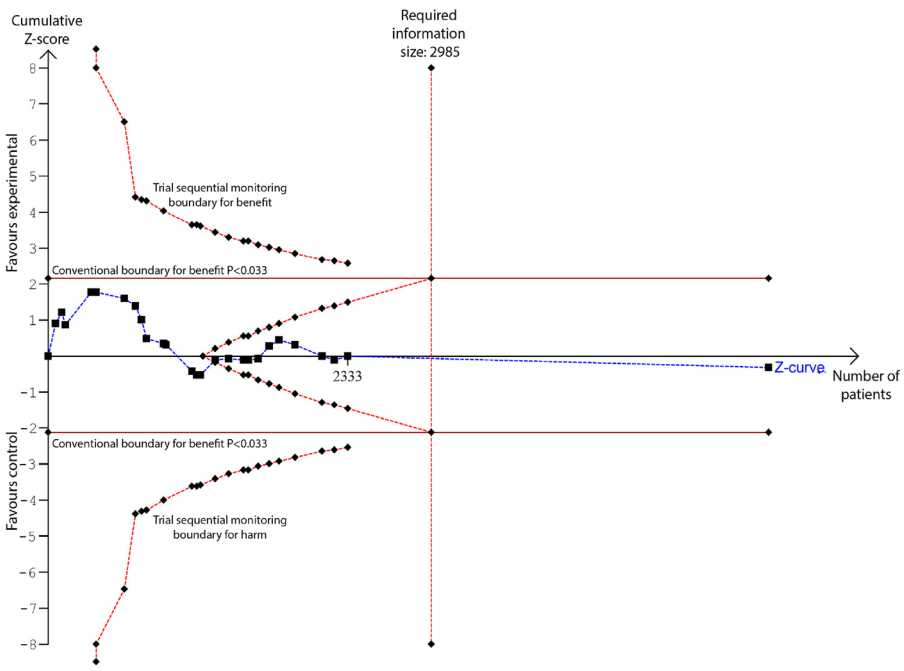


Table 1 Summary of findings

Proton pump inhibitors or histamin-2 receptor antagonists compared to placebo/no prophylaxis for stress ulcer prophylaxis in adult ICU patients										
Certainty assessment							**Summary of findings**			
No. of participants (studies) Follow-up	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Publication bias**	**Overall certainty of evidence**	**Study event rates (%)**		**Relative effect (95% naive CI)**	**Anticipated absolute effects**
							With control	**With PPI/H2RA**		**Risk with control**	**Risk difference with PPI/H2RA**
Mortality—low risk of bias trials											
3557 (3 RCTs)	Not serious	Not serious[a]	Not serious[b]	Not serious[c]	None	⨁⨁⨁⨁ High	537/1790 (30.0%)	557/1767 (31.5%)	RR 1.03 (0.94–1.14)	300 per 1000	9 more per 1000 (18 fewer to 42 more)
Mortality—all trials											
5656 (28 RCTs)	Serious[d]	Not serious[e]	Not serious[f]	Not serious[g]	None	⨁⨁⨁◯ Moderate	725/2714 (26.7%)	769/2942 (26.1%)	RR 1.01 (0.93–1.10)	267 per 1000	3 more per 1000 (19 fewer to 27 more)
GI bleeding—low risk of bias trials											
3596 (3 RCTs)	Not serious	Not serious[h]	Not serious[i]	Not serious[j]	None	⨁⨁⨁⨁ High	157/1797 (8.7%)	95/1799 (5.3%)	RR 0.60 (0.47–0.77)	87 per 1000	35 fewer per 1000 (46 fewer to 20 fewer)
GI bleeding—all trials											
6627 (39 RCTs)	Serious[k]	Serious[l]	Not serious[m]	Not serious[n]	None	⨁⨁◯◯ Low	395/3223 (12.3%)	218/3404 (6.4%)	RR 0.52 (0.45–0.61)	123 per 1000	59 fewer per 1000 (48 fewer to 67 fewer)
Serious adverse events (highest proportion)—low risk of bias trials											
3587 (3 RCTs)	Not serious	Not serious[o]	Very serious[p]	Not serious[q]	None	⨁⨁◯◯ Low	537/1790 (30.0%)	557/1797 (31.0%)	RR 1.03 (0.94–1.14)	300 per 1000	9 more per 1000 (18 fewer to 42 more)
Serious adverse events (highest proportion)—all trials											
6744 (42 RCTs)	Serious[r]	Serious[s]	Very serious[t]	Not serious[u]	Publication bias strongly suspected[v]	⨁◯◯◯ Very low	852/3252 (26.2%)	822/3492 (23.5%)	RR 0.92 (0.85–1.00)	262 per 1000	21 fewer per 1000 (39 fewer to 0 fewer)
Serious adverse events (cumulated)—low risk of bias trials											
3587 (3 RCTs)	Not serious	Serious[w]	Very serious[x]	Serious[y]	None	⨁◯◯◯ Very low	1130/1790 (63.1%)	1073/1797 (59.7%)	RR 1.04 (0.85–1.26)	631 per 1000	25 more per 1000 (95 fewer to 164 more)
Serious adverse events (cumulated)—all trials											
6748 (42 RCTs)	Serious[z]	Serious[aa]	Very serious[ab]	Not serious[ac]	None	⨁◯◯◯ Very low	1627/3254 (50.0%)	1521/3494 (43.5%)	RR 0.89 (0.85–0.93)	500 per 1000	55 fewer per 1000 (75 fewer to 35 fewer)
Health-related quality of life											
0 (0 RCTs)						–					

Table 1 (continued)

Proton pump inhibitors or histamin-2 receptor antagonists compared to placebo/no prophylaxis for stress ulcer prophylaxis in adult ICU patients										
Certainty assessment							Summary of findings			
No. of participants (studies) Follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall certainty of evidence	Study event rates (%)		Relative effect (95% naive CI)	Anticipated absolute effects
							With control	With PPI/H2RA		Risk with control	Risk difference with PPI/H2RA
Myocardial ischaemia											
3291 (1 RCT)	Not serious	Not serious	Not serious	Very serious[ad]	None	⨁⨁◯◯ Low	66/1647 (4.0%)	77/1644 (4.7%)	RR 1.17 (0.85–1.61)	40 per 1000	7 more per 1000 (6 fewer to 24 more)
Pneumonia—low risk of bias trials											
3596 (3 RCTs)	Not serious	Not serious[ae]	Not serious[af]	Serious[ag]	None	⨁⨁⨁◯ Moderate	273/1797 (15.2%)	278/1799 (15.5%)	RR 1.01 (0.87–1.18)	152 per 1000	2 more per 1000 (20 fewer to 27 more)
Pneumonia—all trials											
4951 (16 RCTs)	Serious[ah]	Not serious[ai]	Not serious[aj]	Serious[ak]	None	⨁⨁◯◯ Low	358/2401 (14.9%)	400/2550 (15.7%)	RR 1.07 (0.94–1.21)	149 per 1000	10 more per 1000 (9 fewer to 31 more)
Cl. difficile—low risk of bias trials											
3596 (3 RCTs)	not serious	Not serious[al]	Not serious[am]	very serious[an]	None	⨁⨁◯◯ LOW	26/1797 (1.4%)	22/1799 (1.2%)	RR 0.84 (0.48–1.47)	14 per 1000	2 fewer per 1000 (8 fewer to 7 more)
Cl. difficile—all trials											
3698 (4 RCTs)	Serious[ao]	Not serious[ap]	Not serious[aq]	Very serious[ar]	None	⨁◯◯◯ Very loW	29/1844 (1.6%)	23/1854 (1.2%)	RR 0.78 (0.46–1.34)	16 per 1000	3 fewer per 1000 (8 fewer to 5 more)

CI confidence interval, *cumulated* all reported serious adverse events cumulated in each trial, *highest proportion* highest proportion of reported serious adverse events in each trial, *RR* risk ratio

[a] $I^2 = 0\%$, $P = 0.72$, overlap of confidence intervals

[b] All trials assess PPI versus placebo. Duration of intervention differed slightly

[c] TSA-adjusted CI 0.69, 1.55 with the *Z*-curve reaching futility area for an RRR/RRI of 20%

[d] 25/28 trials had overall high risk of bias

[e] $I^2 = 0\%$, $P = 0.84$, overlap of confidence intervals

[f] 20 trials assessed H2RA and nine trials assessed PPI (no subgroup difference, test-of-interaction $P = 0.51$). Some 22 trials compared intervention to placebo and seven trials compared to no prophylaxis (no subgroup difference, test-of-interaction $P = 0.51$). Duration of intervention differed

[g] 95% CI CI 0.90.93, 1.10 with the *Z*-curve reaching required information size

[h] $I^2 = 0\%$, $P = 0.66$, overlap of confidence intervals

[i] All trials assessed PPI versus placebo. Treatment duration differed slightly

[j] TSA-adjusted CI 0.36, 1.00 with the *Z*-curve reaching the trial sequential monitoring boundary for benefit

[k] 36/39 had overall high risk of bias

Table 1 (continued)

[l] $I^2 = 43\%$, $P = 0.005$. Signs of heterogeneity in forest plot

[m] 11 trials assessed PPI and 28 trials assessed H2RA (no subgrup difference, $P = 0.38$). 28 trials compared intervention to placebo and 11 compared to no prophylaxis (no subgroup difference, $P = 0.59$). Treatment duration differed

[n] TSA-adjusted CI 0.31, 0.84 with the *Z*-curve crossing the trial sequential monitoring boundary for benefit

[o] $I^2 = 0\%$, $P = 0.72$, overlap of confidence intervals

[p] Method used in meta-analysis and under-reporting of serious adverse events in trials

[q] TSA-adjusted CI 0.69, 1.55 with the cumulative *Z*-curve reaching the futility area for an RRR of 20%

[r] 39/42 trials had overall high risk of bias

[s] $I^2 = 44\%$, $P = 0.002$. Signs of heterogeneity in forest plot

[t] Method used in meta-analysis and under-reporting of serious adverse events in trials

[u] TSA-adjusted CI 0.84, 1.02 with the cumulative *Z*-curve crossing the trial sequential monitoring boundary for benefit

[v] Funnel plot indicated asymmetry and Harbord's test did indicated this publication bias ($P = 0.019$)

[w] $I^2 = 53\%$, $P = 0.12$. Signs of heterogeneity in forest plot

[x] Method used in meta-analysis and under-reporting of serious adverse events in trials

[y] TSA-adjusted CI 0.64, 1.68 with the cumulative *Z*-curve not reaching the trial sequential monitoring boundary and not reaching the futility area

[z] 39/42 trials had overall high risk of bias

[aa] $I^2 = 59\%$, $P < 0.00001$. Signs of heterogeneity in forest plot

[ab] Method is an indirect method of estimating SAE

[ac] TSA-adjusted CI was 0.85, 0.94 with the cumulative *Z*-curve reaching the trial sequential monitoring boundary

[ad] Only one trial included, with wide CI around effect estimate. However, it included 3291 analysed patients

[ae] $I^2 = 0\%$, $P = 0.64$, overlap of confidence intervals

[af] All trials assess PPI versus placebo. Treatment duration differed slightly

[ag] According to the 95% CI and TSA-adjusted CI there are still a risk of 20% RRR/RRI

[ah] 13/16 were overall high risk of bias

[ai] $I^2 = 0\%$, $P = 0.50$, overlap of confidence intervals

[aj] Six trials assessed PPI and 10 assessed H2RA (no subgroup difference, $P = 0.18$). 13 trials assess placebo and 3 trials assessed no prophylaxis (no subgroup difference,$P = 0.16$). Duration of intervention differed

[ak] According to the 95% CI and TSA-adjusted CI there are still a risk of 20% RRR/RRI

[al] $I^2 = 0\%$, $P = 0.58$, overlap of confidence intervals

[am] All trials assessed PPI versus placebo. Minor differences in intervention period

[an] TSA was not possible due to too little information. Optimal information size criterion is therefore not met

[ao] 1/4 trials had overall high risk of bias

[ap] $I^2 = 0\%$, $P = 0.59$, overlap of confidence intervals

[aq] All trials assessed PPI versus placebo. Minor differences in intervention period

[ar] TSA was not possible due to too little information. Optimal information size criterion is therefore not met

(See figure on next page.)

Fig. 4 **a** Forest plot of gastrointestinal bleeding in trials with overall low risk of bias versus trials with overall high risk of bias. *Size of squares* for risk ratio reflects weight of trial in pooled analysis. *Horizontal bars* represent 95% confidence intervals. **b** Trial sequential analysis of all 39 trials regardless of risk of bias of the effect of proton pump inhibitors/histamine 2 receptor antagonists versus placebo/no prophylaxis on GI bleeding using a control event proportion of 12.26% (from the included trials), a diversity (D2) of 0%, an alpha of 3.3%, a power of 90%, and relative risk reduction of 20%. The relative risk was 0.52 with a TSA-adjusted CI 0.39, 0.68. As the cumulative *Z*-curve reached the trial sequential monitoring boundary for benefit there is evidence of at a 20% relative risk reduction in the risk of GI bleeding from proton pump inhibitors or histamine 2 receptor antagonists

Cl. difficile enteritis

A total of four trials with 3698 participants reported data on *Cl. difficile* enteritis, including the three trials with overall low risk of bias with 3596 participants.

The meta-analyses of trials with overall low risk of bias and trials regardless of risk of bias were both inconclusive (S37, ESM). TSA highlighted that less than 5% of the required information size had been reached. The certainty of evidence was low/very low due to very serious imprecision and risk of bias (Table 1).

Subgroup analyses of PPI versus H2RA and placebo versus no prophylaxis were not applicable. The sensitivity analyses of missing data and subgroup analyses were consistent with the primary analysis (S38 and S39, ESM).

Discussion

In this updated systematic review, we did not find a difference in mortality between adult ICU patients receiving PPI or H2RA versus placebo/no prophylaxis, and TSA highlighted that the required information size to detect a 20% (and even a 15%) relative difference in mortality had been reached, indicating firm evidence. Furthermore, we found a reduction in the occurrence of any GI bleeding and clinically important GI bleeding, and TSA highlighted that firm evidence for such a reduction in any GI bleeding had been reached; however, this was not the case for clinically important GI bleeding. The effects on the other outcomes, including SAEs, HRQoL, myocardial infarction, pneumonia, and *CI. difficile* enteritis, were inconclusive.

Strengths and limitations

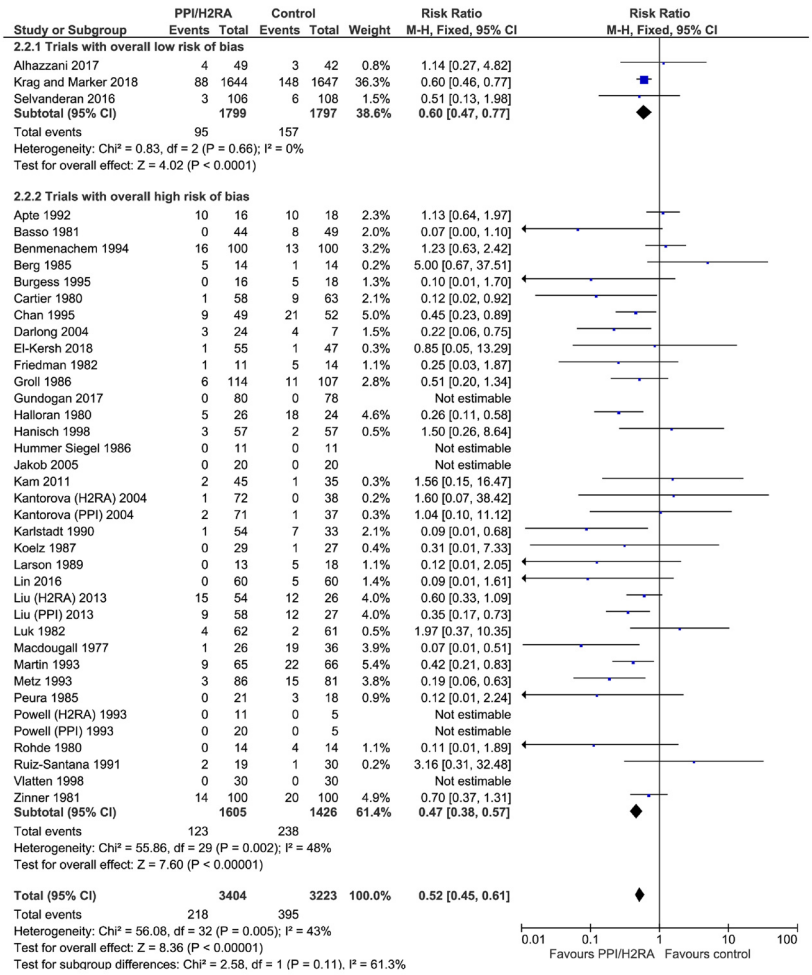
Strengths of this review include the systematic, transparent and robust methodology used, including the use of the Cochrane Handbook [20], the PRISMA statement [21], a prespecified protocol [19], an up-to-date comprehensive literature search, and the independent study selection, data extraction, and risk of bias assessment by two authors. Also, we used TSA to assess the overall risk of random error to increase the reliability of the results of the meta-analysis, and to identify the required information size. Finally, we assessed the certainty of evidence using GRADE.

Limitations of our review include a risk of clinical heterogeneity between trials. Furthermore, statistical heterogeneity was present in the analyses of GI bleeding and SAEs. To account for systematic errors and missing data in the included trials, we conducted subgroup analyses comparing trials of overall high risk of bias with trials of overall low risk of bias, and sensitivity analyses to account for missing data. We cannot exclude a biased effect estimate of the trials of overall high risk of bias; hence, the certainty of evidence for all trials irrespective of risk of bias was downgraded one level for risk of bias. We were unable to include the losses to follow-up from four trials ($n=81$) in the sensitivity analyses exploring uncertainty due to missing data, as the trial reports did not specify to which intervention group these patients belonged. The uncertainty due to loss to follow-up is therefore higher. None of the included trials reported detailed data on SAEs according to the ICH-GCP recommendation [25]; however, four trials reported zero SAEs in both groups, although mortality, clinically important GI bleeding and hospital-acquired pneumonia were reported [12, 52, 64]. Accordingly, SAEs are likely considerably underreported. To estimate the effect on SAEs actually reported in the included trials we conducted two post hoc analyses aiming to estimate the effect on the proportion of patients having one or more SAEs expected to lie between these two extremes. Analysing SAEs according to ICH-GCP may not be optimal in ICU patients who may experience numerous SAEs each day, making it difficult to register them all; thus, a composite outcome as defined by ICH-GCP may be inappropriate. Although we had two co-authors not involved in the SUP-ICU trial to assess the risk of bias in this trial, we acknowledge the potential for indirect conflicts of interests from review authors being involved in the SUP-ICU trial. Finally, limited data on SAEs, HRQoL, myocardial ischemia, pneumonia, and *CI. difficile* enteritis resulted in no firm evidence on the balance between the benefits and harms for these outcomes.

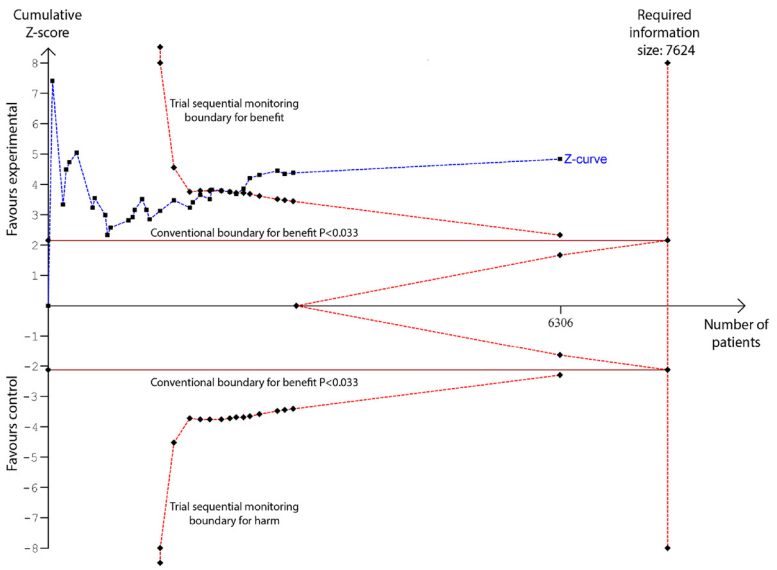
Our results in relation to previous systematic reviews

Previous systematic reviews have not observed a difference in mortality between PPI/H2RA and placebo/no prophylaxis [80–83], which our results, including TSA, confirm. Previous reviews have shown conflicting results regarding the effects of stress ulcer prophylaxis on any GI bleeding [80, 82, 83]. Our results show an absolute

a


PPI/H2RA
Control
Risk Ratio
Risk Ratio
Study or Subgroup Events Total Events Total Weight M-H, Fixed, 95% CI
M-H, Fixed, 95% CI
2.2.1 Trials with overall low risk of bias
Alhazzani 2017 4 49 3 42 0.8% 1.14 [0.27, 4.82]
Krag and Marker 2018 88 1644 148 1647 36.3% 0.60 [0.46, 0.77]
Selvanderan 2016 3 106 6 108 1.5% 0.51 [0.13, 1.98]
Subtotal (95% CI) 1799 1797 38.6% 0.60 [0.47, 0.77]
Total events 95 157
Heterogeneity: Chi² = 0.83, df = 2 (P = 0.66); I² = 0%
Test for overall effect: Z = 4.02 (P < 0.0001)
2.2.2 Trials with overall high risk of bias
Apte 1992 10 16 10 18 2.3% 1.13 [0.64, 1.97]
Basso 1981 0 44 8 49 2.0% 0.07 [0.00, 1.10]
Benmenachem 1994 16 100 13 100 3.2% 1.23 [0.63, 2.42]
Berg 1985 5 14 1 14 0.2% 5.00 [0.67, 37.51]
Burgess 1995 0 16 5 18 1.3% 0.10 [0.01, 1.70]
Cartier 1980 1 58 9 63 2.1% 0.12 [0.02, 0.92]
Chan 1995 9 49 21 52 5.0% 0.45 [0.23, 0.89]
Darlong 2004 3 24 4 7 1.5% 0.22 [0.06, 0.75]
El-Kersh 2018 1 55 1 47 0.3% 0.85 [0.05, 13.29]
Friedman 1982 1 11 5 14 1.1% 0.25 [0.03, 1.87]
Groll 1986 6 114 11 107 2.8% 0.51 [0.20, 1.34]
Gundogan 2017 0 80 0 78 Not estimable
Halloran 1980 5 26 18 24 4.6% 0.26 [0.11, 0.58]
Hanisch 1998 3 57 2 57 0.5% 1.50 [0.26, 8.64]
Hummer Siegel 1986 0 11 0 11 Not estimable
Jakob 2005 0 20 0 20 Not estimable
Kam 2011 2 45 1 35 0.3% 1.56 [0.15, 16.47]
Kantorova (H2RA) 2004 1 72 0 38 0.2% 1.60 [0.07, 38.42]
Kantorova (PPI) 2004 2 71 1 37 0.3% 1.04 [0.10, 11.12]
Karlstadt 1990 1 54 7 33 2.1% 0.09 [0.01, 0.68]
Koelz 1987 0 29 1 27 0.4% 0.31 [0.01, 7.33]
Larson 1989 0 13 5 18 1.1% 0.12 [0.01, 2.05]
Lin 2016 0 60 5 60 1.4% 0.09 [0.01, 1.61]
Liu (H2RA) 2013 15 54 12 26 4.0% 0.60 [0.33, 1.09]
Liu (PPI) 2013 9 58 12 27 4.0% 0.35 [0.17, 0.73]
Luk 1982 4 62 2 61 0.5% 1.97 [0.37, 10.35]
Macdougall 1977 1 26 19 36 3.9% 0.07 [0.01, 0.51]
Martin 1993 9 65 22 66 5.4% 0.42 [0.21, 0.83]
Metz 1993 3 86 15 81 3.8% 0.19 [0.06, 0.63]
Peura 1985 0 21 3 18 0.9% 0.12 [0.01, 2.24]
Powell (H2RA) 1993 0 11 0 5 Not estimable
Powell (PPI) 1993 0 20 0 5 Not estimable
Rohde 1980 0 14 4 14 1.1% 0.11 [0.01, 1.89]
Ruiz-Santana 1991 2 19 1 30 0.2% 3.16 [0.31, 32.48]
Vlatten 1998 0 30 0 30 Not estimable
Zinner 1981 14 100 20 100 4.9% 0.70 [0.37, 1.31]
Subtotal (95% CI) 1605 1426 61.4% 0.47 [0.38, 0.57]
Total events 123 238
Heterogeneity: Chi² = 55.86, df = 29 (P = 0.002); I² = 48%
Test for overall effect: Z = 7.60 (P < 0.00001)
Total (95% CI) 3404 3223 100.0% 0.52 [0.45, 0.61]
Total events 218 395
Heterogeneity: Chi² = 56.08, df = 32 (P = 0.005); I² = 43%
Test for overall effect: Z = 8.36 (P < 0.00001)
Test for subgroup differences: Chi² = 2.58, df = 1 (P = 0.11), I² = 61.3%
0.01 0.1 1 10 100
Favours PPI/H2RA Favours control

b


Cumulative Z-score
Required information size: 7624
Trial sequential monitoring boundary for benefit
Z-curve
Favours experimental
Conventional boundary for benefit P<0.033
6306
Number of patients
Conventional boundary for benefit P<0.033
Favours control
Trial sequential monitoring boundary for harm

difference in any GI bleeding of 3.4%, corresponding to a number needed to treat of 35 (CI from 46 fewer to 20 fewer) in trials with overall low risk of bias. Previous reviews have also shown inconsistent results in clinically important GI bleeding [81, 83]. In accordance with previous reviews, we did not observe a statistically significant difference in hospital-acquired pneumonia, indicating no firm evidence for benefit or harm [80–83]. A recently published systematic review did not report a difference in *CI. difficile* enteritis which is supported by our results [82]. SAEs, HRQoL, and myocardial ischemia have not been assessed in previous reviews.

Clinical implications and perspectives

Nowadays, GI bleeding, including clinically important GI bleeding, is an important but rare event in adult ICUs. Yet, stress ulcer prophylaxis is used in three out of four acutely admitted adult ICU patients [2], and recommendation on its use is conflicting [6, 7].

Our results indicate that, although we did not find an effect of stress ulcer prophylaxis on mortality, GI bleeding is reduced by almost 50% and clinically important bleeding a little less, which could be used as an argument for using PPI/H2RA as a prophylactic intervention in intensive care patients. Conversely, GI bleeding occurs in 12% of intensive care patients and clinically important GI bleeding in only 5% of the patients with placebo or no intervention. Furthermore, as mortality does not seem to be reduced using PPI/H2RA, it could be argued that the prophylactic use is unnecessary and that treatment with antacids should be reserved for patients developing active GI bleeding. Moreover, a pre-planned subgroup analysis in the recently published SUP-ICU trial suggested excess mortality among patients with a Simplified Acute Physiology Score II greater than 53 allocated to PPI compared with placebo, indicating that the most severely ill patients may be harmed from prophylactic PPI [3]. On the other hand, prophylactic PPI does not appear to substantially increase the number of SAEs, including nosocomial infections and myocardial ischemia. Accordingly, additional data on the importance of disease severity on the overall effects of stress ulcer prophylaxis are needed, along with data on long-term outcomes, HRQoL, and an economic analysis [84].

Conclusions

In this updated systematic review, we were able to refute a relative change of 20% of mortality when prophylactic PPI or H2RA were compared with placebo or no prophylaxis in adult ICU patients. GI bleeding was reduced with PPI or H2RA, but firm evidence for a reduction in clinically important GI bleeding was not found. The effects on SAEs, HRQoL, myocardial ischemia, pneumonia, and *CI. difficile* enteritis remain inconclusive.

Discrepancy between protocol and review

We used a power of 90%, and not 80% as reported in the protocol [19], as meta-analyses should use a higher (or the same) power as its included trials to be able to communicate the best available evidence.

We choose to report two post hoc analyses of the effect of PPI/H2RA on SAEs as none of the trials reported these according to the ICH-GCP criteria. Furthermore, we conducted two post hoc subgroup analyses according to dose of PPI and publication year. In addition, we have made a TSA anticipating a 15% RRR of mortality on the meta-analysis of new trials published after our first review [34].

Electronic supplementary material

The online version of this article (https://doi.org/10.1007/s00134-019-05526-z) contains supplementary material, which is available to authorized users.

Author details

[1] Copenhagen Trial Unit, Centre for Clinical Intervention Research, 2100 Copenhagen, Denmark. [2] Centre for Research in Intensive Care, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. [3] Department of Intensive Care, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. [4] Department of Cardiology, Holbaek Hospital, Holbaek, Denmark.

Acknowledgements

MB, SM, JCJ, AP and JW were supported by the public Innovation Fund Denmark (4108-00011B), which did not have any role in study design, data collection, data analysis, data interpretation, or writing of the report. No other sources of financial support were obtained for this review. The authors thank Sanam Safi and Kiran Kumar Katakam, who were not involved in any aspects of the SUP-ICU trial, for extracting data and evaluating risk of bias of this trial. We also wish to thank Maria Hernandez Sierra, Aleksandra Mazur, Ning Liang and Dezhao Kong for translating papers.

Compliance with ethical standards

Conflicts of interest

Marija Barbateskovic: PhD student at the Copenhagen Trial Unit and the Centre for Research in Intensive Care. Søren Marker: PhD student at the Department of Intensive Care at Rigshospitalet and the Centre for Research in Intensive Care. Coordinating investigator of the randomised clinical trial 'Stress Ulcer Prophylaxis in the Intensive Care Unit' (SUP-ICU). Anders Granholm: Coordinating investigator of the SUP-ICU trial. Carl Thomas Anthon: Coordinating investigator of the SUP-ICU trial. Mette Krag: Coordinating investigator of the SUP-ICU trial. Janus Christian Jakobsen: Director of Research, Chief Physician, Department of Cardiology, Holbæk Sygehus, Holbæk, Denmark. Anders Perner: Head of Research at the Department of Intensive Care at Rigshospitalet. The intensive care unit receives support for research from CSL Behring, Fresenius Kabi, Ferring Pharmaceuticals and the Novo Nordisk Foundation. Dr Perner is initiator of the SUP-ICU trial. Jørn Wetterslev: Member of the Copenhagen Trial Unit task force for developing Trial Sequential Analysis theory, manual and software which is presently free-ware at www.ctu.dk/tsa. Dr Wetterslev is member of the SUP-ICU trial steering group. Morten Hylander Møller: Sponsor and initiator of the SUP-ICU trial.

Ethical approval

An approval by an ethics committee was not applicable.

Publisher's Note

Received: 29 October 2018 Accepted: 7 January 2019
Published online: 24 January 2019

References

1. Marik PE, Vasu T, Hirani A, Pachinburavan M (2010) Stress ulcer prophylaxis in the new millennium: a systematic review and meta-analysis. Crit Care Med 38(11):2222–2228
2. Krag M, Perner A, Wetterslev J, Wise M, Borthwick M, Bendel S et al (2015) Prevalence and outcome of gastrointestinal bleeding and use of acid suppressants in acutely ill adult intensive care patients. Intensive Care Med 41(5):833–845
3. Krag M, Marker S, Perner A, Wetterslev J, Wise MP, Schefold JC et al (2018) Pantoprazole in patients at risk for gastrointestinal bleeding in the ICU. N Eng J Med. https://doi.org/10.1056/nejmoa1714919
4. Alhazzani W, Guyatt G, Alshahrani M, Deane AM, Marshall JC, Hall R et al (2017) Withholding pantoprazole for stress ulcer prophylaxis in critically ill patients: a pilot randomized clinical trial and meta-analysis. Crit Care Med 45(7):1121–1129
5. Selvanderan SP, Summers MJ, Finnis ME, Plummer MP, Ali Abdelhamid Y, Anderson MB et al (2016) Pantoprazole or placebo for stress ulcer prophylaxis (pop-up): randomized double-blind exploratory study. Crit Care Med 44(10):1842–1850
6. Rhodes A, Evans LE, Alhazzani W, Levy MM, Antonelli M, Ferrer R et al (2017) Surviving sepsis campaign: international guidelines for management of sepsis and septic shock: 2016. Intensive Care Med 43(3):304–377
7. Madsen KR, Lorentzen K, Clausen N, Oberg E, Kirkegaard PR, Maymann-Holler N et al (2014) Guideline for stress ulcer prophylaxis in the intensive care unit. Dan Med J 61(3):C4811
8. Krag M, Perner A, Wetterslev J, Moller MH (2013) Stress ulcer prophylaxis in the intensive care unit: is it indicated? A topical systematic review. Acta Anaesthesiol Scand 57(7):835–847
9. Marker S, Krag M, Moller MH (2017) What's new with stress ulcer prophylaxis in the ICU? Intensive Care Med 43(8):1132–1134
10. Alhazzani W, Alenezi F, Jaeschke RZ, Moayyedi P, Cook DJ (2013) Proton pump inhibitors versus histamine 2 receptor antagonists for stress ulcer prophylaxis in critically ill patients: a systematic review and meta-analysis. Crit Care Med 41(3):693–705
11. Lin PC, Chang CH, Hsu PI, Tseng PL, Huang YB (2010) The efficacy and safety of proton pump inhibitors vs histamine-2 receptor antagonists for stress ulcer bleeding prophylaxis among critical care patients: a meta-analysis. Crit Care Med 38(4):1197–1205
12. Kantorova I, Svoboda P, Scheer P, Doubek J, Rehorkova D, Bosakova H et al (2004) Stress ulcer prophylaxis in critically ill patients: a randomized controlled trial. Hepatogastroenterology 51(57):757–761
13. Pongprasobchai S, Kridkratoke S, Nopmaneejumruslers C (2009) Proton pump inhibitors for the prevention of stress-related mucosal disease in critically-ill patients: a meta-analysis. J Med Assoc Thai 92(5):632–637
14. Barkun AN, Bardou M, Pham CQ, Martel M (2012) Proton pump inhibitors vs. histamine 2 receptor antagonists for stress-related mucosal bleeding prophylaxis in critically ill patients: a meta analysis. Am J Gastroenterol 107(4):507–520
15. Alshamsi F, Belley-Cote E, Cook D, Almenawer SA, Alqahtani Z, Perri D et al (2016) Efficacy and safety of proton pump inhibitors for stress ulcer prophylaxis in critically ill patients: a systematic review and meta-analysis of randomized trials. Crit Care 20(1):120
16. Liu BL, Li B, Zhang X, Fei Z, Hu SJ, Lin W et al (2013) A randomized controlled study comparing omeprazole and cimetidine for the prophylaxis of stress-related upper gastrointestinal bleeding in patients with intracerebral hemorrhage. J Neurosurg 118(1):115–120
17. Lin CC, Hsu YL, Chung CS, Lee TH (2016) Stress ulcer prophylaxis in patients being weaned from the ventilator in a respiratory care center: a randomized control trial. J Formos Med Assoc 115(1):19–24
18. El-Kersh K, Jalil B, Mcclave SA, Cavallazzi R, Guardiola J, Guilkey K et al (2018) Enteral nutrition as stress ulcer prophylaxis in critically ill patients: a randomized controlled exploratory study. J Crit Care 43:108–113
19. Barbateskovic M, Marker S, Jakobsen JC, Krag M (2018) Stress ulcer prophylaxis in adult intensive care unit patients—a protocol for a systematic review. Acta Anaesthesiol Scand 62(6):744–755
20. Higgins JPT, Green S (eds) (2011) Cochrane handbook for systematic reviews of interventions version 5.1.0. Updated March 2011. The cochrane collaboration, 2011. Available fromwww.handbook.cochrane.org
21. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP et al (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 339:b2700
22. Keus F, Wetterslev J, Gluud C, Van Laarhoven CJ (2010) Evidence at a glance: error matrix approach for overviewing available evidence. BMC Med Res Methodol 10:90
23. Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C (2014) Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. BMC Med Res Methodol 14:120
24. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S et al (2004) Grading quality of evidence and strength of recommendations. BMJ 328(7454):1490
25. International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use (1997). ICH harmonised tripartite guideline. Guideline for good clinical practice. Updated July 2002
26. Savovic J, Turner RM, Mawdsley D, Jones HE, Beynon R, Higgins JPT et al (2018) Association between risk-of-bias assessments and results of randomized trials in cochrane reviews: the ROBES meta-epidemiologic study. Am J Epidemiol 187(5):1113–1122
27. Harbord RM, Egger M, Sterne JA (2006) A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. Stat Med 25(20):3443–3457
28. Thorlund K EJ, Wetterslev J, Brok J, Imberger G, Gluud C. User manual for trial sequential analysis (TSA). Www.Ctu.Dk/Tsa/Files/Tsa_Manual.Pdf. Accessed 15 Oct 2018
29. Wetterslev J, Jakobsen JC, Gluud C (2017) Trial sequential analysis in systematic reviews with meta-analysis. BMC Med Res Methodol 17(1):39
30. Revman (2014) Review manager (Revman) (computer program), version 5.3. The Nordic Cochrane Center, the Cochrane Collaboration, Copenhagen. https://Community.Cochrane.Org/Help/Tools-and-Software/Revman-5/Revman-5-Download. Accessed 15 Oct 2018
31. Wetterslev J, Thorlund K, Brok J, Gluud C (2009) Estimating required information size by quantifying diversity in random-effects model meta-analyses. BMC Med Res Methodol 9:86
32. Mantel N, Haenszel W (1959) Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst 22(4):719–748
33. Demets DL (1987) Methods for combining randomized clinical trials: strengths and limitations. Stat Med 6(3):341–350
34. Dersimonian R, Laird N (1986) Meta-analysis in clinical trials. Control Clin Trials 7(3):177–188
35. Deeks JJ, Higgins JPT (2010) Statistical algorithms in review manager 5. 2010: RevMan 5.3
36. Brok J, Thorlund K, Gluud C, Wetterslev J (2008) Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses. J Clin Epidemiol 61(8):763–769
37. Brok J, Thorlund K, Wetterslev J, Gluud C (2009) Apparently conclusive meta-analyses may be inconclusive—trial sequential analysis adjustment of random error risk due to repetitive testing of accumulating data in apparently conclusive neonatal meta-analyses. Int J Epidemiol 38(1):287–298
38. Higgins JP, Whitehead A, Simmonds M (2011) Sequential methods for random-effects meta-analysis. Stat Med 30(9):903–921
39. Imberger G, Gluud C, Boylan J, Wetterslev J (2015) Systematic reviews of anesthesiologic interventions reported as statistically significant: problems with power, precision, and type 1 error protection. Anesth Analg 121(6):1611–1622
40. Mascha EJ (2015) Alpha, beta, meta: guidelines for assessing power and type i error in meta-analyses. Anesth Analg 121(6):1430–1433

41. Pogue JM, Yusuf S (1997) Cumulating evidence from randomized trials: utilizing sequential monitoring boundaries for cumulative meta-analysis. Control Clin Trials 18(6):580–593
42. Terkawi AS, Mavridis D, Flood P, Wetterslev J, Terkawi RS, Bin Abdulhak AA et al (2016) Does ondansetron modify sympathectomy due to subarachnoid anesthesia?: Meta-analysis, meta-regression, and trial sequential analysis. Anesthesiology 124(4):846–869
43. Thorlund K, Devereaux PJ, Wetterslev J, Guyatt G, Ioannidis JP, Thabane L et al (2009) Can trial sequential monitoring boundaries reduce spurious inferences from meta-analyses? Int J Epidemiol 38(1):276–286
44. Wetterslev J, Thorlund K, Brok J, Gluud C (2008) Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. J Clin Epidemiol 61(1):64–75
45. Thorlund K, Imberger G, Johnston BC, Walsh M, Awad T, Thabane L et al (2012) Evolution of heterogeneity (I2) estimates and their 95% confidence intervals in large meta-analyses. PLoS ONE 7(7):e39471
46. Apte NM, Karnad DR, Medhekar TP, Tilve GH, Morye S, Bhave GG (1992) Gastric colonization and pneumonia in intubated critically ill patients receiving stress ulcer prophylaxis: a randomized controlled trial. Crit Care Med 20(5):590–593
47. Basso N, Bagarani M, Materia A, Fiorani S, Lunardi P, Speranza V (1981) Cimetidine and antacid prophylaxis of acute upper gastrointestinal bleeding in high risk patients: controlled, randomized trial. Am J Surg 141(3):339–341
48. Ben-Menachem T, Fogel R, Patel RV, Touchette M, Zarowitz BJ, Hadzijahic N et al (1994) Prophylaxis for stress-related gastric hemorrhage in the medical intensive care unit: a randomized, controlled, single-blind study. Ann Intern Med 121(8):568–575
49. Van Den Berg B, Van Blankenstein M (1985) Prevention of stress-induced upper gastrointestinal bleeding by cimetidine in patients on assisted ventilation. Digestion 31(1):1–8
50. Burgess P, Larson GM, Davidson P, Brown J, Metz CA (1995) Effect of ranitidine on intragastric ph and stress-related upper gastrointestinal bleeding in patients with severe head injury. Dig Dis Sci 40(3):645–650
51. Cartier F, Gauthier-Lafaye P, Lareng L, Mottin J, Cara M, Passelecq J et al (1980) Cimetideine in patients at risk of stress ulcers: a multi-centre controlled trial. Intensive Care Med 6:54
52. Chan KH, Lai EC, Tuen H, Ngan JH, Mok F, Fan YW et al (1995) Prospective double-blind placebo-controlled randomized trial on the use of ranitidine in preventing postoperative gastroduodenal complications in high-risk neurosurgical patients. J Neurosurg 82(3):413–417
53. Darlong V, Jayalakhsmi TS, Kaul HL, Tandon R (2003) Stress ulcer prophylaxis in patients on ventilator. Trop Gastroenterol 24(3):124–128
54. Domingues SHS, Stoeber GH, Stoeber AC (1985) Ranitidina Injetável Em Pacientes De Alto Risco. Folha Med 91(3):225–228
55. Friedman CJ, Oblinger MJ, Suratt PM, Bowers J, Goldberg SK, Sperling MH et al (1982) Prophylaxis of upper gastrointestinal hemorrhage in patients requiring mechanical ventilation. Crit Care Med 10(5):316–319
56. Groll A, Simon JB, Wigle RD, Taguchi K, Todd RJ, Depew WT (1986) Cimetidine prophylaxis for gastrointestinal bleeding in an intensive care unit. Gut 27(2):135–140
57. Gundogan K, Karakoc E, Teke T, Zerman A, Coruh A, Sungur M (2017) Effects of enteral nutrition on stress ulcer hemorrhage in critically ill patients: multicenter randomized controlled trial. Intensive Care Med Exp 5(2):44
58. Gursoy O, Memis D, Sut N (2008) Effect of proton pump inhibitors on gastric juice volume, gastric ph and gastric intramucosal pH in critically ill patients: a randomized, double-blind placebo-controlled study. Clin Drug Investig 28(12):777–782
59. Halloran LG, Zfass AM, Gayle WE, Wheeler CB, Miller JD (1980) Prevention of acute gastrointestinal complications after severe head injury: a controlled trial of cimetidine prophylaxis. Am J Surg 139(1):44–48
60. Hanisch EW, Encke A, Naujoks F, Windolf J (1998) A randomized, double-blind trial for stress ulcer prophylaxis shows no evidence of increased pneumonia. Am J Surg 176(5):453–457
61. Hummer-Sigiel M, Jacquier A, Girard A, Garric J, Laxenaire MC, Mandorla JY (1986) Ranitidine Pour La Prophylaxie De L'ulcére De Stress Chexz Les Traumatisés Crâniens Graves. Ann Med Nancy l'Est 25:101–103
62. Jakob SM, Parviainen I, Ruokonen E, Uusaro A, Takala J (2005) Lack of effect of ranitidine on gastric luminal Ph and mucosal PCO2 during the first day in the ICU. Acta Anaesthesiol Scand 49(3):390–396
63. Kam J, Modi C, Doraiswamy V, Abdul-Jawad S, Dixit D, Spira T et al (2011) Role of gastrointestinal ulcer prophylaxis in critically ill patients. Am J Gastroenterol 106(suppl. 2):s420
64. Karlstadt RG, Iberti TJ, Silverstein J, Lindenberg L, Bright-Asare P, Rockhold F et al (1990) Comparison of cimetidine and placebo for the prophylaxis of upper gastrointestinal bleeding due to stress-related gastric mucosal damage in the intensive care unit. J Intensive Care Med 5:26–32
65. Koelz HR, Aeberhard P, Hassler H, Kunz H, Wagner HE, Roth F et al (1987) Prophylactic treatment of acute gastroduodenal stress ulceration: low-dose antacid treatment without and with additional ranitidine. Scand J Gastroenterol 22(9):1147–1152
66. Larson GM, Davidson P, Brown J, Wilson T, Bishop A (1989) Comparison of ranitidine versus placebo on 24-hour gastric Ph and upper gastrointestinal (UGI) bleeding in head injury patients. Abstr Am J Gastroenterol 84:1165
67. Luk GD, Summer WR, Messersmith JF (1982) Cimetidine and antacid in prophylaxis of acute gastrointestinal bleeding: a randomized, double-blind, controlled study. Gastroenterology 82:1121
68. Macdougall BR, Bailey RJ, Williams R (1977) H2-receptor antagonists and antacids in the prevention of acute gastrointestinal haemorrhage in fulminant hepatic failure: two controlled trials. Lancet 1(8012):617–619
69. Martin LF, Booth FV, Karlstadt RG, Silverstein JH, Jacobs DM, Hampsey J et al (1993) Continuous intravenous cimetidine decreases stress-related upper gastrointestinal hemorrhage without promoting pneumonia. Crit Care Med 21(1):19–30
70. Metz CA, Livingston DH, Smith JS, Larson GM, Wilson TH (1993) Impact of multiple risk factors and ranitidine prophylaxis on the development of stress-related upper gastrointestinal bleeding: a prospective, multicenter, double-blind, randomized trial. The Ranitidine Head Injury Study Group. Crit Care Med 21(12):1844–1849
71. Nielsen HJ, Witt K, Moesgaard F, Kehlet H (1989) Ranitidine for improvement of delayed hypersensitivity response in patients with sepsis. Acta Chir Scand 155(9):445–449
72. Peura DA, Johnson LF (1985) Cimetidine for prevention and treatment of gastroduodenal mucosal lesions in patients in an intensive care unit. Ann Intern Med 103(2):173–177
73. Powell H, Morgan M, Li SK, Baron JH (1993) Inhibition of gastric acid secretion in the intensive care unit after coronary artery bypass graft. Theor Surg 8:125–130
74. Rigaud D, Accary JP, Chastre J, Mignon M, Laigneau JP, Reinberg A et al (1988) Persistence of circadian rhythms in gastric acid, gastrin, and pancreatic polypeptide secretions despite loss of cortisol and body temperature rhythms in man under stress. Gastroenterol Clin Biol 12(1):12–18
75. Rohde H, Lorenz W, Fischer M (1980) Eine Randomisierte Klinische Studie Zur Stressulkusprophylaxe Mit Cimetidin Beim Schweren Polytrauma. Z Gastroenterol 18(6):328–329
76. Ruiz-Santana S, Ortiz E, Gonzalez B, Bolanos J, Ruiz-Santana AJ, Manzano JL (1991) Stress-induced gastroduodenal lesions and total parenteral nutrition in critically ill patients: frequency, complications, and the value of prophylactic treatment: a prospective, randomized study. Crit Care Med 19(7):887–891
77. Spapen H, Diltoer M, Nguyen DN, Ingels G, Ramet J, Huyghens L (1995) One week treatment with cimetidine does not attenuate the cortisol response to a short corticotropin test in stable intensive care patients: a prospective, randomized, and controlled study. Acta Anaesthesiol Belg 46(3–4):133–140
78. Vlatten A, Wiedeck H, Reinelt H, Stanescu A, Georgieff M (1998) Stressulkus-Prophylaxe Bei Hoch-Risiko-Intensivpatienten. Vergleich Von Omeprazol, Pirenzepin Und Plazebo. Wien Klin Wochenschr Suppl 110(suppl. 1):38
79. Zinner MJ, Zuidema GD, Smith P, Mignosa M (1981) The prevention of upper gastrointestinal tract bleeding in patients in an intensive care unit. Surg Gynecol Obstet 153(2):214–220
80. Krag M, Perner A, Wetterslev J, Wise MP, Hylander Moller M (2014) Stress ulcer prophylaxis versus placebo or no prophylaxis in critically ill

patients: a systematic review of randomised clinical trials with meta-analysis and trial sequential analysis. Intensive Care Med 40(1):11–22
81. Toews I, George AT, Peter JV, Kirubakaran R, Fontes LES, Ezekiel JPB et al (2018) Interventions for preventing upper gastrointestinal bleeding in people admitted to intensive care units. Cochrane Database Syst Rev 6:Cd008687
82. Huang HB, Jiang W, Wang CY, Qin HY, Du B (2018) Stress ulcer prophylaxis in intensive care unit patients receiving enteral nutrition: a systematic review and meta-analysis. Crit Care 22(1):20
83. Alhazzani W, Alshamsi F, Belley-Cote E, Heels-Ansdell D, Brignardello-Petersen R, Alquraini M et al (2018) Efficacy and safety of stress ulcer prophylaxis in critically ill patients: a network meta-analysis of randomized trials. Intensive Care Med 44(1):1–11
84. Macias WL, Nelson DR, Williams M, Garg R, Janes J, Sashegyi A (2005) Lack of evidence for qualitative treatment by disease severity interactions in clinical studies of severe sepsis. Crit Care 9(6):R607–R622

PAPER IV

Cochrane Library

Cochrane Database of Systematic Reviews

Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the intensive care unit (Review)

Barbateskovic M, Schjørring OL, Russo Krauss S, Jakobsen JC, Meyhoff CS, Dahl RM, Rasmussen BS, Perner A, Wetterslev J

Barbateskovic M, Schjørring OL, Russo Krauss S, Jakobsen JC, Meyhoff CS, Dahl RM, Rasmussen BS, Perner A, Wetterslev J.
Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the intensive care unit.
Cochrane Database of Systematic Reviews 2019, Issue 11. Art. No.: CD012631.
DOI: 10.1002/14651858.CD012631.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the intensive care unit

Marija Barbateskovic[1,2], Olav L Schjørring[2,3], Sara Russo Krauss[4], Janus C Jakobsen[2,5,6,7], Christian S Meyhoff[8], Rikke M Dahl[9], Bodil S Rasmussen[2,3], Anders Perner[2,10], Jørn Wetterslev[1,2]

[1]Copenhagen Trial Unit, Centre for Clinical Intervention Research, Department 7812, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. [2]Centre for Research in Intensive Care, Department 7831, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. [3]Department of Anaesthesia and Intensive Care, Aalborg University Hospital, Aalborg, Denmark. [4]Copenhagen Trial Unit, Centre for Clinical Intervention Research, Copenhagen, Denmark. [5]Cochrane Hepato-Biliary Group, Copenhagen Trial Unit, Centre for Clinical Intervention Research, Department 7812, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. [6]Department of Cardiology, Holbaek Hospital, Holbaek, Denmark. [7]Copenhagen Trial Unit, Centre for Clinical Intervention Research, Department 7812, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. [8]Department of Anaesthesia and Intensive Care, Bispebjerg and Frederiksberg Hospital, University of Copenhagen, Copenhagen, Denmark. [9]Department of Anaesthesiology, Herlev Hospital, University of Copenhagen, Herlev, Denmark. [10]Department of Intensive Care, Righospitalet, Copenhagen University Hospital, Copenhagen, Denmark

Contact address: Marija Barbateskovic, Copenhagen Trial Unit, Centre for Clinical Intervention Research, Department 7812, Rigshospitalet, Copenhagen University Hospital, Blegdamsvej 9, Copenhagen, DK-2100, Denmark. marija.barbateskovic@ctu.dk.

Editorial group: Cochrane Emergency and Critical Care Group
Publication status and date: Edited (no change to conclusions), published in Issue 11, 2019.

Citation: Barbateskovic M, Schjørring OL, Russo Krauss S, Jakobsen JC, Meyhoff CS, Dahl RM, Rasmussen BS, Perner A, Wetterslev J. Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the intensive care unit. *Cochrane Database of Systematic Reviews* 2019, Issue 11. Art. No.: CD012631. DOI: 10.1002/14651858.CD012631.pub2.

ABSTRACT

Background

The mainstay treatment for hypoxaemia is oxygen therapy, which is given to the vast majority of adults admitted to the intensive care unit (ICU). The practice of oxygen administration has been liberal, which may result in hyperoxaemia. Some studies have indicated an association between hyperoxaemia and mortality, whilst other studies have not. The ideal target for supplemental oxygen for adults admitted to the ICU is uncertain. Despite a lack of robust evidence of effectiveness, oxygen administration is widely recommended in international clinical practice guidelines. The potential benefit of supplemental oxygen must be weighed against the potentially harmful effects of hyperoxaemia.

Objectives

To assess the benefits and harms of higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the ICU.

Search methods

We identified trials through electronic searches of CENTRAL, MEDLINE, Embase, Science Citation Index Expanded, BIOSIS Previews, CINAHL, and LILACS. We searched for ongoing or unpublished trials in clinical trials registers. We also scanned the reference lists of included studies. We ran the searches in December 2018.

Selection criteria

We included randomized controlled trials (RCTs) that compared higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the ICU. We included trials irrespective of publication type, publication status, and language.

We included trials with a difference between the intervention and control groups of a minimum 1 kPa in partial pressure of arterial oxygen (PaO_2), minimum 10% in fraction of inspired oxygen (FiO_2), or minimum 2% in arterial oxygen saturation of haemoglobin/non-invasive peripheral oxygen saturation (SaO_2/SpO_2).

We excluded trials randomizing participants to hypoxaemia (FiO_2 below 0.21, SaO_2/SpO_2 below 80%, and PaO_2 below 6 kPa) and to hyperbaric oxygen.

Data collection and analysis

Three review authors independently, and in pairs, screened the references retrieved in the literature searches and extracted data. Our primary outcomes were all-cause mortality, the proportion of participants with one or more serious adverse events, and quality of life. None of the trials reported the proportion of participants with one or more serious adverse events according to the International Conference on Harmonisation Good Clinical Practice (ICH-GCP) criteria. Nonetheless, most trials reported several serious adverse events. We therefore included an analysis of the effect of higher versus lower fraction of inspired oxygen, or targets using the highest reported proportion of participants with a serious adverse event in each trial. Our secondary outcomes were lung injury, acute myocardial infarction, stroke, and sepsis.

None of the trials reported on lung injury as a composite outcome, however some trials reported on acute respiratory distress syndrome (ARDS) and pneumonia. We included an analysis of the effect of higher versus lower fraction of inspired oxygen or targets using the highest reported proportion of participants with ARDS or pneumonia in each trial. To assess the risk of systematic errors, we evaluated the risk of bias of the included trials. We used GRADE to assess the overall certainty of the evidence.

Main results

We included 10 RCTs (1458 participants), seven of which reported relevant outcomes for this review (1285 participants). All included trials had an overall high risk of bias, whilst two trials had a low risk of bias for all domains except blinding of participants and personnel.

Meta-analysis indicated harm from higher fraction of inspired oxygen or targets as compared with lower fraction or targets of arterial oxygenation regarding mortality at the time point closest to three months (risk ratio (RR) 1.18, 95% confidence interval (CI) 1.01 to 1.37; I^2 = 0%; 4 trials; 1135 participants; very low-certainty evidence). Meta-analysis indicated harm from higher fraction of inspired oxygen or targets as compared with lower fraction or targets of arterial oxygenation regarding serious adverse events at the time point closest to three months (estimated highest proportion of specific serious adverse events in each trial RR 1.13, 95% CI 1.04 to 1.23; I^2 = 0%; 1234 participants; 6 trials; very low-certainty evidence). These findings should be interpreted with caution given that they are based on very low-certainty evidence.

None of the included trials reported any data on quality of life at any time point.

Meta-analysis indicated no evidence of a difference between higher fraction of inspired oxygen or targets as compared with lower fraction or targets of arterial oxygenation on lung injury at the time point closest to three months (estimated highest reported proportion of lung injury RR 1.03, 95% CI 0.78 to 1.36; I^2 = 0%; 1167 participants; 5 trials; very low-certainty evidence).

None of the included trials reported any data on acute myocardial infarction or stroke, and only one trial reported data on the effects on sepsis.

Authors' conclusions

We are very uncertain about the effects of higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the ICU on all-cause mortality, serious adverse events, and lung injuries at the time point closest to three months due to very low-certainty evidence. Our results indicate that oxygen supplementation with higher versus lower fractions or oxygenation targets may increase mortality. None of the trials reported the proportion of participants with one or more serious adverse events according to the ICH-GCP criteria, however we found that the trials reported an increase in the number of serious adverse events with higher fractions or oxygenation targets. The effects on quality of life, acute myocardial infarction, stroke, and sepsis are unknown due to insufficient data.

PLAIN LANGUAGE SUMMARY

Supplemental oxygen for adults admitted to the intensive care unit

Review question

We set out to assess whether more supplemental oxygen is better than less supplemental oxygen for adults admitted to the intensive care unit (ICU).

Background

Adults admitted to the ICU are critically ill and are at high risk of dying. Oxygen supplementation, or therapy, is given to most adults admitted to ICU, and many are mechanically ventilated. Severe illness can result in a lack of oxygen in the blood, known as hypoxaemia, which puts patients at risk of low tissue levels of oxygen (hypoxia) and organ failure. The use of sedatives and strong pain relief medications can also depress breathing and therefore oxygen levels.

The practice of supplemental oxygen administration has been liberal, possibly resulting in too much oxygen, known as hyperoxia. Despite a lack of robust evidence of effectiveness, supplemental oxygen administration has been widely recommended in international clinical practice guidelines. However, a new guideline recommends against high oxygen levels as some, but not all, clinical studies have indicated a link between hyperoxaemia and an increased risk of dying. The potential benefit of supplemental oxygen must be weighed against the potentially harmful effects of hyperoxaemia.

Study characteristics

We identified 10 randomized controlled trials (studies where participants are randomly allocated to either an experimental or a control group) involving 1458 participants up to December 2018. Seven of the trials (1285 participants) provided findings on the number of deaths, serious adverse events, and lung injuries in the three months following oxygen therapy in the ICU. Lung injury was measured according to participants developing acute respiratory distress syndrome or pneumonia. Five trials included adults admitted to an ICU caring for patients with a range of serious health conditions and one to a surgical ICU. Two trials involved adults with traumatic brain injury; one trial adults after cardiac arrest and resuscitation; and one trial adults with stroke. All participants in six trials received invasive mechanical ventilation directly through a tube into the main airway. In one trial some of the participants were on mechanical ventilation, whilst others received non-invasive oxygen administration. Three trials involved adults receiving non-invasive oxygen. All trials compared more with less oxygen, however using very different levels of oxygen supplementation. Oxygen therapy was given for timeframes ranging from one hour to the length of hospital admission.

Key results

We are uncertain about the effects of higher levels of oxygen as our findings are based on very low-certainty evidence. We found no evidence for a beneficial effect of higher compared with lower supplemental oxygen levels for adults admitted to ICU. Higher levels of oxygen may have increased the risk of death (4 trials; 1135 participants) and serious adverse events (6 trials; 1234 participants). There was no evidence of a difference in lung injuries with the use of higher supplemental oxygen compared with lower supplemental oxygen, but the evidence is very uncertain (5 trials; 1167 participants). None of the included trials reported on quality of life at any time point, acute myocardial infarction, and stroke. Only one trial reported on sepsis.

Certainty of the evidence

The numbers of participants enrolled in the trials were too small to permit a definitive judgement about the review findings. The trials varied in the types of illness of the participants, their associated clinical care, disease severity, the targets for how much oxygen was given, and for how long. Two of the trials had a low risk of bias other than for lack of blinding of participants and personnel. Overall all included trials had a high risk of bias.

Summary of findings for the main comparison. Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the ICU

Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the ICU

Patient or population: adults admitted to the ICU
Setting: trials were conducted in ICU departments in Europe (n = 5); Iran (n = 2); New Zealand (n = 1); Australia, New Zealand, France (n = 1); and Japan (n = 1)
Intervention: higher fraction of inspired oxygen or targets of arterial oxygenation
Comparison: lower fraction of inspired oxygen or targets of arterial oxygenation

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with lower FiO_2 or targets of arterial oxygenation	Risk with higher FiO_2 or targets of arterial oxygenation				
All-cause mortality follow-up: range 1 month to 3 months	Study population		RR 1.18 (1.01 to 1.37)	1135 (4 RCTs)	⊕⊖⊖⊖ Very low[1]	-
	331 per 1000	391 per 1000 (334 to 453)				
Proportion of participants with 1 or more serious adverse events according to International Conference on Harmonisation Good Clinical Practice (ICH-GCP) follow-up: range 3 to 90 days	Study population		RR 1.13 (1.04 to 1.23)	1234 (6 RCTs)	⊕⊖⊖⊖ Very low[2]	Reported results are derived by taking the highest proportion reported in each trial which addresses the lowest possible proportion of participants with 1 or more serious adverse events. The following outcomes and numbers of trials and participants have been included: mortality: 3 trials, 701 participants; pneumonia: 1 trial, 65 participants; proportion of participants with 1 or more serious adverse events: 1 trial, 434 participants; mechanical ventilation (reported as a poor outcome): 1 trial, 34 participants. Meta-analysis from the analysis cumulating all reported serious adverse events which ad-
	430 per 1000	486 per 1000 (447 to 529)				

						dress the highest possible reported proportion of participants with 1 or more serious adverse events showed RR 1.08, 95% CI 0.99 to 1.18.
Quality of life (any valid scale such as the 36-item Short Form Health Survey (SF-36))	Study population		Not estimable	(0 studies)	-	No studies reported this outcome.
	-	-				
Lung injury diagnosed after randomization (composite outcome) follow-up: range 4 to 23 days	Study population		RR 1.03 (0.78 to 1.36)	1167 (5 RCTs)	⊕⊖⊖⊖ Very low[3]	Reported results are derived by taking the highest proportion reported in each trial which addresses the lowest possible proportion of participants with 1 or more lung injuries. The following outcomes and numbers of trials and participants have been included: ARDS: 2 trials, 223 participants; pneumonia: 3 trials, 944 participants. Meta-analysis from the analysis cumulating all reported lung injuries which address the highest possible reported proportion of participants with 1 or more lung injuries showed RR 0.99, 95% CI 0.75 to 1.30.
	128 per 1000	132 per 1000 (100 to 174)				
Acute myocardial infarction diagnosed after randomization	Study population		Not estimable	(0 studies)	-	No studies reported this outcome.
	-	-				
Stroke diagnosed after randomization	Study population		Not estimable	(0 studies)	-	No studies reported this outcome.
	-	-				
Severe sepsis diagnosed after randomization follow-up: 3 days	Study population		RR 1.87 (0.93 to 3.87)	445 (1 study)	⊕⊖⊖⊖ Very low[4]	Meta-analysis was not conducted, as only 1 trial reported on sepsis.
	50 per 1000	94 per 1000 (46 to 189)				

*__The risk in the intervention (higher) group__ (and its 95% CI) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI). __The risk in the control (lower) group__ (and its 95% CI) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

ARDS: acute respiratory distress syndrome; **CI:** confidence interval; **FiO_2:** fraction of inspired oxygen; **ICU:** intensive care unit; **RCT:** randomised controlled trial; **RR:** risk ratio

GRADE Working Group grades of evidence

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

---

[1]Downgraded three levels: one level because of risk of bias, as only one trial was overall low risk of bias except for blinding of participants and personnel (performance bias); one level because of differences in inspiratory oxygen fraction and target of arterial oxygenation in the experimental and control groups between trials; and one level because the optimal information size (OIS) was not reached. Required information size (RIS) is 2623 participants. RIS = OIS when $I^2 = 0$ and alpha is adjusted for multiple outcomes.
[2]Downgraded three levels: one level because of risk of bias, as only one trial was overall low risk of bias except for blinding of participants and personnel (performance bias); one level because of differences in inspiratory oxygen fraction and target of arterial oxygenation in the experimental and control groups between trials; and one level because the optimal information size (OIS) was not reached. Required information size (RIS) is 1577 participants. RIS = OIS when $I^2 = 0$ and alpha is adjusted for multiple outcomes.
[3]Downgraded three levels: one level because of risk of bias, as only one trial was overall low risk of bias except for blinding of participants and personnel (performance bias); one level because of differences in inspiratory oxygen fraction and target of arterial oxygenation in the experimental and control groups between trials; and one level because the optimal information size (OIS) was not reached. Required information size (RIS) is 7656 participants. RIS = OIS when $I^2 = 0$ and alpha is adjusted for multiple outcomes.
[4]Downgraded three levels: one level because of risk of bias; one level because we cannot reject inconsistency due to the inclusion of only one trial; and one level because the optimal information size was not reached.

BACKGROUND

Description of the condition

Hypoxaemia refers to lack of oxygen in the blood and is usually defined in terms of partial pressure of arterial oxygen (PaO_2) or arterial oxygen saturation of haemoglobin (SaO_2) (O'Driscoll 2017). Additionally, the non-invasive peripheral oxygen saturation (SpO_2) measured by pulse oximetry is routinely used. Hypoxaemia refers directly to the levels of oxygen in the blood, whilst the term hypoxia is defined as the lack of oxygen at a cellular level, for example tissues, organs, alveoli, or the body as a whole (O'Driscoll 2017).

In healthy individuals, the normal range for PaO_2 at sea level is 80 mmHg to 100 mmHg (Kratz 2004), with a general decrease with age (Crapo 1999). There is no clear definition of hypoxaemia; the most widely used definitions are a PaO_2 below 60 mmHg or a SaO_2 below 90% (O'Driscoll 2017). However, oxygenation targets below the normal range, and even defined as hypoxaemic, are recommended in adults who are mechanically ventilated with acute respiratory distress syndrome (ARDS) in the intensive care unit (ICU) targeting PaO_2 of 55 mmHg to 80 mmHg or SpO_2 of 88% to 95% (ARDS Network 2000; Brower 2004).

In adults admitted to the ICU, hypoxaemia is a common clinical manifestation of inadequate gas exchange in the lungs (Petersson 2014). The condition can arise primarily from four different mechanisms: hypoventilation, ventilation or perfusion (V/Q) mismatch, intrapulmonary right-to-left blood shunting, or diffusion impairment, or a combination of these (Petersson 2014; Roussos 2003). Hypoventilation in the ICU is typically caused by an acute depression of the central nervous system, either through administration of sedative or analgesic agents, or due to critical illness with indirect (e.g. circulatory, hypoxic, or hypercapnic failure) or direct (e.g. traumatic brain injury, intracranial haemorrhage, or meningoencephalitis) cerebral affection. Hypoxaemia due to hypoventilation is always accompanied by hypercapnia since hypoventilation affects the alveolar clearance of carbon dioxide to a larger degree than the alveolar oxygenation, and hypoventilation does not affect the alveolar-arterial gradient (Petersson 2014; Roussos 2003). V/Q mismatch with a low V/Q ratio evolves when ventilation in certain lung regions is disproportionally decreased as compared to perfusion. This is seen in various conditions (Petersson 2014), including pneumonia, ARDS, pulmonary oedema, and chronic obstructive pulmonary disease (COPD) (Kent 2011). The impact of a low V/Q ratio is partially compensated by physiological hypoxic pulmonary vasoconstriction in the affected segments of the lung (Rodríguez-Roisin 2005). V/Q mismatch with a high V/Q ratio evolves when perfusion in certain lung regions is disproportionally decreased as compared to ventilation, as is classically seen in pulmonary embolism (Petersson 2014), but is also prevalent in COPD, Wagner 1977, and ARDS (Donahoe 2011). Intrapulmonary shunting is the consequence of complete V/Q mismatch with abolished ventilation which allows the passing of blood through sections of the pulmonary vascular bed without being oxygenated. This is seen in all types of pulmonary atelectasis (including absorption atelectasis) and is especially prevalent in ARDS and pneumonia (Petersson 2014). V/Q mismatch and intrapulmonary shunting are the most common causes of hypoxaemia in the ICU (Petersson 2014). Diffusion impairment occurs when the diffusion pathway for oxygen from the alveolar space to the pulmonary capillaries is pathologically increased, either acutely as seen in pneumonia, pulmonary oedema, or ARDS, or chronically as seen in the large group of interstitial lung diseases (Petersson 2014).

Description of the intervention

Administration of supplemental oxygen, defined as the fraction of inspired oxygen (FiO_2) above 0.21, is a frequent intervention in adults admitted to the ICU. Oxygen is often administered during acute conditions in the pre-hospital setting and during hospital admission. Adults admitted to the ICU often receive mechanical ventilation, and oxygen support to correct or prevent hypoxaemia. Treatment is usually a combination of ventilatory and non-ventilatory strategies (Esan 2010; Raoof 2010), where the aim is to reduce morbidity and mortality associated with hypoxaemia by restoring arterial oxygenation to normal values. Due to the administration of oxygen, adults often achieve supranormal levels of PaO_2 (de Graaff 2011; de Jonge 2008; Eastwood 2012; Itagaki 2015; Kraft 2018; Suzuki 2013; Zhang 2016).

Oxygen strategies used to treat hypoxaemia in adults admitted to the ICU are associated with harm in some studies, possibly because adults who receive oxygen in the ICU are the most ill, but it may also be that 'too much' oxygen is as harmful as 'too little' (Kallet 2013). The harms associated with lung injury caused by mechanical ventilation as well as by oxygen toxicity following high FiO_2 may exceed the benefit of normalizing oxygenation (PaO_2 and SaO_2).

How the intervention might work

The purpose of oxygen therapy is to increase oxygen delivery to tissues. Tissue hypoxia can cause cell death, but the precise level at which this occurs has not been determined, and the level may differ between tissues, organs, and individuals (O'Driscoll 2017).

Supplemental oxygen therapy has several potential advantages including maintenance of delivery of oxygen to tissues and prevention of organ dysfunction followed by anoxic injury (Budinger 2013). Several additional beneficial effects of supplemental oxygen have been proposed and include: induction of antioxidant enzymes, anti-inflammatory proteins, anti-inflammatory cytokines and certain growth factors; reduced postoperative infections, neutrophil activation, and markers of cerebral tissue breakdown; anti-apoptotic effects in brain and myocardium; normalization of cerebral extracellular homeostasis; and stabilization of the blood-brain barrier (Tan 2014).

High inspiratory oxygen concentrations have been associated with adverse outcomes in emergency medical conditions in patients with exacerbation of COPD (Austin 2010); after resuscitation after cardiac arrest (Kilgannon 2010); in patients with myocardial infarction (Cabello 2016); and in patients with traumatic brain injury (Brenner 2012). Additionally, treating perioperative adults with high FiO_2 may be associated with increased mortality without reducing surgical site infections in surgical adults (Wetterslev 2015). These adverse outcomes may be caused by postoperative pulmonary complications due to atelectasis formation, Benoit 2002; Rothen 1995a; Rothen 1995b, or pulmonary formation of reactive oxygen species (Chow 2003; Helmerhorst 2015; Kallet 2013). However, they may also be related to decreased local blood flow on normal and non-diseased vasculature induced by hyperoxaemic vasoconstriction (Sjöberg 2013), which has been described in the vascular system, for example in the heart and brain (Kenmure 1971; Watson 2000).

Knowledge about cell biology also suggests that oxygen might have harmful effects. Prolonged exposure to hyperoxia causes lung injury, which is thought to be caused by the production and accumulation of reactive oxygen species that overwhelm natural antioxidant defences and destroy cellular structures (Kallet 2013). Exposure to hyperoxia is associated with a boost in the production of reactive oxygen species, which eventually may overwhelm the cell repair processes, thereby causing cell injury (Crapo 1986). It has been proposed that reactive oxygen species may trigger apoptosis within pulmonary cells leading to necrosis, thereby causing an inflammation which damages lung tissue further (Zaher 2007).

Mechanical ventilation may in itself also be associated with complications including increased risk of pneumonia, impaired cardiac performance, and neuromuscular problems relating to sedation and muscle relaxants (Whitehead 2002). Also, applying pressure to the lungs can cause damage, which is known as ventilator-induced lung injury. Ventilator-associated lung injury has been shown to be augmented by hyperoxia in animal studies (Bailey 2003; Helmerhorst 2017b; Sinclair 2004).

Why it is important to do this review

The mainstay treatment for hypoxaemia is supplemental oxygen therapy, which is given to the vast majority of adults admitted to the ICU. It is estimated that 2 to 3 million adults yearly require mechanical ventilation in the ICU in high-income countries (Adhikari 2010), and is associated with morbidity, Kahn 2010, and mortality (Metnitz 2009; Wunsch 2010).

The current practice of oxygen administration has usually been more liberal and may result in hyperoxaemia or high partial tension of oxygen in the lungs (de Graaff 2011; de Jonge 2008; Itagaki 2015; Kraft 2018; Panwar 2013; Rachmale 2012; Suzuki 2013; Zhang 2016). Some studies have indicated an association between hyperoxaemia and mortality (Dahl 2015; Helmerhorst 2017a; Kilgannon 2010; Meyhoff 2012; Zhang 2016), whilst other studies have not (Bellomo 2011; Eastwood 2012; Kraft 2018; Raj 2013; Young 2012). Two meta-analyses of observational data found an association between hyperoxaemia and mortality after cardiac arrest, stroke, and traumatic brain injury (Damiani 2014), and overall across critically ill adults (Helmerhorst 2015). Permissive hypoxaemia has been studied by Gilbert-Kawai and colleagues (Gilbert-Kawai 2014), who compared permissive hypoxaemia to normoxaemia in critically ill adults in a systematic review but found no relevant randomized controlled trials (RCTs).

Although the possible adverse effects of hyperoxaemia are known, prevention of hypoxia through hyperoxaemia seems to be prioritized (Pannu 2016). The ideal target oxygenation for adults admitted to the ICU is uncertain due to limited evidence from RCTs. Despite a lack of robust evidence of effectiveness, oxygen administration is widely recommended in international clinical practice guidelines (AARC 2002; ARC 2014; Dellinger 2013; O'Driscoll 2017). However, it appears that a change towards a more restrictive approach is under way (Chu 2018; Siemieniuk 2018). Panwar and colleagues, Panwar 2015, and Girardis and colleagues, Girardis 2016, published data on RCTs comparing higher with lower oxygenation targets in adults admitted to the ICU, and Asfar and colleagues, Asfar 2017, published data on an RCT comparing high FiO_2 with lower oxygenation targets throughout the first 24 hours of ICU admission in adults with septic shock. Additional RCTs comparing high versus low targeted oxygen therapy in the critically ill are ongoing and may soon be published (NCT02321072; NCT03174002).

Oxygen is a common intervention in adults admitted to ICU and might have beneficial effects as well as harmful effects (Hafner 2015). The potential benefit of supplemental oxygen must be weighed against the potentially harmful effects of hyperoxaemia (Jakobsen 2013).

OBJECTIVES

To assess the benefits and harms of higher versus lower fraction of inspired oxygen or targets of arterial oxygenation in adults in intensive care units.

METHODS

Criteria for considering studies for this review

Types of studies

We included RCTs, irrespective of publication status, reported outcomes, publication date, and language.

We included unpublished trials only if methodological descriptions and trial data were provided by direct contact with trial authors or in written form.

We excluded randomized cross-over trials and quasi-randomized trials.

Types of participants

We included any adult aged 18 years or older admitted to the ICU. We only included participants if they were admitted to the ICU when randomization was allocated.

Types of interventions

We included trials having a clear differentiation of participants randomized to either a high-target (liberal) or a low-target (conservative) oxygenation strategy. Both mechanically ventilated and non-mechanically ventilated adults were eligible for inclusion. In order to include all relevant trials, we did not use predefined arbitrary thresholds of oxygenation for the two groups.

Intervention group: adults receiving a high-target (liberal) oxygenation strategy administered by any device, the aim of which was exposure to hyperoxia in the lungs, either by high FiO_2 or high-target PaO_2 or SaO_2/SpO_2.

Control group: adults receiving a low-target (conservative) oxygenation strategy administered by any device, the aim of which was to minimize exposure to hyperoxia in the lungs and reduce exposure to high FiO_2 or high-target PaO_2 or SaO_2/SpO_2.

Eligible trials were required to have a difference between the intervention and control groups of minimum 1 kPa in PaO_2, minimum 10% in FiO_2, or minimum 2% in SaO_2/SpO_2, either as aimed or achieved saturation or target. We only required one of these separation criteria to be fulfilled (PaO_2, SaO_2 or FiO_2), either aimed or achieved, for the trial to be eligible for inclusion.

We excluded trials/groups randomized to hypoxaemia (FiO_2 below 0.21, SaO_2/SpO_2 below 80%, and PaO_2 below 6 kPa). We furthermore excluded interventions with hyperbaric oxygen.

Types of outcome measures

Primary outcomes

1. All-cause mortality at the time point closest to three months.
2. Proportion of participants with one or more serious adverse events, defined as a dichotomous outcome according to participants having at least one serious adverse event or none at time point closest to three months. We defined a serious adverse event as any untoward medical occurrence that: resulted in death; was life-threatening; required hospitalization or prolongation of existing hospitalization; resulted in persistent or significant disability; or jeopardized the participant (ICH-GCP 1997). We performed two analyses on the proportion of participants with one or more serious adverse events according to the International Conference on Harmonisation Good Clinical Practice (ICH-GCP) (ICH-GCP 1997). We considered all other adverse events as non-serious.
3. Quality of life (any valid scale such as the 36-item Short Form Health Survey (SF-36)) at the time point closest to three months.

Secondary outcomes

1. Lung injury diagnosed after randomization (composite outcome) at the time point closest to three months. This composite outcome was defined as either: ARDS (defined by the onset of a known clinical insult within one week or acute worsening of respiratory symptoms; chest imaging; origin of oedema; and oxygenation may be mild, moderate, or severe (ARDS Definition Task Force 2012), or as defined by trialists); pulmonary fibrosis (defined as evolved from any cause or as defined by trialists); or pneumonia (defined as pneumonia occurring 48 hours or more after admission in non-intubated participants or pneumonia arising more than 48 to 72 hours after endotracheal intubation (ATS 2005), or as defined by trialists). As a secondary analysis, we analysed each component of the composite outcome separately. We performed two analyses on the proportion of participants with one or more lung injury.
2. Acute myocardial infarction diagnosed after randomization at the time point closest to three months (defined as the demonstration of myocardial cell death due to significant and sustained ischaemia (Thygesen 2012), or as defined by trialists).
3. Stroke diagnosed after randomization at the time point closest to three months (defined as central nervous system infarction, ischaemic stroke, silent central nervous system infarction, intracerebral haemorrhage, stroke caused by intracerebral haemorrhage, silent cerebral haemorrhage, subarachnoid haemorrhage, stroke caused by subarachnoid haemorrhage, stroke caused by cerebral venous thrombosis, and stroke not otherwise specified (Sacco 2013), or as defined by trialists).
4. Severe sepsis diagnosed after randomization at the time point closest to three months (defined as sepsis plus sepsis-induced organ dysfunction or tissue hypoperfusion (Dellinger 2013), or as defined by trialists).

Search methods for identification of studies

Electronic searches

We identified eligible RCTs through literature searching with systematic and sensitive search strategies specifically designed to identify relevant RCTs without restrictions to language, publication year, and journal.

We searched the following databases:

1. Cochrane Central Register of Controlled Trials (CENTRAL) (Issue 12, 2018) (Appendix 1);
2. MEDLINE (Ovid, 1946 to 20 December 2018) (Appendix 2);
3. Embase (Ovid, 1974 to 20 December 2018) (Appendix 3);
4. Science Citation Index (Web of Science, 1900 to 20 December 2018) (Appendix 4);
5. BIOSIS Previews (Web of Science, 1969 to 20 December 2018) (Appendix 5);
6. Cumulative Index to Nursing and Allied Health Literature (CINAHL) (EBSCO, 1981 to 20 December 2018) (Appendix 6);
7. Latin American and Caribbean Health Science Information database (LILACS) (1982 to 20 December 2018) (Appendix 7).

Searching other resources

We manually screened the reference lists of included trial reports, reviews, relevant papers, randomized and non-randomized trials, and editorials for potentially relevant trials.

Furthermore, we searched for ongoing and unpublished trials using the following trial registers:

1. US National Institutes of Health Ongoing Trials Register ClinicalTrials.gov (clinicaltrials.gov) (searched 21 December 2018);
2. World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) (www.who.int/ictrp/en/) (searched 21 December 2018);
3. EU Clinical Trials Register (www.clinicaltrialsregister.eu/) (searched 21 December 2018);
4. Australian New Zealand Clinical Trials Registry (ANZCTR) (www.anzctr.org.au/) (searched 21 December 2018).

Data collection and analysis

We used the following methods for data collection and data analyses.

Selection of studies

Three review authors (MB, OLS or SRK), independently and in pairs, screened each title and abstract of all reports identified by the searches. We obtained the full texts of those reports deemed potentially relevant and assessed these for inclusion in the review. Any disagreements were resolved by consensus or by consulting another review author (JW) when necessary.

Data extraction and management

Three review authors (MB, OLS or SRK), independently and in pairs, extracted predefined data of the included trials using a data collection form that was specifically designed and piloted by the review team (Appendix 8). We collected the following data:

1. trial: country, duration of the trial, date of publication, and type of trial;
2. participants: numbers randomized, numbers analysed, numbers lost to follow-up or withdrawn, type of population, mean or median age, sex, inclusion criteria, and exclusion criteria;
3. interventions: intervention, comparator, and concomitant interventions;
4. outcomes: predefined primary and secondary outcomes.

Any disagreements concerning the extracted data were resolved by discussion or by consulting a third review author (JW) when necessary. Where required, we contacted corresponding authors to clarify issues relating to data reporting or if further study details were needed.

Assessment of risk of bias in included studies

At least two review authors (MB, OLS or SRK) independently assessed the methodological quality of each included trial, as defined by the design of the trial and reporting. Any disagreements were resolved by discussion. We assessed the risk of bias according to the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011a), employing the criteria described in Appendix 9.

We assessed the following risk of bias domains for all included trials: random sequence generation, allocation sequence concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective outcome reporting, other potential sources of bias, and overall risk of bias. In addition, we assessed the domains blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, and selective outcome reporting for each outcome, which permitted an assessment of the risk of bias for each result. Based on this assessment, we defined the included trials and each outcome result as low risk of bias if all bias domains were judged as at low risk of bias.

We provided a summary assessment of the risk of bias across trials and for each important outcome (across domains) by preparing a 'Summary of findings' table, 'Risk of bias' graph, and a 'Risk of bias' summary figure (Higgins 2011a).

Measures of treatment effect

We calculated the risk ratio (RR) with 95% confidence interval (CI) and Trial Sequential Analysis (TSA) CI, adjusted for multiple outcomes, sparse data, and repetitive testing for dichotomous outcomes. For continuous outcomes, we planned to include both end scores and change scores in the analyses; we would use end scores if both were reported. We planned to calculate the mean difference (MD) and standardized mean difference (SMD) with 95% CIs and TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing for continuous outcomes.

Unit of analysis issues

Had we found a multi-arm trial that compared, for example, three different oxygenation targets, we would have combined the two experimental intervention groups of the study (if they each fulfilled the minimum difference compared with the control group of 1 kPa in PaO_2, 10% in FiO_2, and 2% in SaO_2/SpO_2) into a single group and compared these to the control group. If only one of the experimental groups fulfilled the minimum difference to the control, we would have compared this group to the control group.

For multi-arm trials that compare, for example, three different oxygenation targets, where the control group is the middle group, and the minimum difference in oxygenation target was fulfilled, we planned to compare the higher oxygenation group to the control group, as the lower group would be excluded due to being randomized to an extreme permissive hypoxaemia.

For cluster-randomized trials, we planned to define the ICU as the unit of allocation, and we would use the generic inverse-variance method in Review Manager 5 to calculate effect estimates for these trials (Review Manager 2014).

Dealing with missing data

We contacted trial investigators of the original reports for important missing data.

We did not impute missing data for any outcomes in the primary analysis, and we did not use intention-to-treat data if the original report did not contain such data.

If trial reports did not report standard deviations (SD), we would calculate the SDs using data from the trial report if possible.

We used imputed data in the sensitivity analysis for dichotomous and continuous outcomes (see Sensitivity analysis).

Assessment of heterogeneity

We assessed signs of heterogeneity by visual inspection of the forest plots.

We assessed the presence of statistical heterogeneity using the Chi^2 test with significance set at $P < 0.10$, and by measuring the quantities of heterogeneity using the I^2 statistic (Higgins 2003). Overall, we considered an I^2 statistic of 0% to 40% as not important, 30% to 60% as moderate, 50% to 90% as substantial, and 75% to 100% as considerable heterogeneity (Higgins 2011a). High statistical heterogeneity is generally more prevalent when meta-analysing continuous outcomes (Alba 2016). Because we anticipated large clinical heterogeneity as well as statistical heterogeneity, we generally preferred to use a random-effects model. However, if one or two trials dominate the acquired evidence (e.g. with more than 80% of the randomized participants) (Higgins 2002; MAGIC 2002; Woods 2002), the random-effects model may grossly overestimate the intervention effect; in such a situation, we would primarily report the results from a fixed-effect model. Hence, we primarily reported the result from the model with the most conservative point estimate of the two (Jakobsen 2014a), being the estimate closest to zero effect. If the two estimates were approximately equal, we used the estimate with the widest CI.

We explored potential clinical heterogeneity by conducting the prespecified subgroup analyses (see Subgroup analysis and investigation of heterogeneity).

Assessment of reporting biases

We planned to visually assess funnel plots for signs of asymmetry if an analysis included 10 or more trials (Higgins 2011a; Jakobsen 2014a).

We planned to test asymmetry within dichotomous outcomes using the Harbord test (Harbord 2006), and for continuous outcomes

using the asymmetry test (Egger 1997). We would also use adjusted rank correlation (Begg 1994).

Data synthesis

Meta-analysis

We undertook the systematic review according to the recommendations provided in the *Cochrane Handbook for Systematic Reviews of Interventions* and the eight-step assessment suggested by Jakobsen and colleagues (Higgins 2011a; Jakobsen 2014a), including TSA and calculation of Bayes factors. We performed meta-analysis of outcomes with comparable effect measures where more than one trial was included. If clinical and statistical heterogeneity were large or unexpected, we planned to reconsider performing meta-analysis. We used the statistical software Review Manager 5 provided by Cochrane and the TSA software version 0.9 CTU to meta-analyse data (Review Manager 2014; TSA 2011).

Assessment of significance

We assessed our intervention effects with both random-effects model meta-analyses, Deeks 2010; DerSimonian 1986; Mantel 1959, and fixed-effect model meta-analyses, DeMets 1987; Mantel 1959, and reported the most conservative estimate, being the point estimate closest to no effect or the estimate with the widest CI.

We used three co-primary outcomes and therefore considered $P \leq 0.025$ as statistically significant analysing the primary outcomes (Jakobsen 2014a; Jakobsen 2016). We used four co-secondary outcomes and therefore considered $P \leq 0.02$ as statistically significant analysing the secondary outcomes (Jakobsen 2014a). We used the eight-step procedure to assess if the thresholds for significance were crossed (Jakobsen 2014a).

Trial Sequential Analysis (TSA)

The chance of type I error (a false-positive finding) is increased when multiple testing is done (e.g. when analysing multiple primary and secondary outcomes or repeated testing of the data). In small studies, notably for binary outcomes, type I error is likely because the effect estimates tend to be more unstable (Mascha 2015). In meta-analyses the chance of finding a type I error is increased when they are updated over time when new trials are added (Mascha 2015). Cochrane recommends updating systematic reviews when, for example, new trials are available that will or might change the findings or credibility of the review, making it highly important to adjust for the multiplicity issue.

Current practice often uses a 0.05 significance criterion each time meta-analyses are updated, thus increasing the overall chance of a type I error (Mascha 2015). In addition, type II error (the probability of missing true findings) is a problem in many meta-analyses due to sparse data. Statistically significant meta-analyses with few participants have low reliability, and the interventional effect is often overrated (Turner 2013). In a random sample of 50 meta-analyses of anaesthesiology interventions with dichotomous outcome variables, Imberger and colleagues found 88% of the meta-analyses to be underpowered, meaning that although significant at $P < 0.05$, the meta-analyses should have included more participants (Imberger 2015). Furthermore, only 32% of the meta-analyses preserved the risk of type I error at 5% or less when powered for detecting a relative risk of 20% between groups (Imberger 2015).

Consequently, cumulative meta-analyses are at risk of producing random errors due to sparse data and multiple testing of accumulating data (Brok 2008; Brok 2009; Higgins 2011b; Imberger 2015; Mascha 2015; Pogue 1997; Terkawi 2016; Thorlund 2009; Wetterslev 2008), and TSA, Imberger 2016; TSA 2011, can be applied to assess this risk (Gluud 2011). The required information size and the required number of trials (i.e. the number of participants and trials needed in a meta-analysis to detect or reject an a priori prespecified realistic intervention effect) can be calculated to minimize random errors (Kulinskaya 2014; Wetterslev 2009). The required information size takes into account the event proportion in the control group, the assumption of a plausible relative risk reduction (RRR), and the heterogeneity variance of the meta-analysis (Turner 2013; Wetterslev 2009). Trial Sequential Analysis enables testing for significance to be conducted each time a new trial is included in the meta-analysis. On the basis of the required information size and the required number of trials, trial sequential monitoring boundaries can be constructed. This enables determination of the statistical inference concerning cumulative meta-analysis that has not yet reached the required information size (Imberger 2015; Mascha 2015; Terkawi 2016; Wetterslev 2008).

Firm evidence for benefit or harms may be established if the trial sequential monitoring boundary is crossed before reaching the required information size, in which case further trials may turn out to be superfluous. In contrast, if the boundary is not surpassed, the determination can be made that it is necessary to continue with further trials before a certain intervention effect can be detected or rejected. TSA can also assess firm evidence for lack of the postulated intervention effect, which occurs when the cumulative Z-score crosses the trial sequential monitoring boundaries for futility.

We used relatively conservative estimations of the anticipated intervention effect estimates in order to reduce the risk of random error (Jakobsen 2014a). Large anticipated intervention effects lead to small required information sizes, and the thresholds for significance will be less strict after the information size has been reached (Jakobsen 2014a).

We analysed all primary and secondary outcomes with TSA. We estimated the diversity (meta-analytic heterogeneity-adjustment factor) and calculated the required information size (Wetterslev 2009), based on the proportion of participants with an outcome in the control group. In addition, we used a family-wise error rate (FWER) of 5% (Jakobsen 2014a), leading to a statistical significance level of 2.5% for each of the co-primary outcomes, a beta of 10%, and a diversity (D^2), Wetterslev 2009, suggested by the trials in the meta-analysis (Jakobsen 2014a). We have presented TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing (Gluud 2011). As a sensitivity analysis, we used a diversity of 20% if the actual measured heterogeneity was zero because in this case heterogeneity will most likely increase when further trials are added until the required information size is reached. As anticipated intervention effects for the primary and secondary outcomes in the TSA, we used realistic a priori RRR of 20% or a 20% relative risk increase (RRI). Furthermore, we used an RRR or an RRI based on the confidence limit closest to null effect in the 95% CI in the traditional meta-analysis.

Bayes factor

A low P value indicates that an observed result is unlikely given the null hypothesis is true (Jakobsen 2014b). In meta-analyses, a low

P value can be misleading if there is also a low probability that data are compatible with an anticipated intervention effect (e.g. RRR or RRI of 20%). Bayes factor may be used to consider whether the probability that the actual measured difference in the effect of the compared interventions results from an a priori anticipated 'true' difference (Jakobsen 2014a). We calculated Bayes factors for the co-primary outcomes, which is the ratio between the probability of the meta-analysis result given the null hypothesis (H_0) is true divided by the probability of the meta-analysis result given the alternative hypothesis (H_A) is true using a Bayes factor calculator (Bayes factor calculator 2014). A high Bayes factor indicates that the meta-analysis result is produced by an intervention effect that is lower than the anticipated intervention effect, and thus the meta-analysis result should be interpreted with caution. A low Bayes factor together with a low P value corresponds to a high probability of an intervention effect similar to or greater than the anticipated intervention effect used in the calculation of the required information size. A Bayes factor less than 0.1 (a tenfold higher likelihood of compatibility with the alternative hypothesis than with the null hypothesis) has been suggested as the threshold for significance (Jakobsen 2014b).

Subgroup analysis and investigation of heterogeneity

We meta-analysed all included trials regardless of oxygenation strategy (PaO_2, SaO_2, SpO_2, FiO_2). We believed a meta-analysis of the specified strategies was feasible, as the amount of oxygen absorbed overlaps to a great extent. Whether FiO_2 is raised, or the aim is a higher target PaO_2, the result is that more oxygen is delivered, and the PaO_2 will be elevated in both strategies. However, we recognize that, especially in adults with ARDS, there are individuals where it would be extremely difficult to reach a predefined target of PaO_2 by either strategy, but both strategies would certainly expose the lungs to high oxygen levels, whilst other individuals may subsequently develop different PaO_2 levels with the two strategies.

We sought to determine if the efficacy and safety of the treatment options were influenced by types of ICU populations and type of oxygen administration.

We performed the following subgroup analyses.

1. According to different types of oxygen interventions:
   a. oxygen level defined by FiO_2 (as defined and set by trialists);
   b. oxygenation target measured using PaO_2 (as defined by trialists);
   c. oxygenation target measured using SaO_2 or SpO_2 (as defined by trialists);
   d. oxygenation target measured using either PaO_2 or SaO_2 or SpO_2 (as defined by trialists).
2. According to FiO_2 or oxygenation/target in the higher-oxygen-administration group:
   a. low targets defined as FiO_2 of 0.5 or lower or PaO_2 of 10 kPa or lower or SaO_2/SpO_2 of 95% or lower;
   b. high targets defined as FiO_2 above 0.5 or PaO_2 above 10 kPa or SaO_2/SpO_2 above 95%.
3. According to FiO_2 or oxygenation/target in the lower-oxygen-administration group:
   a. low targets defined as FiO_2 between or at 0.21 to 0.30 or PaO_2 between or at 6 kPa to 8 kPa or SaO_2/SpO_2 between or at 85% to 90%;
   b. high targets defined as FiO_2 above 0.30 to 0.40 or PaO_2 above 8 kPa to 10 kPa or SaO_2/SpO_2 above 90%.
4. According to ICU population:
   a. medical;
   b. surgical;
   c. mixed;
   d. adults with any respiratory failure;
   e. adults with any cerebral disease;
   f. adults with any heart disease;
   g. adults with any trauma;
   h. adults with COPD.
5. According to oxygen delivery system:
   a. invasive mechanical ventilation with endotracheal tube;
   b. any non-invasive oxygen administration.

Sensitivity analysis

To assess the potential impact of bias, we planned to conduct a sensitivity analysis for each outcome excluding trials with overall 'high risk of bias'.

To assess the potential impact of the missing data for dichotomous outcomes, we performed the two following analyses:

1. 'best-worst-case' scenario: we assumed that all participants lost to follow-up in the experimental group survived, had no serious adverse event, and had no morbidity; and all participants with missing outcomes in the control group did not survive, had a serious adverse event, and had morbidity;
2. 'worst-best-case' scenario: we assumed that all participants lost to follow-up in the experimental group did not survive, had a serious adverse event, and had morbidity; and all participants with missing outcomes in the control group did survive, had no serious adverse event, and had no morbidity.

Results from both scenarios are presented in the review.

To assess the potential impact of the missing data for continuous outcomes, we planned to perform the two following analyses:

1. 'best-worst-case' scenario: we assumed that all participants lost to follow-up in the experimental group had mean (from participants with follow-up) + 2 × SD, and all participants with missing outcomes in the control group had mean (from participants with follow-up) − 2 × SD;
2. 'worst-best-case' scenario: we assumed that all participants lost to follow-up in the experimental group had mean (from participants with follow-up) − 2 × SD, and all participants with missing outcomes in the control group had mean (from participants with follow-up) + 2 × SD (Jakobsen 2014a).

To assess the potential impact of missing SDs for continuous outcomes, we planned to perform the following sensitivity analyses: where SDs were missing and it was not possible to calculate them, we planned to impute SDs from trials with similar populations and

low risk of bias. If there were no such trials, we would impute SDs from trials with a similar population. As the final option, we planned to impute SDs from all trials.

1. To assess the potential impact of meta-analysing trials comparing two low targets (FiO_2 below 0.5 or PaO_2 below 10 kPa or SaO_2/SpO_2 below 95%) or two high targets (FiO_2 above 0.5 or PaO_2 above 10 kPa or SaO_2/SpO_2 above 95%), we performed sensitivity analysis excluding trials comparing two low targets or two high targets.
2. To assess the impact of longer follow-up, we performed analyses at maximum follow-up.

'Summary of findings' tables and GRADE

We used the GRADE system to assess the certainty of the body of evidence associated with each of the primary outcomes (all-cause mortality, proportion of participants with one or more serious adverse events, quality of life) and secondary outcomes (lung injury, acute myocardial infarction, stroke, sepsis) by constructing Summary of findings for the main comparison (Guyatt 2008), employing GRADEpro GDT software (GRADEpro GDT). For each primary and secondary outcome, we planned first to present summaries of findings in RCTs with an overall low risk of bias, and second results in all trials.

The GRADE approach appraises the certainty of a body of evidence based on the extent to which one can be confident that an estimate of effect or association reflects the item being assessed. The measure of a body of evidence considers within-study risk of bias, directness of the evidence, heterogeneity of the data, precision of effect estimates (Jakobsen 2014a), and risk of publication bias. We did not expect to identify any trials using adequate blinding of participants and personnel due to the practice of administration of oxygen. Hence, we planned to base our primary conclusions on the results of the analyses of the primary outcomes with low risk of bias in all 'Risk of bias' domains except 'blinding of participants and personnel'.

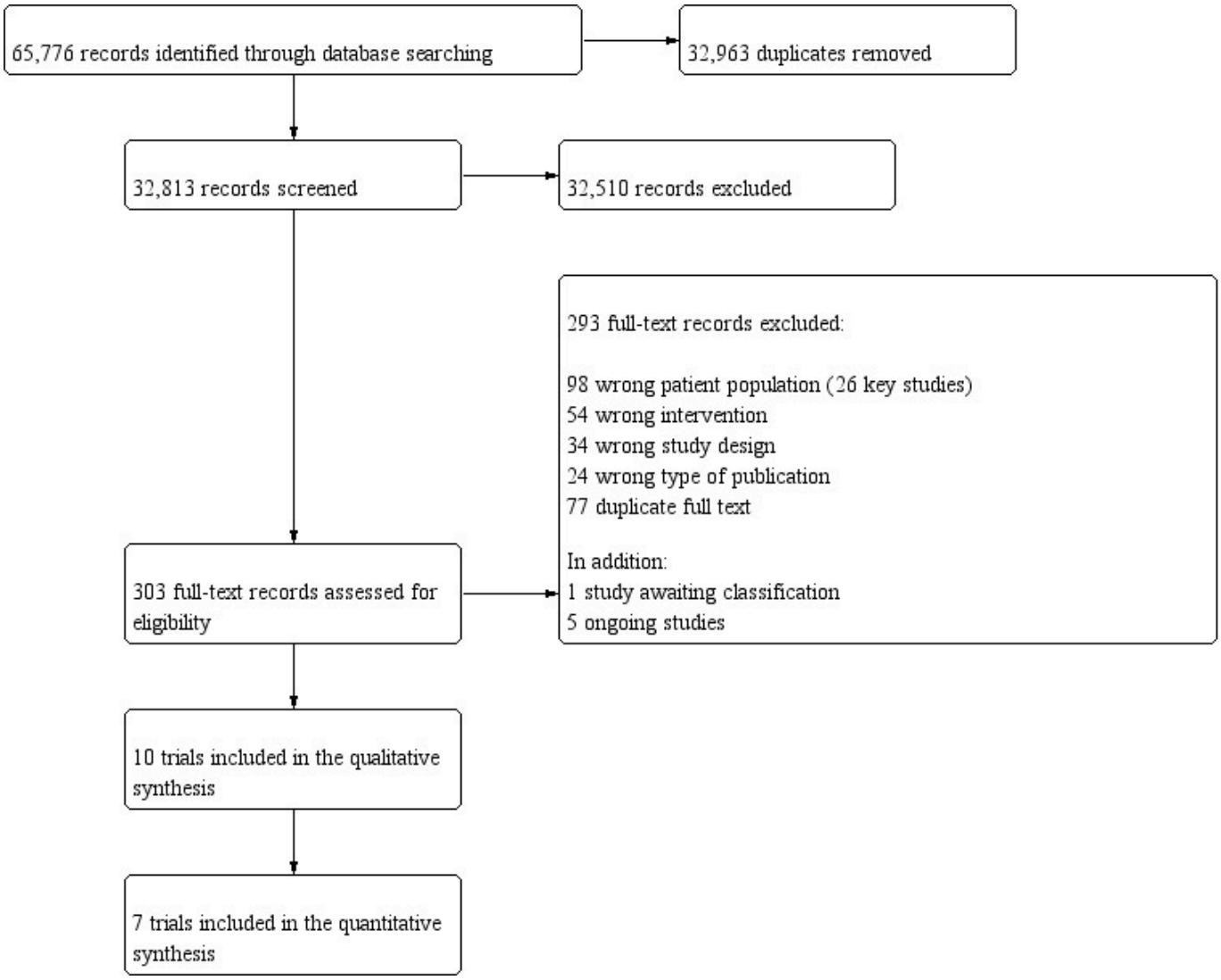
RESULTS

Description of studies

Results of the search

We screened 32,813 titles and abstracts, which included forward and backward citation searches, clinical trials registers, and grey literature. We obtained 303 full-text reports to assess eligibility (Figure 1) and excluded 293 references (98 wrong population, 54 wrong intervention, 34 wrong study design, 24 wrong type of publication, 77 duplicate full text, 5 ongoing studies, and 1 study awaiting classification) from the meta-analyses.

Figure 1. Study flow diagram.



Ten RCTs including a total of 1458 participants fulfilled our inclusion criteria. We approached all 10 corresponding authors to request missing or unclear information and received a reply from six. Detailed descriptions are shown in the Characteristics of included studies table.

See Figure 1.

Included studies

We included 10 RCTs involving a total of 1458 participants randomly assigned to a higher versus lower fraction of inspired oxygen or targets of arterial oxygenation. Seven trials reported on outcomes for this review (Characteristics of included studies).

Trial characteristics

Seven trials reported on mortality (1285 participants) (Asfar 2017; Girardis 2016; Gomersall 2002; Jakkula 2018; Lång 2018; Mazdeh 2015; Panwar 2016).

The same seven trials reported on the proportion of participants with one or more serious adverse events or any serious adverse event (1285 participants).

Five trials reported on lung injury (1167 participants) (Asfar 2017; Girardis 2016; Jakkula 2018; Lång 2018; Panwar 2016), and one trial reported on sepsis (445 participants) (Girardis 2016). Three trials did not report on any of our outcomes (Ishii 2018; Taher 2016; Young 2017). Eight trials used a two-arm, parallel-group design, and two trials used a two-factorial design. The trials were published from 2002 to 2018. Five trials were conducted in Europe; two in Iran; one in New Zealand; one in Australia, New Zealand, and France; and one in Japan.

See Characteristics of included studies.

Participants

The number of participants in the trials ranged from 36 to 480. The approximate weighted mean age of participants was 61 years, and the approximate mean proportion of men was 64%.

All trials included adults admitted to the ICU. Five trials included adults admitted to a multidisciplinary ICU (Asfar 2017; Girardis 2016; Gomersall 2002; Panwar 2016; Young 2017), and one to a surgical ICU (Ishii 2018). Two trials included adults with traumatic brain injury (Lång 2018; Taher 2016); one trial adults after cardiac arrest and resuscitation (Jakkula 2018); and one trial adults with stroke (Mazdeh 2015). Six trials included adults receiving invasive mechanical ventilation; three trials adults receiving any non-invasive oxygen administration; and one trial both adults on invasive mechanical ventilation and adults receiving non-invasive oxygen administration.

Funding

Seven trials were funded by public grants (Asfar 2017; Girardis 2016; Gomersall 2002; Lång 2018; Mazdeh 2015; Panwar 2016; Taher 2016; Young 2017); one trial did not report how it was funded (Ishii 2018); and one trial was funded by public and private funds and specified that funding bodies had no input regarding the design, management, or reporting of the trial (Jakkula 2018).

Experimental intervention

Of the 10 included trials, four trials randomized participants to higher versus lower oxygen using FiO_2 (Ishii 2018; Lång 2018; Mazdeh 2015; Taher 2016); five trials randomized participants to an oxygenation target (Girardis 2016; Gomersall 2002; Jakkula 2018; Panwar 2016; Young 2017); and one trial randomized participants to a specific FiO_2 in the experimental group and to target an oxygen saturation in the control group (Table 1) (Asfar 2017).

Of the five trials using FiO_2 in the experimental group, two trials used a FiO_2 of 1.0 (Asfar 2017; Ishii 2018); one used FiO_2 of 0.80 (Taher 2016); one used FiO_2 of 0.70 (Lång 2018); and one trial used FiO_2 of 0.50 (Mazdeh 2015). Of the five trials aiming to reach a target in the experimental group, one trial targeted an SpO_2 of 97% to 100% (Girardis 2016); one trial targeted an SpO_2 of ≥ 96% (Panwar 2016); one trial targeted a PaO_2 above 9.0 kPa (67.5 mmHg) (Gomersall 2002); one trial targeted 20 to 25 kPa (150 to 187.5 mmHg) (Jakkula 2018); and one trial randomized participants to standard care (no specific measures taken to avoid high FiO_2 or SpO_2; however, FiO_2< 0.30 was discouraged) (Young 2017).

Two trials were categorized as using a low target in the experimental (higher) group (Gomersall 2002; Mazdeh 2015), and seven trials were categorized as using a high target in the experimental group (Asfar 2017; Girardis 2016; Ishii 2018; Jakkula 2018; Lång 2018; Panwar 2016; Taher 2016). One trial could not be categorized according to our definitions, as no specific target was used (Young 2017).

Comparator intervention

Three trials used FiO_2 in the control group; one trial used expected FiO_2 to achieve a PaO_2 of 100 mmHg (13.3 kPa) (Ishii 2018); one trial used FiO_2 of 0.40 (Lång 2018); and one trial used FiO_2 of 0.50 (Table 1) (Taher 2016). Six trials used a target in the control group: one trial used SpO_2 88% to 92% (Panwar 2016); one trial used SaO_2 between 88% and 95% (Asfar 2017); one trial used SpO_2 between 94% and 98% (Girardis 2016); one trial used PaO_2 of > 6.6 kPa (50 mmHg) (Gomersall 2002); one trial used SpO_2 between 95% and 98% (Jakkula 2018); and one trial used SaO_2/SpO_2 between 91% to 96% (Young 2017). One trial used no supplemental oxygen (Mazdeh 2015).

Six trials were categorized as using a low target in the control group (Asfar 2017; Gomersall 2002; Mazdeh 2015; Panwar 2016; Taher 2016; Young 2017), and four trials were categorized as using a high target in the control group (Girardis 2016; Ishii 2018; Jakkula 2018; Lång 2018).

Excluded studies

We excluded RCTs of higher versus lower fraction of inspired oxygen or targets of arterial oxygenation that were conducted in populations not being admitted to an ICU. We listed the reasons for exclusion of 26 key excluded studies, which included RCTs of higher versus lower oxygen tensions for participants who were critically ill but not admitted to the ICU, as detailed in the Characteristics of excluded studies table.

Awaiting classification

One trial is awaiting classification (ICU-ROX 2019). This study was ongoing at the time of the search and will be included in future updates of this review. See Characteristics of studies awaiting classification.

Ongoing studies

We identified five ongoing trials (NCT02321072; NCT02713451; NCT03141099; NCT03174002; NCT03287466), which we will include in future updates of this review. See Characteristics of ongoing studies.

Risk of bias in included studies

Two trials had low risk of bias in all domains, except for blinding of participants and personnel. The remaining eight trials had high or unclear risk of bias in one or more bias domains other than blinding of participants and personnel. See the 'Risk of bias' graph (Figure 2) and 'Risk of bias' summary (Figure 3).

Figure 2. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.

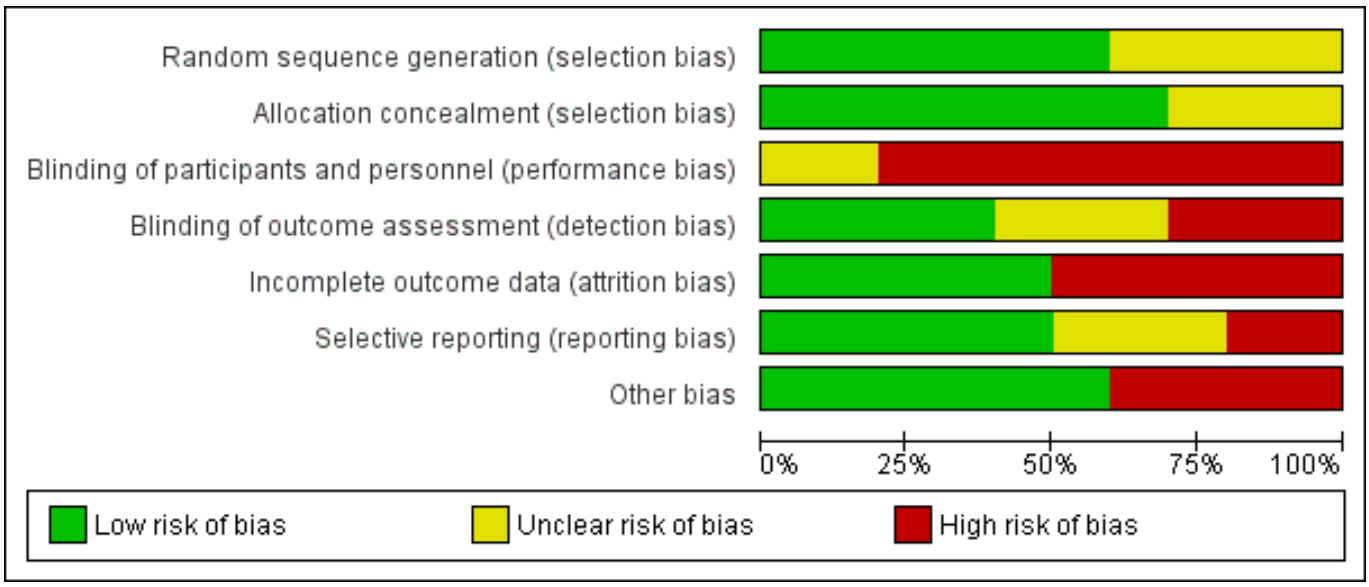


Figure 3. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.

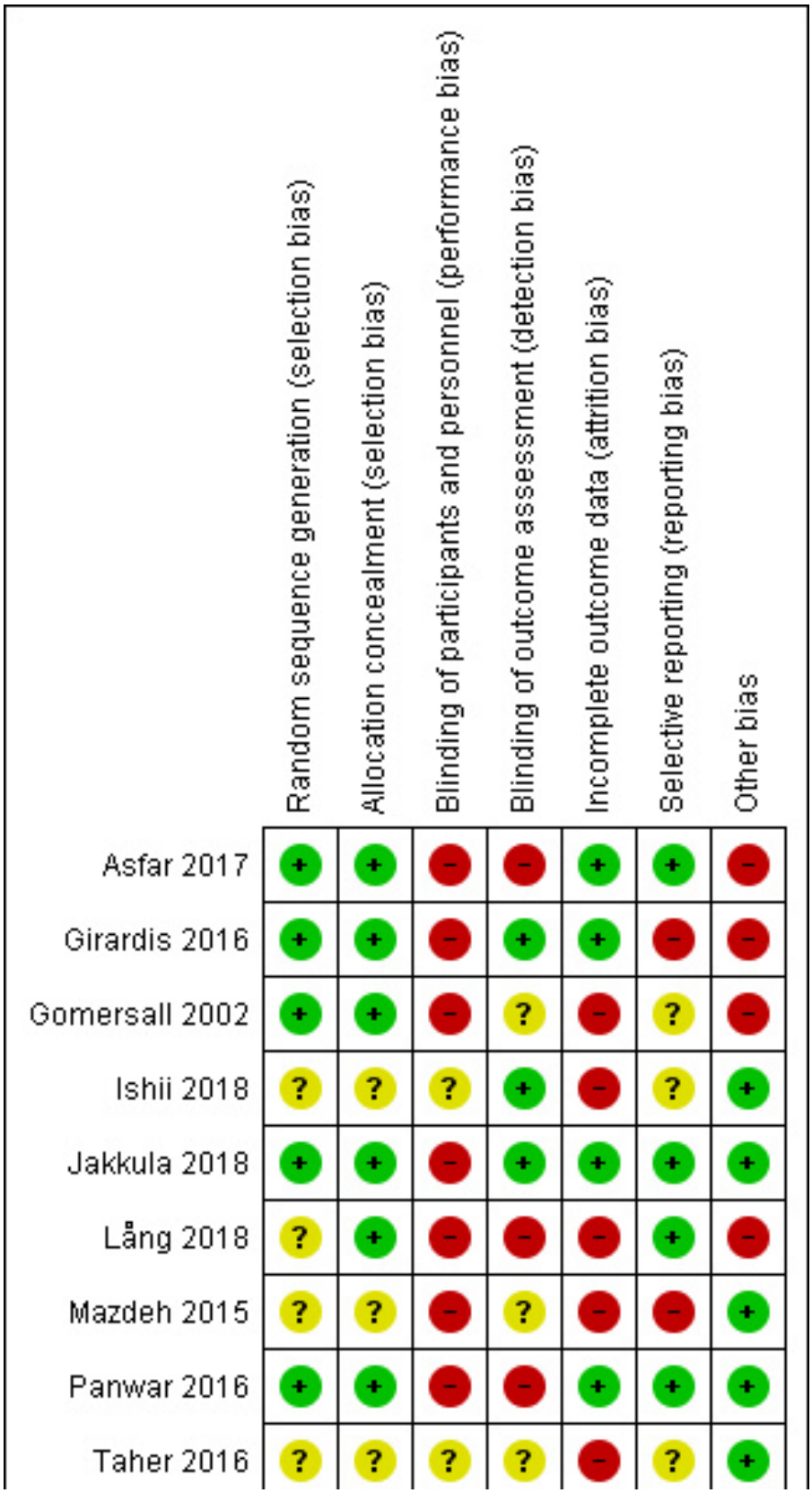


Figure 3. (Continued)

Allocation

Generation of the allocation sequence

Six trials described generation of the allocation sequence adequately, using computer-generated random numbers. Four trials did not describe the method of sequence generation and were considered to have an unclear risk of bias.

Allocation concealment

Seven trials described adequate allocation concealment, whilst three trials did not describe whether allocation concealment was adequate and were thus judged as having an unclear risk of bias.

Blinding

Blinding of participants and personnel

We judged no trials as having a low risk of bias for blinding of participants and personnel. Three trials blinded participants; five trials did not blind participants and personnel to the interventions; and two trials did not describe whether participants and personnel were blinded to the intervention and were thus judged as having an unclear risk of bias.

Blinding of outcome assessors

Four trials described adequate blinding of outcome assessors; three trials did not describe blinding of outcome assessors and were thus judged as at unclear risk of bias; and three trials used non-blinded outcome assessors.

Incomplete outcome data

Five trials provided numbers and reasons for dropouts and withdrawals or reported no dropouts or withdrawals, whilst five trials were judged as at high risk of bias due either to a high number of dropouts or lost to follow-up, dropouts and participants lost to follow-up not specified by allocation group, or participants being excluded due to mortality or lost to follow-up.

Selective reporting

Five trials were registered before randomization and reported on predefined outcomes; three trials provided insufficient information to determine if they had registered their trial or published a protocol before randomization; and two trials were judged as at high risk of bias due to being registered retrospectively.

Seven trials reported on all-cause mortality; one trial reported on proportion of participants with one or more serious adverse events, and seven trials reported on individual serious adverse events; no trials reported on quality of life; no trials reported on proportion of participants with lung injury, but five trials reported on either ARDS or pneumonia; no trials reported on acute myocardial infarction or stroke; and one trial reported on sepsis.

Other potential sources of bias

We assessed three trials as at high risk of bias due to early stopping: one trial was stopped after a pre-planned interim analysis for a reason that was not prespecified; one trial was stopped after an interim analysis that was not pre-planned; and one trial was stopped early due to lack of funding and slow recruitment.

We judged one trial as at high risk of bias due to a difference in co-interventions between groups, in which the participants in the low-oxygen tension group also received doxapram if they developed an acidosis with pH < 7.2, whereas those in the high-oxygen tension group received doxapram if they developed symptomatic acidosis.

We assessed two trials as at unclear risk of bias for this domain: one trial did not describe funding sources, and one trial was very poorly reported.

Overall risk of bias

We judged all included trials as at overall high risk of bias. Our assessment of risk of bias of the published trial reports is shown in Figure 2 and Figure 3 (Characteristics of included studies).

Effects of interventions

See: **Summary of findings for the main comparison** Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the ICU

See Summary of findings for the main comparison

Primary outcomes

All-cause mortality

Time point closest to three months

Four of 10 trials with a total of 1135 participants and a mean follow-up of 2 months (range 1 to 3 months) reported on all-cause mortality (Asfar 2017; Girardis 2016; Jakkula 2018; Panwar 2016).

A total of 39.1% in the higher group versus 33.1% in the lower group died. Meta-analysis showed evidence of a harmful effect of higher fraction of inspired oxygen or targets compared with lower fraction or targets of arterial oxygenation when assessing mortality (random-effects model risk ratio (RR) 1.18, 95% confidence interval (CI) 1.01 to 1.37; I^2 = 0%; 1135 participants; 4 trials; Analysis 1.1; very low-certainty evidence).

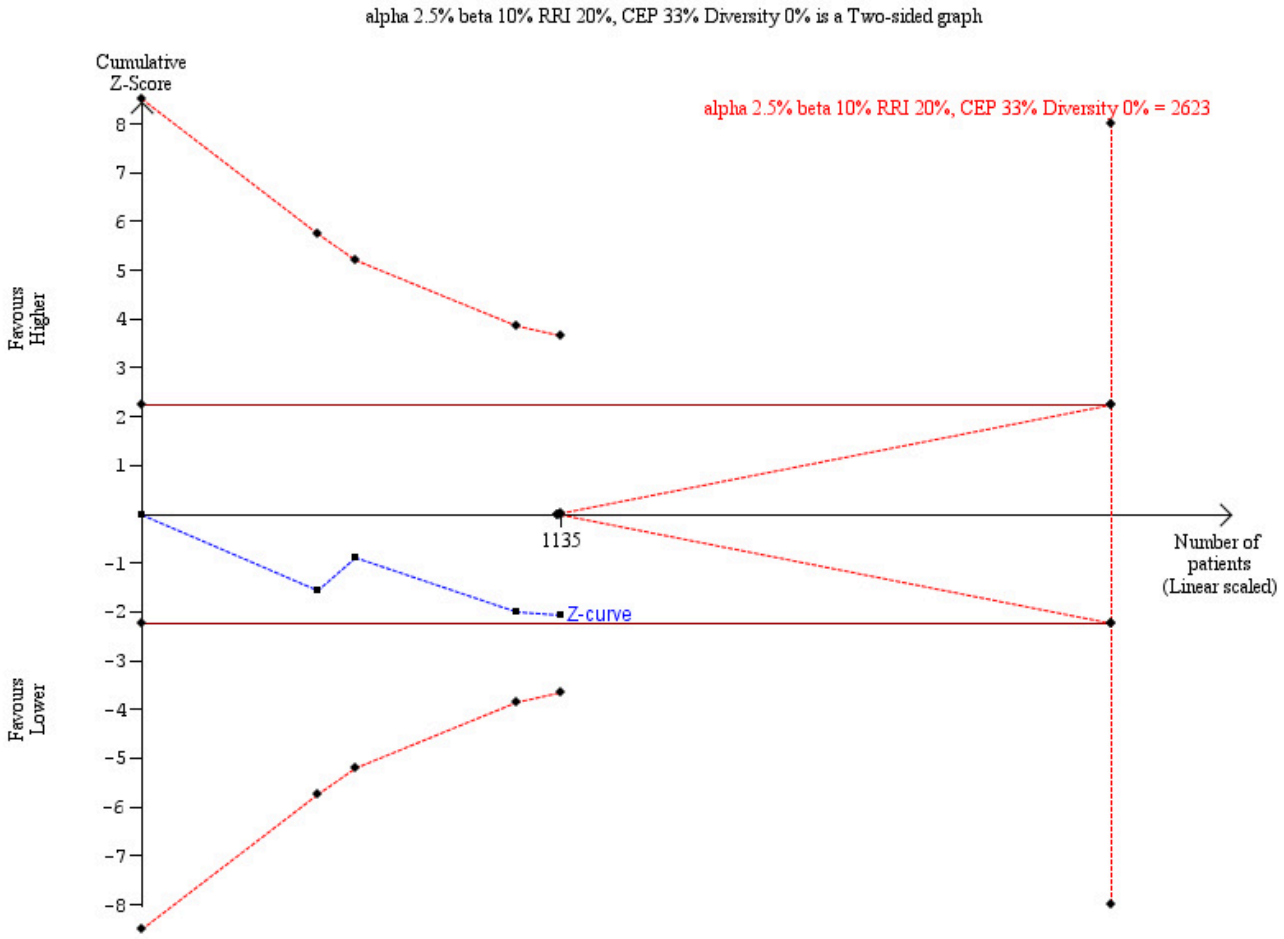
Heterogeneity

Neither visual inspection of the forest plot nor inconsistency factor (I^2 = 0%; 95% CI 0.00 to 0.59; P = 0.63) indicated statistical heterogeneity.

Trial Sequential Analysis

Trial Sequential Analysis showed that with an anticipated RRI of 20%, mortality in the control group of 33%, a type 1 error of 2.5%, a type 2 error of 10%, and a diversity of 0%, the required information size was 2623 participants. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility. This indicated that considering sparse data and repetitive testing, evidence was insufficient to confirm or refute a 20% RRI of higher versus lower oxygen (Figure 4). The TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing, for the intervention effect was 0.88 to 1.57.

Figure 4. Trial Sequential Analysis of the effects of higher versus lower fraction of inspired oxygen or targets of arterial oxygenation on the risk of mortality at the time point closest to three months. The analysis was based on a control event proportion (CEP) of 33%, a relative risk increase (RRI) of 20%, a type 1 error (alpha) of 2.5%, a type 2 error (beta) of 10%, and a diversity of 0%. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility.



Bayes factor

Bayes factors are presented in Table 2.

Sensitivity analyses

We were unable to perform the sensitivity analysis excluding trials at overall high risk of bias (except for blinding of participants and personnel) as only one trial reporting on mortality was at overall low risk of bias (except for blinding of participants and personnel) (Jakkula 2018).

The sensitivity analysis excluding trials comparing two low targets or two high targets indicated no evidence of a difference in the effect of higher versus lower oxygen on all-cause mortality (RR 1.11, 95% CI 0.92 to 1.35; I^2 = 0%; 537 participants; 2 trials; Analysis 1.2).

The sensitivity analysis based on missing data indicated that incomplete outcome data alone had the potential to influence the results:

- best-worst-case scenario random-effects meta-analysis: RR 1.13, 95% CI 0.97 to 1.31; 1149 participants; 4 trials; Analysis 1.3;
- worst-best-case scenario random-effects meta-analysis: RR 1.21, 95% CI 1.04 to 1.41; 1149 participants; 4 trials; Analysis 1.4).

However, both sensitivity analyses indicated harm of higher versus lower oxygen supplementation. Data were imputed for four trials (Asfar 2017; Girardis 2016; Jakkula 2018; Panwar 2016).

Subgroup analyses

We found no evidence of a difference in subgroup analyses according to different types of oxygen interventions (Analysis 1.5); FiO_2 or oxygenation target in the higher oxygen-administration group (analysis not applicable; Analysis 1.6); FiO_2 or oxygenation target in the lower oxygen-administration group (Analysis 1.7); ICU population (Analysis 1.8); and oxygen delivery system (Analysis 1.9).

Maximum follow-up

Seven of 10 trials with a total of 1285 participants and a mean follow-up of 3.33 months (range 1 month to 6 months) reported all-

cause mortality. A total of 36.41% in the higher group versus 31.39% in the lower group died. Meta-analysis showed evidence of a harmful effect of higher fraction of inspired oxygen or targets compared with lower fraction or targets of arterial oxygenation when assessing mortality (random-effects model RR 1.16, 95% CI 1.00 to 1.35; I^2 = 0%; 1285 participants; 7 trials; Analysis 2.1; very low-certainty evidence).

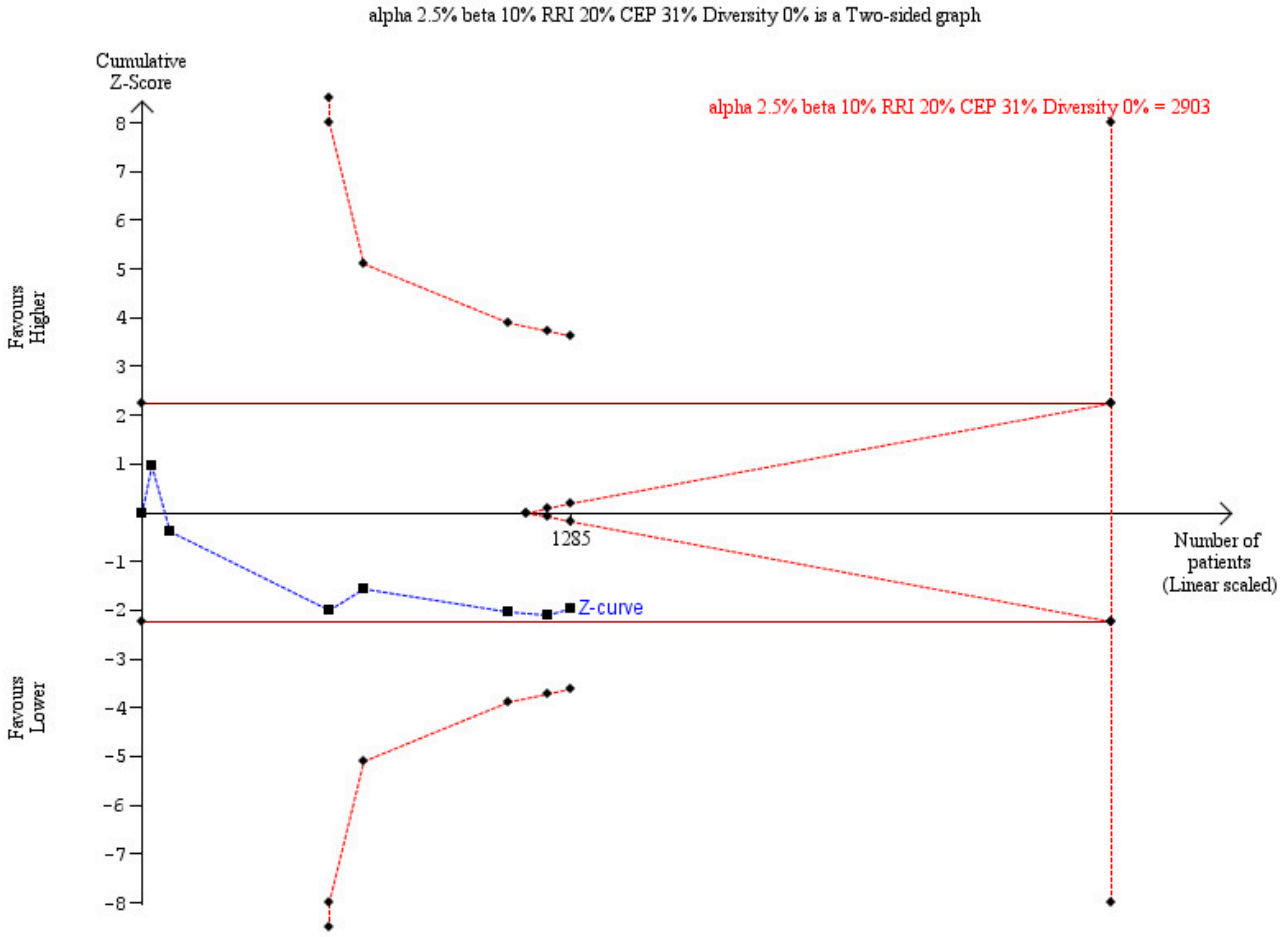
Heterogeneity

Neither visual inspection of the forest plot nor inconsistency factor (I^2 = 0%; 95% CI 0.00 to 0.46; P = 0.76) indicated any heterogeneity.

Trial Sequential Analysis

Trial Sequential Analysis showed that with an anticipated RRI of 20%, mortality in the control group of 31%, a type 1 error of 2.5%, a type 2 error of 10%, and a diversity of 0%, the required information size was 2903 participants. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility. This indicated that considering sparse data and repetitive testing, evidence was insufficient to confirm or refute a 20% RRI of higher versus lower oxygen (Figure 5). The TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing, for the intervention effect was 0.88 to 1.53.

Figure 5. Trial Sequential Analysis of the effects of higher versus lower fraction of inspired oxygen or targets of arterial oxygenation on the risk of mortality at maximum follow-up. The analysis was based on a control event proportion (CEP) of 31%, a relative risk increase (RRI) of 20%, a type 1 error (alpha) of 2.5%, a type 2 error (beta) of 10%, and a diversity of 0%. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility.



Bayes factor

Bayes factors are presented in Table 2.

Sensitivity analyses

We were unable to perform the sensitivity analysis excluding trials at overall high risk of bias (except for blinding of participants and personnel) as only one trial reporting on mortality was at overall low risk of bias (except for blinding of participants and personnel) (Jakkula 2018).

The sensitivity analysis excluding trials comparing two low targets or two high targets indicated a harmful effect of higher versus lower oxygen on all-cause mortality (RR 1.11, 95% CI 0.92 to 1.35; I^2 = 0%; 537 participants; 2 trials; Analysis 2.2).

The sensitivity analysis on missing data indicated that incomplete outcome data alone had the potential to influence the results:

- best-worst-case scenario random-effects meta-analysis: RR 1.11, 95% CI 0.96 to 1.28; 1306 participants; 7 trials; Analysis 2.3;
- worst-best-case scenario random-effects meta-analysis: RR 1.21, 95% CI 1.05 to 1.41; 1306 participants, 7 trials; Analysis 2.4.

However, both analyses indicated harm of higher versus lower oxygen supplementation. Data were imputed for six trials (Asfar 2017; Girardis 2016; Gomersall 2002; Jakkula 2018; Lång 2018; Panwar 2016).

Subgroup analyses

We found no evidence of a difference in subgroup analyses according to different types of oxygen interventions (Analysis 2.5); FiO_2 or oxygenation/target in the higher oxygen-administration group (Analysis 2.6); FiO_2 or oxygenation/target in the lower oxygen-administration group (Analysis 2.7); ICU population (Analysis 2.8); and oxygen delivery system (Analysis 2.9).

Proportion of participants with one or more serious adverse events

One of 10 trials reported on the proportion of participants with one or more serious adverse events as a composite outcome, according to our primary analysis on the proportion of participants with one or more serious adverse events (Asfar 2017). A total of 85% in the higher group versus 76% in the lower group had at least one serious adverse event. Another six trials, Girardis 2016; Gomersall 2002; Jakkula 2018; Lång 2018; Mazdeh 2015; Panwar 2016, reported on outcomes categorized by us as serious adverse events according to the ICH-GCP definition (ICH-GCP 1997).

As the reporting of serious adverse events as a combined outcome was not carried out according to the ICH-GCP recommendation, we estimated the reported proportion of participants with one or more serious adverse events in two ways:

1. by choosing the one specific serious adverse event with the highest proportion reported in each trial that addresses the lowest possible proportion of participants with one or more serious adverse events (somehow a best-case scenario);
2. by cumulating all reported serious adverse events, assuming that participants only experience one serious adverse event (the number of participants in each group will constitute a maximum), address the highest possible reported proportion of participants with one or more serious adverse events (somehow a worst-case scenario).

Time point closest to three months (follow-up range 3 days to 90 days)

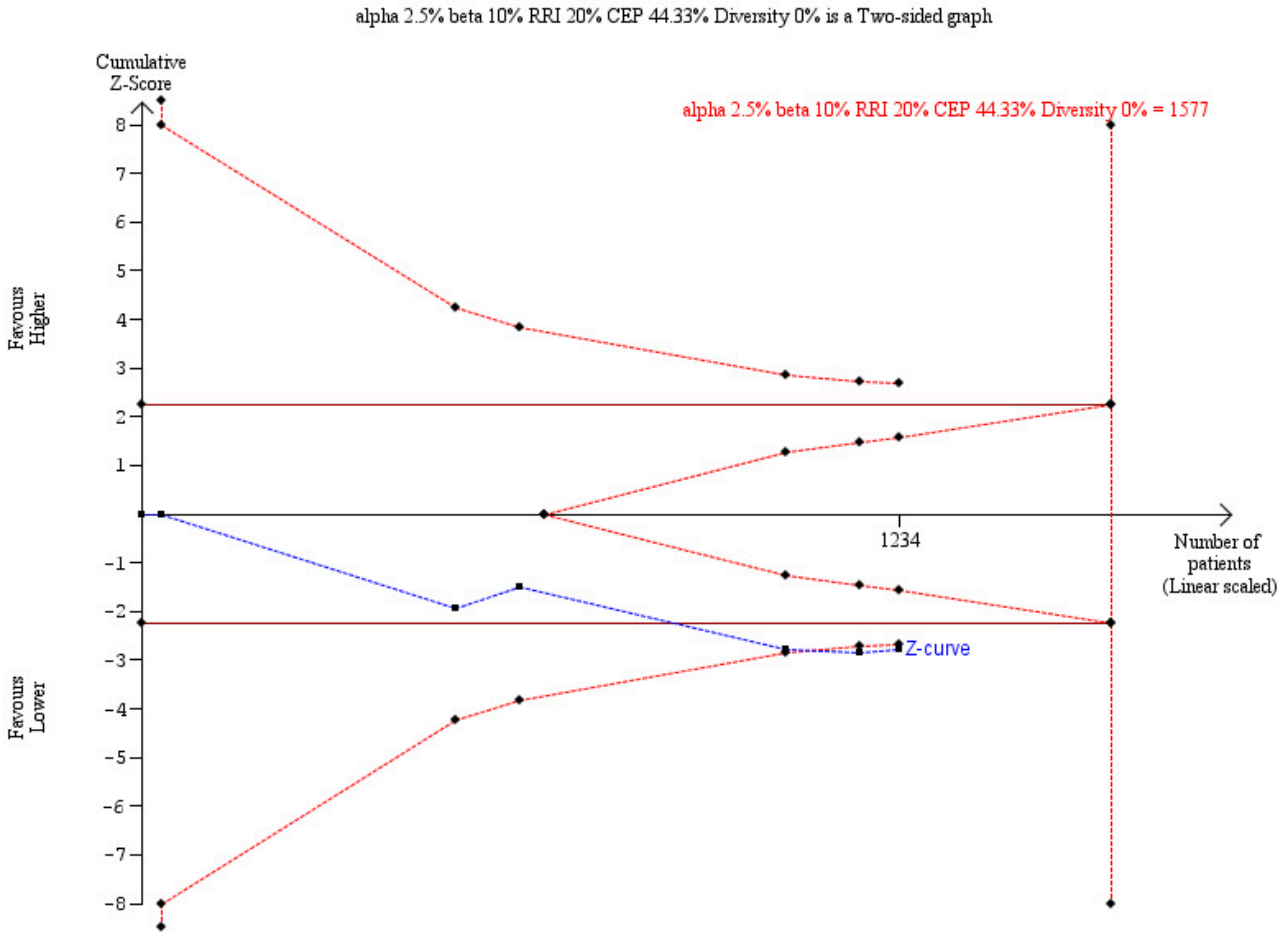
Meta-analysis showed evidence of a harmful effect of higher fraction of inspired oxygen or targets compared with lower fraction or targets of arterial oxygenation when assessing the estimated highest reported proportion of specific serious adverse events in each trial (random-effects model RR 1.13, 95% CI 1.04 to 1.23; I^2 = 0%; 1234 participants; 6 trials; Analysis 3.1; very low-certainty evidence). Individual types of serious adverse events included mortality (Girardis 2016; Jakkula 2018; Panwar 2016); proportion of participants with one or more serious adverse events (Asfar 2017); mechanical ventilation (reported as a poor outcome) (Gomersall 2002); and pneumonia (Lång 2018).

Meta-analysis showed no evidence of a difference of higher fraction of inspired oxygen or targets compared with lower fraction or targets of arterial oxygenation when assessing the estimated cumulated number of serious adverse events (random-effects model RR 1.08, 95% CI 0.99 to 1.18; I^2 = 49%; 1234 participants; 6 trials; Analysis 3.2; very low-certainty evidence). Individual types of serious adverse events included mortality; ARDS; pneumonia; sepsis; respiratory failure; cardiovascular failure; liver failure; renal failure; bloodstream infection; respiratory infection; surgical site infection; peripheral arterial thrombosis, pneumothorax; ventricular arrhythmias; new infections (composite outcome: when events were reported individually, they were not included in the analysis); haemodynamic instability; mechanical ventilation; severe hypercapnia and respiratory acidosis ($PaCO_2$ > 10 kPa and pH < 7.15); and unexplained brain oedema on computed tomography (CT) scan.

Trial Sequential Analysis

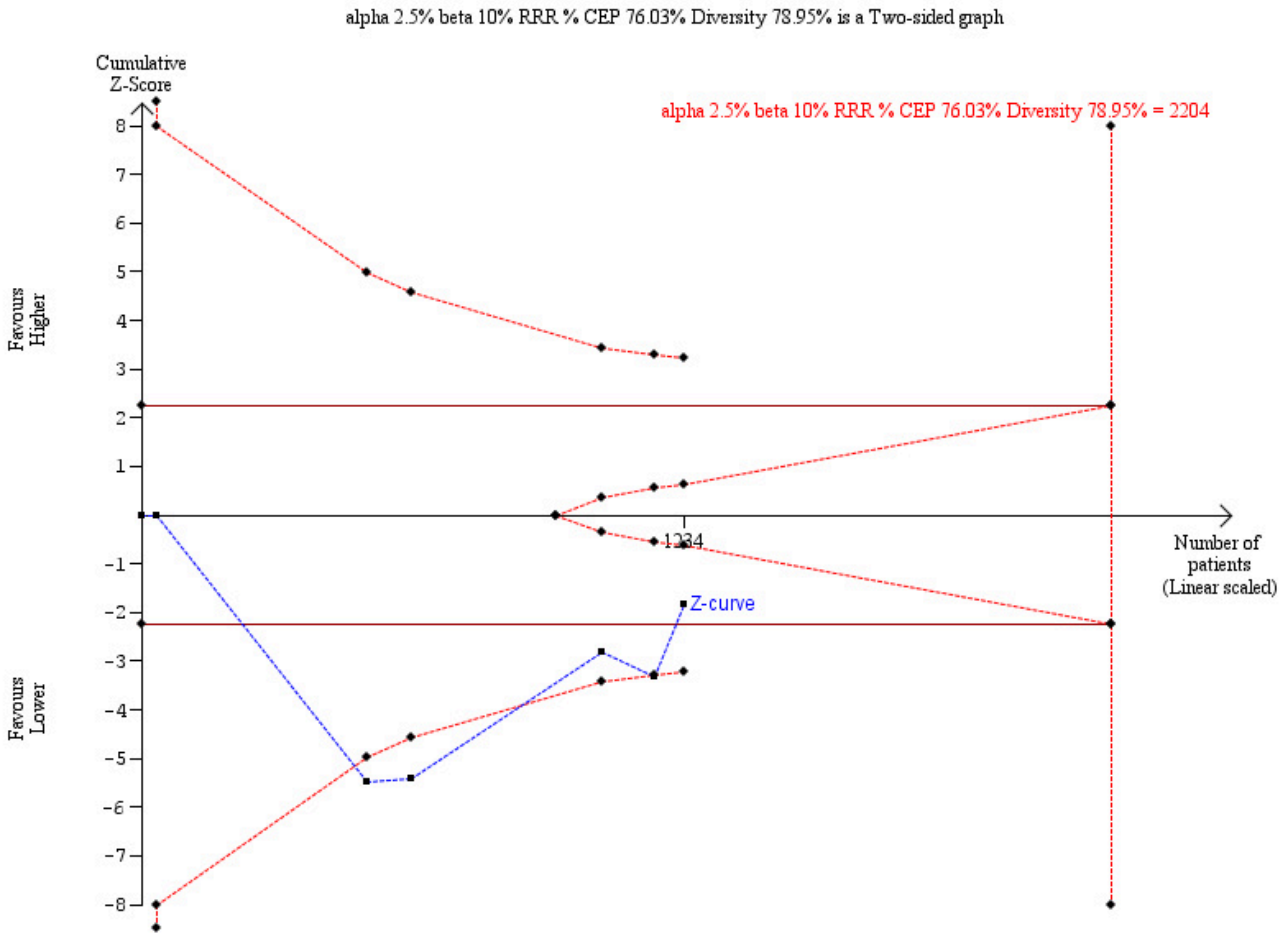
Trial Sequential Analysis of the estimated highest reported proportion of serious adverse events showed that with an anticipated RRI of 20%, serious adverse events in the control group of 44.33%, a type 1 error of 2.5%, a type 2 error of 10%, and a diversity of 0%, the required information size was 1577 participants (Figure 6). The cumulative Z-curve crossed the trial sequential monitoring boundary for harm, indicating there is evidence that higher versus lower oxygen may increase the relative risk of participants with one or more serious adverse events at three months follow-up. The TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing, for the intervention effect was 1.00 to 1.27.

Figure 6. Trial Sequential Analysis of the effects of higher versus lower fraction of inspired oxygen or targets of arterial oxygenation on the risk of the estimated highest reported proportion of serious adverse events at the time point closest to three months. The analysis was based on a control event proportion (CEP) of 44.33%, a relative risk increase (RRI) of 20%, a type 1 error (alpha) of 2.5%, a type 2 error (beta) of 10% and a diversity of 0%. The cumulative Z-curve crossed the trial sequential monitory boundary for harm.



Trial Sequential Analysis of the estimated cumulated number of serious adverse events showed that with an anticipated RRR of 20%, serious adverse events in the control group of 76.03%, a type 1 error of 2.5%, a type 2 error of 10%, and a diversity of 78.95%, the required information size was 2204 (Figure 7). The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility (although reaching futility boundary). This indicated that considering sparse data and repetitive testing, evidence was insufficient to confirm or refute a 20% RRR of higher versus lower oxygen. The TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing, for the intervention effect was 0.94 to 1.25.

Figure 7. Trial Sequential Analysis of the effects of higher versus lower fraction of inspired oxygen or targets of arterial oxygenation on the risk of the estimated cumulated proportion of serious adverse events at time point closest to three months. The analysis was based on a control event proportion (CEP) of 76.03%, a relative risk reduction (RRR) of 20%, a type 1 error (alpha) of 2.5%, a type 2 error (beta) of 10%, and a diversity of 78.95%. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility.



Bayes factor

Bayes factors are presented in Table 2.

Maximum follow-up (follow-up range 6 days to 6 months)

Meta-analysis showed evidence of a harmful effect of higher fraction of inspired oxygen or targets compared with lower fraction or targets of arterial oxygenation when assessing the estimated highest reported proportion of serious adverse events (random-effects model RR 1.13, 95% CI 1.04 to 1.23; I^2 = 0%; 1285 participants; 7 trials; Analysis 4.1). Individual types of serious adverse events included mortality (Girardis 2016; Jakkula 2018; Lång 2018; Mazdeh 2015; Panwar 2016); proportion of participants with one or more serious adverse events (Asfar 2017); and mechanical ventilation (Gomersall 2002).

Meta-analysis showed no evidence of a difference between higher fraction of inspired oxygen or targets compared with lower fraction or targets of arterial oxygenation when assessing the estimated cumulated number of serious adverse events (random-effects model RR 1.07, 95% CI 0.97 to 1.18; I^2 = 49%; 1285 participants; 7 trials; Analysis 4.2). Individual types of serious adverse events included mortality; ARDS; pneumonia; sepsis; respiratory failure; cardiovascular failure; liver failure; renal failure; bloodstream infection; respiratory infection; surgical site infection; peripheral arterial thrombosis, pneumothorax; ventricular arrhythmias; new infections (composite outcome: when events were reported individually, they were not included in the analysis); cardiac arrhythmia; coma; haemodynamic instability; mechanical ventilation; severe hypercapnia and respiratory acidosis ($PaCO_2$ > 10 kPa and pH < 7.15); and unexplained brain oedema on CT scan.

Trial Sequential Analysis

Trial Sequential Analysis of the estimated highest reported proportion of serious adverse events showed that with an anticipated RRI of 20%, serious adverse events in the control group of 43.38%, a type 1 error of 2.5%, a type 2 error of 10%, and a diversity of 0%, the required information size was 1644 participants. The cumulative Z-curve crossed the trial sequential monitoring boundary for harm.

This indicated that there was firm evidence that higher versus lower oxygen increases serious adverse events at maximum follow-up. The TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing, for the intervention effect was 1.01 to 1.27.

Trial Sequential Analysis of the estimated cumulated number of serious adverse events showed that with an anticipated RRR of 20%, serious adverse events in the control group of 74.92%, a type 1 error of 2.5%, a type 2 error of 10%, and a diversity of 82.80%, the required information size was 2826 participants. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility (although reaching futility boundary). This indicated that considering sparse data and repetitive testing, evidence was insufficient to confirm or refute a 20% RRI of higher versus lower oxygen. The TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing, for the intervention effect was 0.92 to 3.01.

Bayes factor

Bayes factors are presented in Table 2.

Quality of life (any valid scale such as the 36-item Short Form Health Survey (SF-36))

None of the included trials reported any data on quality of life at any time point.

Secondary outcomes

Lung injury

None of the 10 included trials reported any data on lung injury (as a composite outcome defined as either ARDS, pulmonary fibrosis, or pneumonia) at any time point. Five of the 10 trials reported on specific lung outcomes: ARDS (Jakkula 2018; Lång 2018; Panwar 2016); pulmonary fibrosis not reported; pneumonia (Asfar 2017; Girardis 2016; Lång 2018), during index admission.

We estimated the reported proportion of participants with one or more lung injury in two ways:

1. by choosing the one specific lung injury event with the highest proportion reported in each trial that addresses the lowest possible proportion of participants with one or more lung injuries (somehow a best-case scenario);
2. by cumulating all reported lung injury events, assuming that participants only experience one lung injury event (the number of participants in each group will constitute a maximum), address the highest possible reported proportion of participants with one or more lung injuries (somehow a worst-case scenario).

Time point closest to three months (follow-up range median 4 days to median 23 days)

Meta-analysis showed no evidence of a difference between higher fraction of inspired oxygen or targets compared with lower fraction or targets of arterial oxygenation when assessing the estimated highest reported proportion of lung injury (fixed-effect model RR 1.03, 95% CI 0.78 to 1.36; I^2 = 0%; 1167 participants; 5 trials; Analysis 5.1; very low-certainty evidence).

Meta-analysis showed no evidence of a difference between higher fraction of inspired oxygen or targets compared with lower fraction or targets of arterial oxygenation when assessing the estimated cumulated number of lung injury events (fixed-effect model RR 0.99, 95% CI 0.75 to 1.30; I^2 = 0%; 1167 participants; 5 trials; Analysis 5.2; very low-certainty evidence).

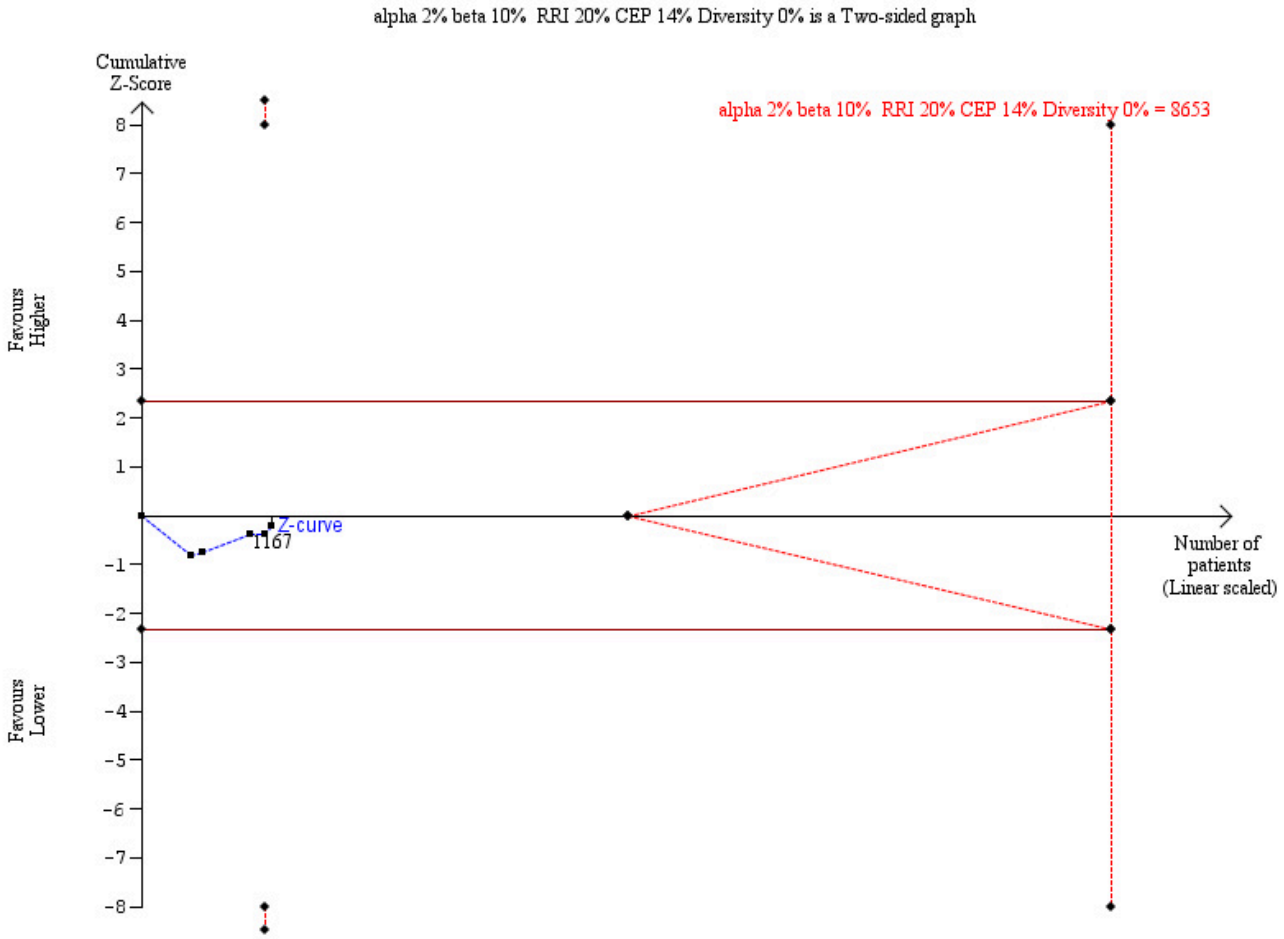
Three of 10 trials with a total of 288 participants reported ARDS. A total of 10.7% in the lower group versus 8.1% in the higher group had ARDS. Meta-analysis showed no evidence of a difference between higher fraction of inspired oxygen or targets compared with lower fraction or targets of arterial oxygenation when assessing ARDS (random-effects model RR 0.79, 95% CI 0.28 to 2.20; I^2 = 16%; 288 participants; 3 trials; Analysis 5.3; very low-certainty evidence).

Three of 10 trials with a total of 944 participants reported pneumonia. A total of 14.7% in the lower group versus 15.2% in the higher group had pneumonia. Meta-analysis showed no evidence of a difference between higher fraction of inspired oxygen or targets compared with lower fraction or targets of arterial oxygenation when assessing pneumonia (fixed-effect model RR 1.03, 95% CI 0.76 to 1.40; I^2 = 0%; 944 participants; 3 trials; Analysis 5.4; very low-certainty evidence).

Trial Sequential Analysis

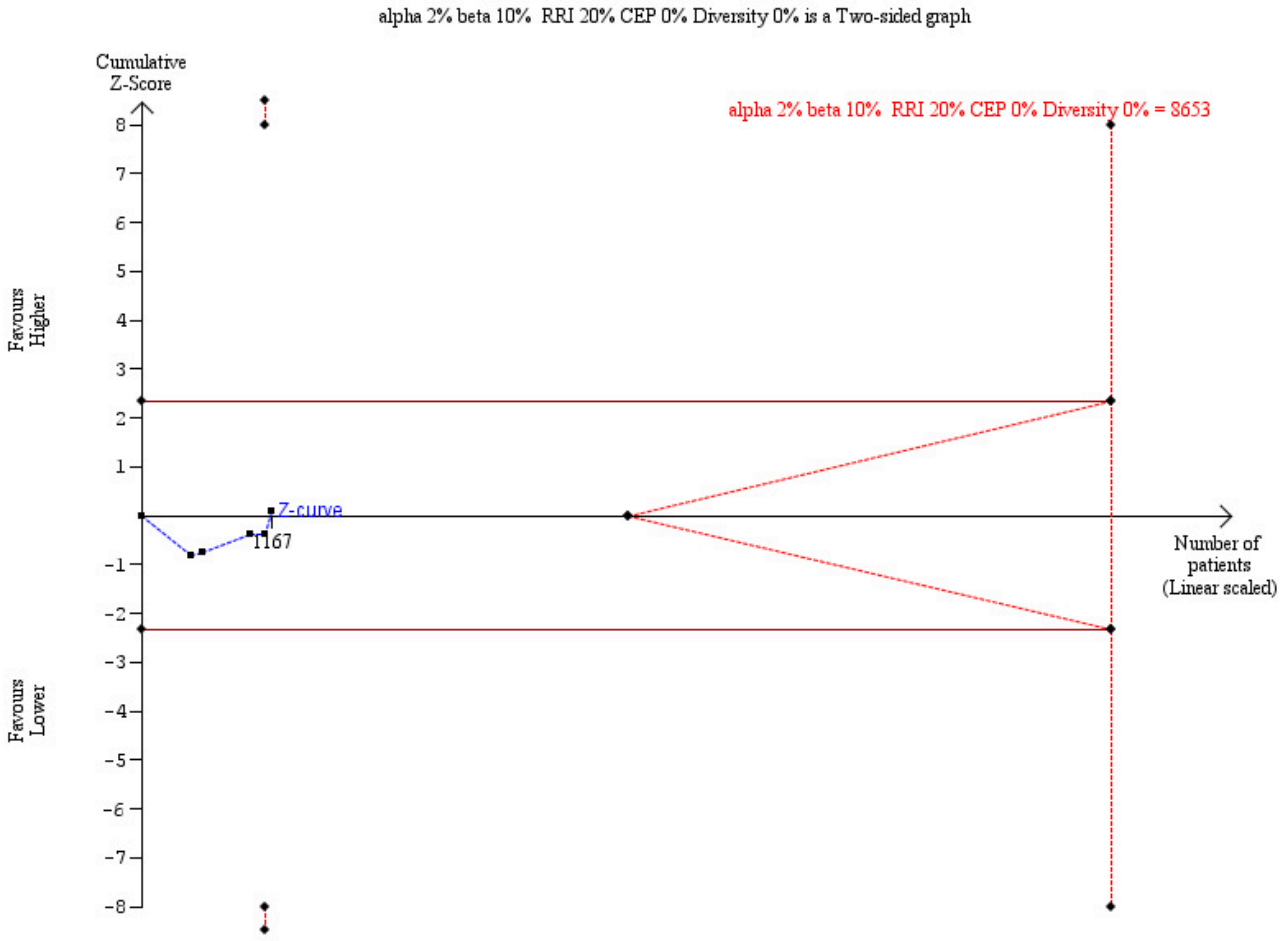
Trial Sequential Analysis of the estimated highest reported proportion of lung injuries showed that with an anticipated RRI of 20%, lung injury in the control group of 14%, a type 1 error of 2%, a type 2 error of 10%, and a diversity of 0%, the required information size was 8653 participants (Figure 8). The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility. This indicated that considering sparse data and repetitive testing, evidence was insufficient to confirm or refute a 20% RRR or 20% RRI for benefit or harm of higher versus lower oxygen. The TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing, for the intervention effect was 0.33 to 3.23.

Figure 8. Trial Sequential Analysis of the effects of higher versus lower fraction of inspired oxygen or targets of arterial oxygenation on the risk of the estimated highest reported proportion of lung injury at the time point closest to three months. The analysis was based on a control event proportion (CEP) of 14%, a relative risk increase (RRI) of 20%, a type 1 error (alpha) of 2%, a type 2 error (beta) of 10%, and a diversity of 0%. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility.



Trial Sequential Analysis of the estimated cumulated number of lung injuries showed that with an anticipated RRI of 20%, lung injury in the control group of 14%, a type 1 error of 2%, a type 2 error of 10%, and a diversity of 0%, the required information size was 8653 participants (Figure 9). The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility. This indicated that considering sparse data and repetitive testing, evidence was insufficient to confirm or refute a 20% RRR or 20% RRI for benefit or harm of higher versus lower oxygen. The TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing, for the intervention effect was 0.32 to 3.05.

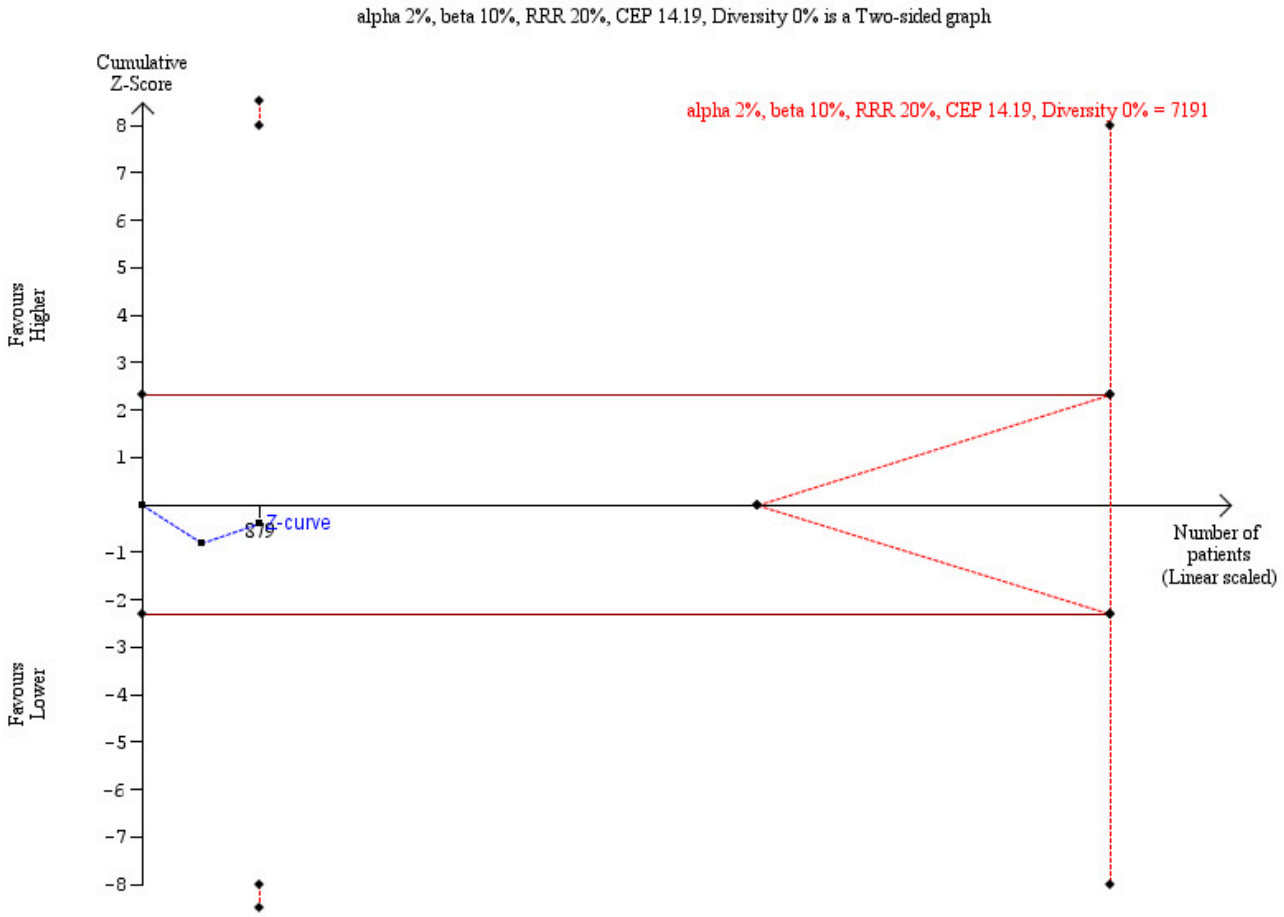
Figure 9. Trial Sequential Analysis of the effects of higher versus lower fraction of inspired oxygen or targets of arterial oxygenation on the risk of the estimated cumulated proportion of lung injury at the time point closest to three months. The analysis was based on a control event proportion (CEP) of 14%, a relative risk increase (RRI) of 20%, a type 1 error (alpha) of 2%, a type 2 error (beta) of 10%, and a diversity of 0%. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility.



We were unable to conduct Trial Sequential Analysis of ARDS due to insufficient information (1.34%). The required information size was 21,533 participants.

Trial Sequential Analysis of pneumonia showed that with an anticipated RRI of 20%, pneumonia in the control group of 14%, a type 1 error of 2%, a type 2 error of 10%, and a diversity of 0%, the required information size was 10,200 participants (Figure 10). The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility. This indicated that considering sparse data and repetitive testing, evidence was insufficient to confirm or refute a 20% RRR or 20% RRI for benefit or harm of higher versus lower oxygen. The TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing, for the intervention effect was 0.30 to 3.57.

Figure 10. Trial Sequential Analysis of the effects of higher versus lower fraction of inspired oxygen or targets of arterial oxygenation on the risk of pneumonia at the time point closest to three months. The analysis was based on a control event proportion (CEP) of 14%, a relative risk reduction (RRR) of 20%, a type 1 error (alpha) of 2%, a type 2 error (beta) of 10%, and a diversity of 0%. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility.



Maximum follow-up

None of the 10 trials reported any data on lung injury (as a composite outcome defined as either ARDS, pulmonary fibrosis, or pneumonia), including specific lung outcomes (ARDS, pulmonary fibrosis, or pneumonia), with longer follow-up than during index admission.

Acute myocardial infarction

None of the included trials reported any data on acute myocardial infarction at any time point.

Stroke

None of the included trials reported any data on stroke at any time point.

Sepsis

One trial reported on sepsis during ICU stay (median 6 days; interquartile range 1 to 11) (Girardis 2016). A total of 9.78% in the higher group versus 5.00% in the lower group had sepsis (RR 1.87, 95% CI 0.93 to 3.87; 1 study; 445 participants; very low-certainty evidence).

DISCUSSION

Summary of main results

We included 10 trials that randomized a total of 1458 participants in this systematic review. Seven trials with a total of 1285 participants contributed data to the analyses. We found no evidence for a beneficial effect of higher versus lower supplemental oxygen for adults admitted to the ICU.

Mortality seems to have been increased with higher supplemental oxygen at the time point closest to three months follow-up (RR 1.18, 95% CI 1.01 to 1.37; 4 studies; 1135 participants; I^2 = 0%; Analysis 1.1; very low-certainty of evidence) (Summary of findings for the main comparison). Trial Sequential Analysis, considering multiple outcomes, sparse data, and repetitive testing, revealed that the information size required to detect or reject an RRI of 20% was not achieved (Figure 11). When mortality was analysed at maximum follow-up, the traditional meta-analysis indicated increased mortality

with higher supplemental oxygen (Analysis 2.1), but TSA highlighted that the required information size to detect or reject a 20% RRI in mortality was not achieved (Figure 11).

The estimated highest reported proportion of serious adverse events at the time point closest to three months follow-up was significantly increased with higher supplemental oxygen (RR 1.13, 95% CI 1.04 to 1.23; 6 studies; 1234 participants; I^2 = 0%; Analysis 3.1; very low-certainty evidence). However, the estimated cumulated number of serious adverse events at the time point closest to three months follow-up did not show evidence of a difference (RR 1.08, 95% CI 0.99 to 1.18; 6 studies; 1234 participants; I^2 = 49%; Analysis 3.2; very low-certainty evidence). Trial Sequential Analysis showed that the monitoring boundary for harm for a 20% RRI was crossed when serious adverse events were analysed as the estimated highest proportion (Figure 6). However, when analysed as the estimated cumulated number of serious adverse events, the TSA revealed that the information size required to detect or reject an RRI of 20% was not achieved (Figure 7).

When serious adverse events were analysed at maximum follow-up, the traditional meta-analysis again showed that serious adverse events were increased with higher supplemental oxygen when analysed as the highest proportion (Analysis 4.1), but were not significantly increased when analysed as cumulated events (Analysis 4.2). Trial Sequential Analysis again showed that the monitoring boundary for harm for a 20% RRI was crossed when serious adverse events were analysed as estimated highest proportion, and when analysed as estimated cumulated number of serious adverse events, the TSA again revealed that the information size required to detect or reject an RRI of 20% was not achieved.

There was no evidence of a difference in lung injury with higher supplemental oxygen when analysed as a composite outcome nor as individual components of the composite outcome, but the evidence is very uncertain (Analysis 5.1; Analysis 5.2). However, TSA, considering multiple outcomes, sparse data, and repetitive testing, revealed that only 13% of the required information size was reached to detect or reject a 20% RRI, and that neither conventional nor trial sequential monitoring boundaries for benefit, harm, and futility had been crossed (Figure 8; Figure 9).

Only one trial reported on sepsis. Based on this one trial, we found that sepsis was not affected by higher supplemental oxygen (RR 1.87, 95% CI 0.93 to 3.87; 1 study; 445 participants; very low-certainty evidence).

No trials reported on quality of life, acute myocardial infarction, or stroke.

Overall completeness and applicability of evidence

We included all RCTs up to December 2018 comparing higher to lower oxygen fractions or targets of oxygenation in adults admitted to the ICU.

We found that clinical heterogeneity, especially relating to the intervention, but also to the population and setting, was present. Six trials were conducted in Europe, Australia, and New Zealand (Asfar 2017; Girardis 2016; Jakkula 2018; Lång 2018; Panwar 2016; Young 2017), two in Iran (Mazdeh 2015; Taher 2016), one in Hong Kong (Gomersall 2002), and one in Japan (Ishii 2018). The trials were conducted from 1994, Gomersall 2002, to 2018, Ishii 2018; Young 2017. Mean age spanned from 44 years, Lång 2018, to 68 years, Gomersall 2002, and the percentage of males versus females spanned from 49%, Jakkula 2018, to 84%, Lång 2018. All participants were admitted to the ICU; however, some trials included participants admitted to the ICU regardless of condition, whilst others included specific populations: five trials included adults from multidisciplinary ICUs (Asfar 2017; Girardis 2016; Gomersall 2002 Panwar 2016; Young 2017); two included adults with traumatic brain injury (Lång 2018; Taher 2016); one included adults admitted to a surgical ICU (Ishii 2018); one included adults with acute stroke (Mazdeh 2015); and one included adults resuscitated during out-of-hospital cardiac arrest (Jakkula 2018). In addition, disease severity differed, for example median Acute Physiology, Age, Chronic Health Evaluation II (APACHE II) of 22, Lång 2018, and median APACHE II of 28, Jakkula 2018. Furthermore, the interventions varied to a great extent. The duration of the intervention ranged from one hour, in Ishii 2018, to the entire duration of ICU admission, in Girardis 2016. The intervention targets compared also differed, and only three trials assessed targets categorized by us as higher versus lower oxygen fractions or targets of oxygenation (Asfar 2017; Panwar 2016; Taher 2016).

In general, statistical heterogeneity was low or moderate and was not explained by our subgroup analyses. Our sensitivity analysis on missing data (best-worst-case scenario and worst-best-case scenario) revealed that incomplete outcome data alone had the potential to influence the results on mortality; however, both analyses indicated harm with higher versus lower oxygen supplementation.

Only two trials had low risk of bias in all domains except for blinding of participants and personnel (Jakkula 2018; Young 2017). Only one of these trials contributed data to the meta-analyses (Jakkula 2018). The meta-analyses on mortality and lung injuries did not reach the required information size to detect or reject a 20% RRR or RRI. Trial Sequential Analysis on serious adverse events revealed that the trial sequential monitoring boundary for harm was crossed in one analysis (Figure 6), but not in the other (Figure 7).

Seven trials contributed data to the analyses on mortality and serious adverse events, and five trials contributed data to the analyses on lung injuries. No trials reported on quality of life, acute myocardial infarction, and stroke, and only one trial reported on sepsis.

Quality of the evidence

We used GRADE to assess the certainty of the evidence for the results on all-cause mortality, serious adverse events, quality of life, lung injury, acute myocardial infarction, stroke, and sepsis at the time point closest to three months (Summary of findings for the main comparison).

The GRADE assessment showed that the certainty of evidence was very low for mortality due to serious risk of bias, indirectness, and imprecision.

The certainty of the evidence was very low for the estimated highest reported proportion of serious adverse events due to serious risk of bias, indirectness, and imprecision. Trial Sequential Analysis showed that the trial sequential monitoring boundary for harm was crossed; hence, even with strict control of random errors, disregarding risk of bias, there is evidence that higher versus lower oxygen tensions increases the risk of serious adverse events by at least 20%.

The certainty of the evidence was very low for lung injury due to serious risk of bias, indirectness, and imprecision.

The certainty of the evidence was very low for sepsis due to serious risk of bias, inconsistency, and imprecision.

The certainty of the evidence for quality of life, acute myocardial infarction, and stroke was not estimable due to lack of data.

Potential biases in the review process

Strengths

We included trials regardless of publication type, publication status, language, and choice of outcomes. In all cases we contacted relevant trial authors if additional information was needed.

We used predefined, up-to-date systematic review methodology, and the methodology was not changed during the review process. We used GRADE to assess the certainty of the evidence and TSA as a sensitivity analysis with adjusted thresholds for significance to strictly control the risk of random errors; we thoroughly assessed the risk of bias of each trial to evaluate the risk of systematic errors (bias); and we used an eight-step procedure to assess if the thresholds for statistical and clinical significance were crossed (Jakobsen 2014a). This adds further robustness to our results and conclusions. We also tested the robustness of our results with sensitivity analyses.

We conducted two post hoc analyses that estimated the effects of higher versus lower oxygen supplementation on risk of having one or more serious adverse events and lung injury.

Limitations

We identified a high risk of clinical heterogeneity, especially within the interventions. The most obvious limitation was that trials did not use the same definition of lower targets and higher targets. Some trials used a fixed FiO_2, whilst others used a target interval, and the achieved oxygen saturation may end up being high even though participants were allocated to the lower group. Furthermore, the targets used in some trials were not adequately different to be categorized as trials comparing real high to real low targets. That being said, statistical heterogeneity seemed to be low.

Our 'Risk of bias' assessment showed that none of the included trials had an overall low risk of bias and none were fully blinded, which was not unexpected due to the complexity and difficulties of blinding interventions of oxygen supplementation for participants and personnel. Nevertheless, only data from one trial used blinded outcome assessors, which may still be used when blinding of participants and personnel is not feasible (Pocock 2015). Inadequate blinding is therefore a limitation in the included trials, as it is associated with exaggeration of beneficial intervention effects and underestimation of harmful effects (Hrobjartsson 2014; Savovic 2018). We thus could not rule out a biased effect estimate of the included trials. As a result, we downgraded the certainty of the evidence for all trials one level for risk of bias.

Only one trial reported serious adverse events as a composite outcome of participants with one or more serious adverse events. To estimate the effect on serious adverse events reported in the included trials, we conducted two analyses to estimate the effect on the proportion of participants having one or more serious adverse events, which may be expected to lie between these two extremes. None of the trials reported on lung injuries as a composite outcome, and thus the same method was applied. Each component was analysed separately for the lung injury outcome, but this was not done for serious adverse events. Each component of composite outcomes may not have similar degrees of severity, and therefore could bias the results of the outcome (Garattini 2016). If, for example, more severe serious adverse events occur in one intervention group, and other less severe serious adverse events occur in the other intervention group, then there is a risk of overlooking actual severity differences between the compared groups when analysing the composite outcome.

Furthermore, the analyses estimating the highest proportion of serious adverse events/lung injuries imply that participants included in the highest proportion also include participants having other serious adverse events. For example, if mortality is the highest proportion, then it is implied that all the participants that did not die did not experience another serious adverse event; this analysis thus underestimates the proportion of participants with one or more serious adverse events, as participants not included in the highest proportion would be expected to experience other serious adverse events not included in the highest proportion. In addition, the analyses estimating the cumulated proportion of serious adverse events/lung injuries imply that all participants who experience a serious adverse event had only this specific serious adverse event, which overestimates the proportion of participants with one or more serious adverse events, since a minimum of one participant would be expected to have more than one serious adverse event.

Only seven relatively small trials contributed data to the meta-analyses. An insufficient number of trials precluded an assessment of publication bias. Although we did not observe statistically significant heterogeneity in our subgroup analyses, they were naturally relatively small, thus we cannot exclude the possibility of subgroup differences.

Agreements and disagreements with other studies or reviews

Systematic reviews of observational data have found an association between hyperoxaemia and mortality in critically ill adults (Damiani 2014; Helmerhorst 2015), which has launched the initiation of several RCTs. Some meta-analyses of RCTs have been published in recent years (Cabello 2016; Chu 2018; Sepehrvand 2018; You 2018).

Critical illness of adults in the reviews is often defined differently or represented as subgroups. We included trials assessing adults admitted to and randomized in the ICU, whereas other reviews also included other settings, such as trauma, surgery, or pre-hospital initiated oxygen supplementation. Previous meta-analyses consistently report that too much supplemental oxygen may be/is harmful or not beneficial. However, it appears that none of these meta-analyses included proper bias risk assessment in their conclusions/recommendations. Limitations due to clinical heterogeneity are to a greater or lesser extent highlighted in the reviews, but these also seem not to be reflected in the conclusions. We performed TSA in order to control the risk of random errors in a cumulative meta-analysis and to prevent premature statements regarding the superiority of higher versus lower oxygen supplementation, which was also used by Chu and colleagues but without adjusting for multiple outcomes and using a possible inadequate power of 80% (Chu 2018).

Despite methodological discrepancies between our review and other meta-analyses and reviews, we agree with recently published reviews reporting a possible association between high oxygenation targets and mortality. However, we did not find the available evidence to be of high certainty (Chu 2018). Furthermore, we did not find that the current evidence necessitates a clinical practice guideline recommending a specific target of FiO_2, SpO_2, and PaO_2, particularly due to the very high heterogeneity in the types of interventions in the trials included in this review. (Rasmussen 2018; Siemieniuk 2018).

AUTHORS' CONCLUSIONS

Implications for practice

We are very uncertain about the effects of higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the intensive care unit on all-cause mortality, serious adverse events, lung injuries, and sepsis at the time point closest to three months due to very low-certainty evidence. Our results suggest that oxygen supplementation with higher versus lower fractions or oxygenation targets may increase mortality. None of the included trials reported the proportion of participants with one or more serious adverse events according to the International Conference on Harmonisation Good Clinical Practice (ICH-GCP) criteria; however, we found an increase in the number of serious adverse events reported by the trials with higher fractions of inspired oxygen or oxygenation targets using strict control of the risk of random errors. The effects of the interventions on quality of life, acute myocardial infarction, and stroke were inconclusive due to lack of data.

Implications for research

Randomized controlled trials assessing the benefits and harms of higher versus lower oxygen supplementation are needed. Such trials should be conducted with the lowest possible risk of bias, low risk of other design errors, and low risk of random errors. Future trials should focus their assessments on multidisciplinary intensive care units and critically ill adults in general and not only subgroups of this population group (Barbateskovic 2018). Oxygen supplementation is standard care, and the assessed intervention and duration should therefore reflect clinically relevant and accepted supplemental oxygen targets (Schjørring 2018). Furthermore, trials should aim to differentiate the intervention groups so that trials are in fact comparing higher versus lower oxygenation targets, and if possible by stratifying according to presence or absence of hypoxaemia at baseline. Patient-centred clinical outcomes should also be reported.

ACKNOWLEDGEMENTS

We would like to thank Arash Afshari (Content Editor), Nathan Pace (Statistical Editor), Daniel Martin and Rakshit Panwar (Peer Reviewers), Janet Wale (Consumer Referee), Janne Vendt (Information Specialist), Jane Cracknell and Teo Quay (Managing Editors), and Harald Herkner (Co-ordinating Editor) for their help and editorial advice during the preparation of this systematic review.

REFERENCES

References to studies included in this review

Asfar 2017 *{published data only}*

Asfar P, Schortgen F, Boisramé-Helms J, Charpentier J, Guérot E, Megarbane B, et al. Hyperoxia and hypertonic saline in patients with septic shock (HYPERS2S): a two-by-two factorial, multicentre, randomised, clinical trial. *Lancet. Respiratory Medicine* 2017;**5**(3):180-90. [PUBMED: 28219612]

Girardis 2016 *{published data only}*

Girardis M, Busani S, Damiani E, Donati A, Rinaldi L, Marudi A, et al. Effect of conservative vs conventional oxygen therapy on mortality among patients in an intensive care unit: the oxygen-ICU randomized clinical trial. *JAMA* 2016;**316**(15):1583-9. [PUBMED: 27706466]

Gomersall 2002 *{published data only}*

Gomersall CD, Joynt GM, Freebairn RC, Lai CK, Oh TE. Oxygen therapy for hypercapnic patients with chronic obstructive pulmonary disease and acute respiratory failure: a randomized, controlled pilot study. *Critical Care Medicine* 2002;**30**(1):113-6. [PUBMED: 11905405]

Ishii 2018 *{published data only}*

Ishii K, Morimatsu H, Hyodo T, Ono K, Hidaka H, Koyama Y, et al. Relationship between inspired oxygen concentration and atelectasis formation after extubation. *Critical Care Medicine* 2018;**46**(1 Suppl 1):533. [DOI: 10.1097/01.ccm.0000529104.66235.9e]

Jakkula 2018 *{published data only}*

Jakkula P, Reinikainen M, Hästbacka J, Loisa P, Tiainen M, Pettilä V, et al. Targeting two different levels of both arterial carbon dioxide and arterial oxygen after cardiac arrest and resuscitation: a randomised pilot trial. *Intensive Care Medicine* 2018;**44**(12):2112-21. [PUBMED: 30430209]

Lång 2018 *{published data only}*

Lång M, Skrifvars MB, Siironen J, Tanskanen P, Ala-Peijari M, Koivisto T, et al. A pilot study of hyperoxemia on neurological injury, inflammation and oxidative stress. *Acta Anaesthesiologica Scandinavica* 2018;**62**(6):801-10. [PUBMED: 29464691]

Mazdeh 2015 *{published data only}*

Mazdeh M, Taher A, Torabian S, Seifirad S. Effects of normobaric hyperoxia in severe acute stroke: a randomized controlled clinical trial study. *Acta Medica Iranica* 2015;**53**(11):676-80. [PUBMED: 26786987]

Panwar 2016 *{published data only}*

Panwar R, Hardie M, Bellomo R, Barrot L, Eastwood GM, Young PJ, et al. Conservative versus liberal oxygenation targets for mechanically ventilated patients. A pilot multicenter randomized controlled trial. *American Journal of Respiratory and Critical Care Medicine* 2016;**193**(1):43-51. [PUBMED: 26334785]

Taher 2016 *{published data only}*

Taher A, Pilehvari Z, Poorolajal J, Aghajanloo M. Effects of normobaric hyperoxia in traumatic brain injury: a randomized controlled clinical trial. *Trauma Monthly* 2016;**21**(1):e26772. [PUBMED: 27218057]

Young 2017 *{published data only}*

Young PJ, Mackle DM, Bailey MJ, Beasley RW, Bennett VL, Deane AM, et al. Intensive care unit randomised trial comparing two approaches to oxygen therapy (ICU-ROX): results of the pilot phase. *Critical Care and Resuscitation* 2017;**19**(4):344-54. [PUBMED: 29202261]

References to studies excluded from this review

Ali 2013 *{published data only}*

Ali K, Warusevitane A, Lally F, Sim J, Sills S, Pountain S, et al. The stroke oxygen pilot study: a randomized controlled trial of the effects of routine oxygen supplementation early after acute stroke - effect on key outcomes at six months. *PLOS ONE* 2013;**8**(6):e59274. [PUBMED: 23755093]

Amar 1994 *{published data only}*

Amar D, Greenberg MA, Menegus MA, Breitbart S. Should all patients undergoing cardiac catheterization or percutaneous transluminal coronary angioplasty receive oxygen?. *Chest* 1994;**105**(3):727-32. [PUBMED: 8131533]

Austin 2010 *{published data only}*

Austin MA, Wills KE, Blizzard L, Walters EH, Wood-Baker R. Effect of high flow oxygen on mortality in chronic obstructive pulmonary disease patients in prehospital setting: randomised controlled trial. *BMJ* 2010;**341**:c5462. [PUBMED: 20959284]

Bickel 2011 *{published data only}*

Bickel A, Gurevits M, Vamos R, Ivry S, Eitan A. Perioperative hyperoxygenation and wound site infection following surgery for acute appendicitis: a randomized, prospective, controlled trial. *Archives of Surgery* 2011;**146**(4):464-70. [PUBMED: 21502457]

Bray 2018 *{published data only}*

Bray JE, Hein C, Smith K, Stephenson M, Grantham H, Finn J. Oxygen titration after resuscitation from out-of-hospital cardiac arrest: a multi-centre, randomised controlled pilot study (the EXACT pilot trial). *Resuscitation* 2018;**128**:211-5. [PUBMED: 29684433]

Hofmann 2017 *{published data only}*

Hofmann R, James SK, Jernberg T, Lindahl B, Erlinge D, Witt N, et al. Oxygen therapy in suspected acute myocardial infarction. *New England Journal of Medicine* 2017;**377**(13):1240-9. [PUBMED: 28844200]

Huynh Ky 2017 *{published data only}*

Huynh Ky M, Bouchard PA, Morin J, L'Her E, Sarrazin JF, Lellouche F. Closed-loop adjustment of oxygen flowrate with FreeO2 in patients with acute coronary syndrome: comparison

of automated titration with FreeO2 (set at two SpO2 target) and of manual titration. A randomized controlled study. *Annals of Intensive Care* 2017;**7**(Suppl 1):O59.

Khoshnood 2018 *{published data only}*

Khoshnood A, Carlsson M, Akbarzadeh M, Bhiladvala P, Roijer A, Nordlund D, et al. Effect of oxygen therapy on myocardial salvage in ST elevation myocardial infarction: the randomized SOCCER trial. *European Journal of Emergency Medicine* 2018;**25**(2):78-84. [PUBMED: 27893526]

Khosnood 2017 *{published data only}*

Khoshnood A, Akbarzadeh M, Roijer A, Meurling C, Carlsson M, Bhiladvala P, et al. Effects of oxygen therapy on wall-motion score index in patients with ST elevation myocardial infarction - the randomized SOCCER trial. *Echocardiography* 2017;**34**(8):1130-7. [PUBMED: 28664557]

Kuisma 2006 *{published data only}*

Kuisma M, Boyd J, Voipio V, Alaspää A, Roine RO, Rosenberg P. Comparison of 30 and the 100% inspired oxygen concentrations during early post-resuscitation period: a randomised controlled pilot study. *Resuscitation* 2006;**69**(2):199-206. [PUBMED: 16500018]

Meyhoff 2009 *{published data only}*

Meyhoff CS, Wetterslev J, Jorgensen LN, Henneberg SW, Høgdall C, Lundvall L, et al. Effect of high perioperative oxygen fraction on surgical site infection and pulmonary complications after abdominal surgery: the PROXI randomized clinical trial. *JAMA* 2009;**302**(14):1543-50. [PUBMED: 19826023]

Padma 2010 *{published data only}*

Padma MV, Bhasin A, Bhatia R, Garg A, Singh MB, Tripathi M, et al. Normobaric oxygen therapy in acute ischemic stroke: a pilot study in Indian patients. *Annals of Indian Academy of Neurology* 2010;**13**(4):284-8. [PUBMED: 21264137]

Perrin 2011 *{published data only}*

Perrin K, Wijesinghe M, Healy B, Wadsworth K, Bowditch R, Bibby S, et al. Randomised controlled trial of high concentration versus titrated oxygen therapy in severe exacerbations of asthma. *Thorax* 2011;**66**(11):937-41. [PUBMED: 21597111]

Ranchord 2012 *{published data only}*

Ranchord AM, Argyle R, Beynon R, Perrin K, Sharma V, Weatherall M, et al. High-concentration versus titrated oxygen therapy in ST elevation myocardial infarction: a pilot randomized controlled trial. *American Heart Journal* 2012;**163**(2):168-75. [PUBMED: 22305833]

Rawles 1976 *{published data only}*

Rawles JM, Kenmure AC. Controlled trial of oxygen in uncomplicated myocardial infarction. *BMJ* 1976;**1**(6018):1121-3. [PUBMED: 773507]

Rodrigo 2003 *{published data only}*

Rodrigo GJ, Rodriquez Verde M, Peregalli V, Rodrigo C. Effects of short-term 28% and 100% oxygen on PaCO2 and peak expiratory flow rate in acute asthma: a randomized trial. *Chest* 2003;**124**(4):1312-7. [PUBMED: 14555560]

Roffe 2010 *{published data only}*

Roffe C, Sills S, Pountain SJ, Allen M. A randomized controlled trial of the effect of fixed-dose routine nocturnal oxygen supplementation on oxygen saturation in patients with acute stroke. *Journal of Stroke and Cerebrovascular Diseases* 2010;**19**(1):29-35. [PUBMED: 20123224]

Roffe 2017 *{published data only}*

Roffe C, Nevatte T, Sim J, Bishop J, Ives N, Ferdinand P, et al. Effect of routine low-dose oxygen supplementation on death and disability in adults with acute stroke: the stroke oxygen study randomized clinical trial. *JAMA* 2017;**318**(12):1125-35. [PUBMED: 28973619]

Sills 2003 *{published data only}*

Sills S, Halim M, Roffe C. A pilot study of routine nocturnal oxygen supplementation in patients with acute stroke. *Age and Ageing* 2003;**32**(Suppl 2):ii41.

Singhal 2005 *{published data only}*

Singhal AB, Benner T, Roccatagliata L, Koroshetz WJ, Schaefer PW, Lo EH, et al. A pilot study of normobaric oxygen therapy in acute ischemic stroke. *Stroke* 2005;**36**(4):797-802. [PUBMED: 15761201]

Singhal 2013 *{published data only}*

Singhal A. A phase IIb clinical trial of normobaric oxygen therapy (NBO) in acute ischemic stroke (AIS). *Neurology* 2013;**80**(Suppl 7):S02.001.

Stub 2014 *{published data only}*

Stub D, Smith K, Bernard S, Nehme Z, Stephenson M, Bray JE, et al. Air versus oxygen in S-segment-elevation myocardial infarction. *Circulation* 2015;**131**(24):2143-50. [PUBMED: 26002889]

Ukholkina 2005 *{published data only}*

Ukholkina GB, Kostianov IIu, Kuchkina NV, Grendo EP, Gofman IaB. Effect of oxygenotherapy used in combination with reperfusion in patients with acute myocardial infarction. *Kardiologiia* 2005;**45**(5):59. [PUBMED: 16007057]

Wu 2014 *{published data only}*

Wu J, Nevatte T, Roffe C. The stroke oxygen supplementation (S02S) study: comparison of postal and telephone responses of 12 months questionnaire follow up. *International Journal of Stroke* 2014;**9**(Suppl 4):37.

Young 2014 *{published data only}*

Young P, Bailey M, Bellomo R, Bernard S, Dicker B, Freebairn R, et al. HyperOxic Therapy OR NormOxic Therapy after out-of-hospital cardiac arrest (HOT OR NOT): a randomised controlled feasibility trial. *Resuscitation* 2014;**85**(12):1686-91. [PUBMED: 25261605]

Zughaft 2013 *{published data only}*

Zughaft D, Bhiladvala P, Van Dijkman A, Harnek J, Madsen Hardig B, Bjork J. The analgesic effect of oxygen during percutaneous coronary intervention (the OXYPAIN Trial). *Acute Cardiac Care* 2013;**15**(3):63-8. [PUBMED: 23957447]

References to studies awaiting assessment

ICU-ROX 2019 *{published data only}*

ICU-ROX Investigators and the Australian and New Zealand Intensive Care Society Clinical Trials Group, Mackle D, Bellomo R, Bailey M, Beasley R, Deane A, Eastwood G, Finfer S, Freebairn R, King V, Linke N, Litton E, McArthur C, McGuinness S, Panwar R, Young P. Conservative Oxygen Therapy during Mechanical Ventilation in the ICU. *New England Journal of Medicine* 2019:[Epub ahead of print]. [PUBMED: 31613432]

References to ongoing studies

NCT02321072 *{published data only}*

NCT02321072. Optimal oxygenation in the intensive care unit (O2-ICU). clinicaltrials.gov/ct2/show/NCT02321072 (first received 9 December 2014).

NCT02713451 *{published data only}*

NCT02713451. Liberal oxygenation versus conservative oxygenation in ARDS (LOCO2). clinicaltrials.gov/ct2/show/NCT02713451 (first received 18 March 2016).

NCT03141099 *{published data only}*

NCT03141099. Blood pressure and OXygenation targets after OHCA (BOX). clinicaltrials.gov/ct2/show/NCT03141099 (first received 4 May 2017).

NCT03174002 *{published data only}*

NCT03174002. Handling oxygenation targets in the intensive care unit (HOT-ICU). clinicaltrials.gov/ct2/show/NCT03174002 (first received 2 June 2017).

NCT03287466 *{published data only}*

Targeted OXYgen therapy in critical illness (TOXYC). clinicaltrials.gov/ct2/show/NCT03287466 (first received 19 September 2017).

Additional references

AARC 2002

Kallstrom TJ, American Association for Respiratory Care. AARC clinical practice guideline. Oxygen therapy for adults in the acute care facility - 2002 revision & update. *Respiratory Care* 2002;**47**(6):717-20. [PUBMED: 12078655]

ACTRN12613000505707

ACTRN12613000505707. Feasibility and safety of conservative versus liberal oxygen targets in the mechanically ventilated patients. anzctr.org.au/Trial/Registration/TrialReview.aspx?id=364185 (first received 5 May 2013).

ACTRN12615000957594

ACTRN12615000957594. A multicentre, randomised, single-blinded clinical trial comparing the effect of conservative oxygen therapy with standard care on ventilator-free days in mechanically ventilated adults in the intensive care unit.. https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=369229&isReview=true (first received 1 September 2015).

Adhikari 2010

Adhikari NK, Fowler RA, Bhagwanjee S, Rubenfeld GD. Critical care and the global burden of critical illness in adults. *Lancet* 2010;**376**(9749):1339-46. [PUBMED: 20934212]

Alba 2016

Alba AC, Alexander PE, Chang J, MacIsaac J, DeFry S, Guyatt GH. High statistical heterogeneity is more frequent in meta-analysis of continuous than binary outcomes. *Journal of Clinical Epidemiology* 2016;**70**:129-35. [PUBMED: 26386323]

ARC 2014

Australian Resuscitation Council. Guideline 11.6.1. Targeted oxygen therapy in adult advanced life support, 2014. resus.org.au/download/section_11/anzcor-guideline-11-6-1-targeted-oxygen-therapy-jan16.pdf (accessed 17 December 2015).

ARDS Definition Task Force 2012

ARDS Definition Task Force, Ranieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, Caldwell E, Fan E, et al. Acute respiratory distress syndrome: the Berlin Definition. *JAMA* 2012;**307**(23):2526-33. [PUBMED: 22797452]

ARDS Network 2000

Acute Respiratory Distress Syndrome Network. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. *New England Journal of Medicine* 2000;**342**(18):1301-8. [PUBMED: 10793162]

ATS 2005

American Thoracic Society, Infectious Diseases Society of America. Guidelines for the management of adults with hospital-acquired, ventilator-associated, and healthcare-associated pneumonia. *American Journal of Respiratory and Critical Care Medicine* 2005;**171**(4):388-416. [PUBMED: 15699079]

Bailey 2003

Bailey TC, Martin EL, Zhao L, Veldhuizen RA. High oxygen concentrations predispose mouse lungs to the deleterious effects of high stretch ventilation. *Journal of Applied Physiology* 2003;**94**(3):975-82. [PUBMED: 12571129]

Barbateskovic 2018

Barbateskovic M, Schjørring OL, Jakobsen JC, Meyhoff CS, Rasmussen BS, Perner A, et al. Oxygen supplementation for critically ill patients - a protocol for a systematic review. *Acta Anaesthesiologica Scandinavica* 2018;**62**(7):1020-30. [PUBMED: 29708586]

Bayes factor calculator 2014

Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C. Bayes factor calculator. www.ctu.dk/tools-and-links/bayes-factor-calculation.aspx (accessed 13 August 2019).

Begg 1994

Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994;**50**(4):1088-101. [PUBMED: 7786990]

Bellomo 2011

Bellomo R, Bailey M, Eastwood GM, Nichol A, Pilcher D, Hart GK. Arterial hyperoxia and in-hospital mortality after resuscitation from cardiac arrest. *Critical Care* 2011;**15**(2):R90. [PUBMED: 21385416]

Benoit 2002

Benoît Z, Wicky S, Fischer JF, Frascarolo P, Chapuis C, Spahn DR, et al. The effect of increased FIO(2) before tracheal extubation on postoperative atelectasis. *Anesthesia and Analgesia* 2002;**95**(6):1777-81. [PUBMED: 12456458]

Brenner 2012

Brenner M, Stein D, Hu P, Kufera J, Wooford M, Scalea T. Association between early hyperoxia and worse outcomes after traumatic brain injury. *Archives of Surgery* 2012;**147**(11):1042-6. [PUBMED: 22801994]

Brok 2008

Brok J, Thorlund K, Gluud C, Wetterslev J. Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses. *Journal of Clinical Epidemiology* 2008;**61**(8):763-9. [PUBMED: 18411040]

Brok 2009

Brok J, Thorlund K, Wetterslev J, Gluud C. Apparently conclusive meta-analyses may be inconclusive - trial sequential analysis adjustment of random error risk due to repetitive testing of accumulating data in apparently conclusive neonatal meta-analyses. *International Journal of Epidemiology* 2009;**38**(1):287-98. [PUBMED: 18824466]

Brower 2004

Brower RG, Lanken PN, MacIntyre N, Matthay MA, Morris A, Ancukiewicz M, et al. Higher versus lower positive end-expiratory pressures in patients with the acute respiratory distress syndrome. *New England Journal of Medicine* 2004;**351**(4):327-36. [PUBMED: 15269312]

Budinger 2013

Budinger GR, Mutlu GM. Balancing the risks and benefit of oxygen therapy in critically ill adults. *Chest* 2013;**143**(4):1151-62. [PUBMED: 23546490]

Cabello 2016

Cabello JB, Burls A, Emparanza JI, Bayliss S, Quinn T. Oxygen therapy for acute myocardial infarction. *Cochrane Database of Systematic Reviews* 2016, Issue 12. [DOI: 10.1002/14651858.CD007160.pub4]

Chow 2003

Chow CW, Herrera Abreu MT, Suzuki T, Downey GP. Oxidative stress and acute lung injury. *American Journal of Respiratory Cell and Molecular Biology* 2003;**29**(4):427-31. [PUBMED: 14500253]

Chu 2018

Chu DK, Kim LH, Young PJ, Zamiri N, Almenawer SA, Jaeschke R, et al. Mortality and morbidity in acutely ill adults treated with liberal versus conservative oxygen therapy (IOTA): a systematic review and meta-analysis. *Lancet* 2018;**391**(10131):1693-705. [PUBMED: 29726345]

Crapo 1986

Crapo JD. Morphologic changes in pulmonary oxygen toxicity. *Annual Review of Physiology* 1986;**48**:721-31. [PUBMED: 3518622]

Crapo 1999

Crapo RO, Jensen RL, Hegewald M, Tashkin DP. Arterial blood gas reference values for sea level and an altitude of 1,400 meters. *American Journal of Respiratory and Critical Care Medicine* 1999;**160**:1525-31. [PUBMED: 10556115]

Dahl 2015

Dahl RM, Grønlykke L, Haase N, Holst LB, Perner A, Wetterslev J, et al. Variability in targeted arterial oxygenation levels in patients with severe sepsis or septic shock. *Acta Anaesthesiologica Scandinavia* 2015;**59**(7):859-69. [PUBMED: 25914095]

Damiani 2014

Damiani E, Adrario E, Girardis M, Romano R, Pelaia P, Singer M, et al. Arterial hyperoxia and mortality in critically ill patients: a systematic review and meta-analysis. *Critical Care* 2014;**18**(6):711. [PUBMED: 25532567]

de Graaff 2011

de Graaff AE, Dongelmans DA, Binnekade JM, de Jonge E. Clinicians' response to hyperoxia in ventilated patients in a Dutch ICU depends on the level of FiO2. *Intensive Care Medicine* 2011;**37**(1):46-51. [PUBMED: 20878146]

de Jonge 2008

de Jonge E, Peelen L, Keijzers PJ, Joore H, de Lange D, van der Voort PH, et al. Association between administered oxygen, arterial partial oxygen pressure and mortality in mechanically ventilated intensive care unit patients. *Critical Care* 2008;**12**(6):R156. [PUBMED: 19077208]

Deeks 2010

Deeks JJ, Higgins JPT. Statistical algorithms in Review Manager 5. https://training.cochrane.org/handbook/statistical-methods-revman5 2010:1-11.

Dellinger 2013

Dellinger RP, Levy MM, Rhodes A, Annane D, Gerlach H, Opal SM, et al. Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock: 2012. *Critical Care Medicine* 2013;**41**(2):580-637. [PUBMED: 23353941]

DeMets 1987

DeMets DL. Methods for combining randomized clinical trials: strengths and limitations. *Statistics in Medicine* 1987;**6**(3):341-50. [PUBMED: 3616287]

DerSimonian 1986

DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986;**7**(3):177-88. [PUBMED: 3802833]

Donahoe 2011

Donahoe M. Acute respiratory distress syndrome: a clinical review. *Pulmonary Circulation* 2011;**1**(2):192-211. [PUBMED: 22034606]

Eastwood 2012

Eastwood G, Bellomo R, Bailey M, Taori G, Pilcher D, Young P, et al. Arterial oxygen tension and mortality in mechanically ventilated patients. *Intensive Care Medicine* 2012;**38**(1):91-8. [PUBMED: 22127482]

Egger 1997

Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;**315**(7109):629-34. [PUBMED: 9310563]

Esan 2010

Esan A, Hess DR, Raoof S, George L, Sessler CN. Severe hypoxemic respiratory failure: part 1 - ventilatory strategies. *Chest* 2010;**137**(5):1203-16. [PUBMED: 20442122]

Garattini 2016

Garattini S, Jakobsen JC, Wetterslev J, Bertelé V, Banzi R, Rath A, et al. Evidence-based clinical practice: overview of threats to the validity of evidence and how to minimise them. *European Journal of Internal Medicine* 2016;**32**:13-21. [PUBMED: 27160381]

Gilbert-Kawai 2014

Gilbert-Kawai ET, Mitchell K, Martin D, Carlisle J, Grocott MPW. Permissive hypoxaemia versus normoxaemia for mechanically ventilated critically ill patients. *Cochrane Database of Systematic Reviews* 2014, Issue 5. [DOI: 10.1002/14651858.CD009931.pub2]

Gluud 2011

Gluud C, Thorlund K, Engstrøm J, Wetterslev J, Brok J, Imberger G. User manual for Trial Sequential Analysis (TSA). www.ctu.dk/tsa/files/tsa_manual.pdf (accessed February 2016).

GRADEpro GDT [Computer program]

McMaster University (developed by Evidence Prime). GRADEpro GDT. Version accessed 25 March 2019. Hamilton (ON): McMaster University (developed by Evidence Prime), 2015.

Guyatt 2008

Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;**336**(7650):924-6. [PUBMED: 18436948]

Hafner 2015

Hafner S, Beloncle F, Koch A, Radermacher P, Asfar P. Hyperoxia in intensive care, emergency, and peri-operative medicine: Dr. Jekyll or Mr. Hyde? A 2015 update. *Annals of Intensive Care* 2015;**5**(1):42. [PUBMED: 26585328]

Harbord 2006

Harbord RM, Egger M, Sterne JA. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. *Statistics in Medicine* 2006;**25**(20):3443-57. [PUBMED: 16345038]

Helmerhorst 2015

Helmerhorst HJ, Roos-Blom MJ, van Westerloo DJ, de Jonge E. Association between arterial hyperoxia and outcome in subsets of critical illness: a systematic review, meta-analysis, and meta-regression of cohort studies. *Critical Care Medicine* 2015;**43**(7):1508-19. [PUBMED: 25855899]

Helmerhorst 2017a

Helmerhorst HJ, Arts DL, Schultz MJ, van der Voort PH, Abu-Hanna A, de Jonge E, et al. Metrics of arterial hyperoxia and associated outcomes in critical care. *Critical Care Medicine* 2017;**45**(2):187-95. [PUBMED: 27763912]

Helmerhorst 2017b

Helmerhorst HJF, Schouten LRA, Wagenaar GTM, Juffermans NP, Roelofs JJTH, Schultz MJ, et al. Hyperoxia provokes a time- and dose-dependent inflammatory response in mechanically ventilated mice, irrespective of tidal volumes. *Intensive Care Medicine Experimental* 2017;**5**(1):27. [PUBMED: 28550659]

Higgins 2002

Higgins JP, Spiegelhalter DJ. Being sceptical about meta-analyses: a Bayesian perspective on magnesium trials in myocardial infarction. *International Journal of Epidemiology* 2002;**31**(1):96-104. [PUBMED: 11914302]

Higgins 2003

Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**(7414):557-60. [PUBMED: 12958120]

Higgins 2011a

Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Higgins 2011b

Higgins JP, Whitehead A, Simmonds M. Sequential methods for random-effects meta-analysis. *Statistics in Medicine* 2011;**30**(9):903-21. [PUBMED: 21472757]

Hrobjartsson 2014

Hrobjartsson A, Emanuelsson F, Skou Thomsen AS, Hilden J, Brorson S. Bias due to lack of patient blinding in clinical trials. A systematic review of trials randomizing patients to blind and nonblind sub-studies. *International Journal of Epidemiology* 2014;**43**(4):1272-83. [PUBMED: 24881045]

ICH-GCP 1997

International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use (ICH) adopts consolidated guideline on good clinical practice in the conduct of clinical trials on medicinal products for human use. *International Digest of Health Legislation* 1997;**48**(2):231-4. [PUBMED: 11656783]

Imberger 2015

Imberger G, Gluud C, Boylan J, Wetterslev J. Systematic reviews of anesthesiologic interventions reported as statistically significant: problems with power, precision, and type 1 error protection. *Anesthesia and Analgesia* 2015;**121**(6):1611-22. [PUBMED: 26579662]

Imberger 2016

Imberger G, Thorlund K, Gluud C, Wetterslev J. False-positive findings in Cochrane meta-analyses with and without application of trial sequential analysis: an empirical review. *BMJ Open* 2016;**6**(8):e011890. [PUBMED: 27519923]

IRCT201212199647N2

IRCT201212199647N2. Effect of high percent oxygen therapy in compared without oxygen therapy on rehabilitation in the first 12 hours of admission in patients with stroke: a randomized clinical trial. en.irct.ir/trial/10208 (first received 3 November 2013).

Itagaki 2015

Itagaki T, Nakano Y, Okuda N, Izawa M, Onodera M, Imanaka H, et al. Hyperoxemia in mechanically ventilated, critically ill subjects: incidence and related factors. *Respiratory Care* 2015;**60**(3):335-40. [PUBMED: 25389354]

Jakobsen 2013

Jakobsen JC, Gluud C. The necessity of randomized clinical trials. *British Journal of Medicine and Medical Research* 2013;**3**(4):1453-68. [DOI: 10.9734/BJMMR/2013/3208]

Jakobsen 2014a

Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC Medical Research Methodology* 2014;**14**:120. [PUBMED: 25416419]

Jakobsen 2014b

Jakobsen JC, Gluud C, Winkel P, Lange T, Wetterslev J. The thresholds for statistical and clinical significance – a five-step procedure for evaluation of intervention effects in randomised clinical trials. *BMC Medical Research Methodology* 2014;**14**:34. [DOI: 10.1186/1471-2288-14-34]

Jakobsen 2016

Jakobsen JC, Wetterslev J, Lange T, Gluud C. Viewpoint: taking into account risks of random errors when analysing multiple outcomes in systematic reviews [editorial]. www.cochranelibrary.com/cdsr/doi/10.1002/14651858.ED000111/full 18 March 2016. [DOI: 10.1002/14651858.ED000111; PUBMED: 27030037]

Kahn 2010

Kahn JM, Benson NM, Appleby D, Carson SS, Iwashyna TJ. Long-term acute care hospital utilization after critical illness. *JAMA* 2010;**303**(22):2253-9. [PUBMED: 20530778]

Kallet 2013

Kallet RH, Matthay MA. Hyperoxic acute lung injury. *Respiratory Care* 2013;**58**(1):123-41. [PUBMED: 23271823]

Kenmure 1971

Kenmure AC, Beatson JM, Cameron AJ, Horton PW. Effects of oxygen on myocardial blood flow and metabolism. *Cardiovascular Research* 1971;**5**(4):483-9. [PUBMED: 5160452]

Kent 2011

Kent BD, Mitchell PD, McNicholas WT. Hypoxemia in patients with COPD: cause, effects, and disease progression. *International Journal of COPD* 2011;**6**:199-208. [PUBMED: 21660297]

Kilgannon 2010

Kilgannon JH, Jones AE, Shapiro NI, Angelos MG, Milcarek B, Hunter K, et al. Association between arterial hyperoxia following resuscitation from cardiac arrest and in-hospital mortality. *JAMA* 2010;**303**(21):2165-71. [PUBMED: 20516417]

Kraft 2018

Kraft F, Andel H, Gamper J, Markstaller K, Ullrich R, Klein KU. Incidence of hyperoxia and related in-hospital mortality in critically ill patients: a retrospective data analysis. *Acta Anaesthesiologica Scandinavica* 2018;**62**(3):347-56. [PUBMED: 29210062]

Kratz 2004

Kratz A, Ferraro M, Sluss PM, Lewandrowski KB. Case records of the Massachusetts General Hospital. Weekly clinicopathological exercises. Laboratory reference values. *New England Journal of Medicine* 2004;**351**(15):1548-63. [PUBMED: 15470219]

Kulinskaya 2014

Kulinskaya E, Wood E. Trial sequential methods for meta-analysis. *Research Synthesis Methods* 2014;**5**(3):212-20. [PUBMED: 26052847]

MAGIC 2002

Magnesium in Coronaries (MAGIC) Trial Investigators. Early administration of intravenous magnesium to high-risk patients with acute myocardial infarction in the Magnesium in Coronaries (MAGIC) Trial: a randomised controlled trial. *Lancet* 2002;**360**(9341):1189-96. [PUBMED: 12401244]

Mantel 1959

Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute* 1959;**22**(4):719-48. [PUBMED: 13655060]

Mascha 2015

Mascha EJ. Alpha, beta, meta: guidelines for assessing power and type I error in meta-analyses. *Anesthesia and Analgesia* 2015;**121**(6):1430-3. [DOI: 10.1213/ANE.0000000000000993; PUBMED: 26579648]

Metnitz 2009

Metnitz PG, Metnitz B, Moreno RP, Bauer P, Del Sorbo L, Hoermann C, et al. Epidemiology of mechanical ventilation: analysis of the SAPS 3 database. *Intensive Care Medicine* 2009;**35**(5):816-25. [PUBMED: 19288079]

Meyhoff 2012

Meyhoff CS, Jorgensen LN, Wetterslev J, Christensen KB, Rasmussen LS, PROXI Trial Group. Increased long-term mortality after a high perioperative inspiratory oxygen fraction during abdominal surgery: follow-up of a randomized clinical trial. *Anesthesia and Analgesia* 2012;**115**(4):849-54. [PUBMED: 22798533]

NCT01201291

NCT01201291. Impact of Inspired Oxygen Fraction on Outcome in Patients With Traumatic Brain Injury (BRAINOXY). clinicaltrials.gov/ct2/show/NCT01201291 (first received 14 September 2010).

NCT01319643

NCT01319643. Normal Oxygenation Versus Hyperoxia in the Intensive Care Unit (ICU) (OXYGEN-ICU). clinicaltrials.gov/ct2/show/NCT01319643 (first received 22 March 2011).

NCT01722422

NCT01722422. Hyperoxia and Hypertonic Saline in Septic Shock (Hyper2S). clinicaltrials.gov/ct2/show/study/NCT01722422 (first received 6 November 2012).

NCT02698917

NCT02698917. Carbon Dioxide, Oxygen and Mean Arterial Pressure After Cardiac Arrest and Resuscitation (COMACARE). https://clinicaltrials.gov/ct2/show/NCT02698917 (first received 4 March 2016).

O'Driscoll 2017

O'Driscoll BR, Howard LS, Earis J, Mak V, British Thoracic Society Emergency Oxygen Guideline Group, BTS Emergency Oxygen Guideline Development Group. BTS guideline for oxygen use in adults in healthcare and emergency settings. *Thorax* 2017;**72**(Suppl 1):ii1-90. [PUBMED: 28507176]

Pannu 2016

Pannu SR, Dziadzko MA, Gajic O. How much oxygen? Oxygen titration goals during mechanical ventilation. *American Journal of Respiratory and Critical Care Medicine* 2016;**193**(1):4-5. [DOI: 10.1164/rccm.201509-1810ED; PUBMED: 26720783]

Panwar 2013

Panwar R, Capellier G, Schmutz N, Davies A, Cooper DJ, Bailey M, et al. Current oxygenation practice in ventilated patients - an observational cohort study. *Anaesthesia and Intensive Care* 2013;**41**(4):505-14. [PUBMED: 23808511]

Panwar 2015

Panwar R, Hardie M, Bellomo R, Barrot L, Eastwood GM, Young PJ, et al. Conservative versus liberal oxygenation targets for mechanically ventilated patients. A pilot multicenter randomized controlled trial. *American Journal of Respiratory and Critical Care Medicine* 2016;**193**(1):43-51. [PUBMED: 26334785]

Petersson 2014

Petersson J, Glenny RW. Gas exchange and ventilation-perfusion relationships in the lung. *European Respiratory Journal* 2014;**44**(4):1023-41. [PUBMED: 25063240]

Pocock 2015

Pocock SJ, Clayton TC, Stone GW. Design of major randomized trials: part 3 of a 4-part series on statistics for clinical trials. *Journal of the American College of Cardiology* 2015;**66**(24):2757-66. [PUBMED: 26700838]

Pogue 1997

Pogue JM, Yusuf S. Cumulating evidence from randomized trials: utilizing sequential monitoring boundaries for cumulative meta-analysis. *Controlled Clinical Trials* 1997;**18**(6):580-93. [PUBMED: 9408720]

Rachmale 2012

Rachmale S, Li G, Wilson G, Malinchoc M, Gajic O. Practice of excessive F(IO(2)) and effect on pulmonary outcomes in mechanically ventilated patients with acute lung injury. *Respiratory Care* 2012;**57**(11):1887-93. [PUBMED: 22613692]

Raj 2013

Raj R, Bendel S, Reinikainen M, Kivisaari R, Siironen J, Lång M, et al. Hyperoxemia and long-term outcome after traumatic brain injury. *Critical Care* 2013;**17**(4):R177. [PUBMED: 23958227]

Raoof 2010

Raoof S, Goulet K, Esan A, Hess DR, Sessler CN. Severe hypoxemic respiratory failure: part 2 - nonventilatory strategies. *Chest* 2010;**137**(6):1437-48. [PUBMED: 20525656]

Rasmussen 2018

Rasmussen BS, Perner A, Wetterslev J, Meyhoff CS, Schjørring OL. Oxygenation targets in acutely ill patients: still a matter of debate. *Lancet* 2018;**392**(10163):2436-7. [PUBMED: 30527413]

Review Manager 2014 [Computer program]

Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager 5 (RevMan 5). Version 5.3. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Rodríguez-Roisin 2005

Rodríguez-Roisin R, Roca J. Mechanisms of hypoxemia. *Intensive Care Medicine* 2005;**31**(8):1017-9. [PUBMED: 16052273]

Rothen 1995a

Rothen HU, Sporre B, Engberg G, Wegenius G, Högman M, Hedenstierna G. Influence of gas composition on recurrence of atelectasis after a reexpansion maneuver during general anesthesia. *Anesthesiology* 1995;**82**(4):832-42. [PUBMED: 7717553]

Rothen 1995b

Rothen HU, Sporre B, Engberg G, Wegenius G, Reber A, Hedenstierna G. Prevention of atelectasis during general anaesthesia. *Lancet* 1995;**345**(8962):1387-91. [PUBMED: 7760608]

Roussos 2003

Roussos C, Koutsoukou A. Respiratory failure. *European Respiratory Journal* 2003;**47**:3S-14S. [PUBMED: 14621112]

Sacco 2013

Sacco RL, Kasner SE, Broderick JP, Caplan LR, Connors JJ, Culebras A, et al. An updated definition of stroke for the 21st century: a statement for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke* 2013;**44**(7):2064-89. [PUBMED: 23652265]

Savovic 2018

Savovic J, Turner RM, Mawdsley D, Jones HE, Beynon R, Higgins JPT, et al. Association between risk-of-bias assessments and results of randomized trials in Cochrane Reviews: the ROBES meta-epidemiologic study. *American Journal of Epidemiology* 2018;**187**(5):1113-22. [PUBMED: 29126260]

Schjørring 2018

Schjørring OL, Toft-Petersen AP, Kusk KH, Mouncey P, Sørensen EE, Berezowicz P, et al. Intensive care doctors' preferences for arterial oxygen tension levels in mechanically ventilated patients. *Acta Anaesthesiologica Scandinavica* 2018;**62**(10):1443-51. [PUBMED: 29926908]

Sepehrvand 2018

Sepehrvand N, James SK, Stub D, Khoshnood A, Ezekowitz JA, Hofmann R. Effects of supplemental oxygen therapy in patients with suspected acute myocardial infarction: a meta-analysis of randomised clinical trials. *Heart* 2018;**104**(20):1694-8. [PUBMED: 29599378]

Siemieniuk 2018

Siemieniuk RAC, Chu DK, Kim LH, Güell-Rous MR, Alhazzani W, Soccal PM, et al. Oxygen therapy for acutely ill medical patients: a clinical practice guideline. *BMJ* 2018;**363**:k4169. [PUBMED: 30355567]

Sinclair 2004

Sinclair SE, Altemeier WA, Matute-Bello G, Chi EY. Augmented lung injury due to interaction between hyperoxia and mechanical ventilation. *Critical Care Medicine* 2004;**32**(12):2496-501. [PUBMED: 15599157]

Sjöberg 2013

Sjöberg F, Singer M. The medical use of oxygen: a time for critical reappraisal. *Journal of Internal Medicine* 2013;**274**(6):505-28. [DOI: 10.1111/joim.12139; PUBMED: 24206183]

Suzuki 2013

Suzuki S, Eastwood GM, Peck L, Glassford NJ, Bellomo R. Current oxygen management in mechanically ventilated patients: a prospective observational cohort study. *Journal of Critical Care* 2013;**28**(5):647-54. [PUBMED: 23683560]

Tan 2014

Tan HL, Wijeweera O. Oxygen in critical care. *Trends in Anaesthesia and Critical Care* 2014;**4**:102-8. [DOI: 10.1016/j.tacc.2014.05.001]

Terkawi 2016

Terkawi AS, Mavridis D, Flood P, Wetterslev J, Terkawi RS, Bin Abdulhak AA, et al. Does ondansetron modify sympathectomy due to subarachnoid anesthesia?: meta-analysis, meta-regression, and trial sequential analysis. *Anesthesiology* 2016;**124**(4):846-69. [DOI: 10.1097/ALN.0000000000001039; PUBMED: 26835645]

Thorlund 2009

Thorlund K, Devereaux PJ, Wetterslev J, Guyatt G, Ioannidis JPA, Thabane L, et al. Can trial sequential monitoring boundaries reduce spurious inferences from meta-analyses?. *International Journal of Epidemiology* 2009;**38**(1):276-86. [PUBMED: 18824467]

Thygesen 2012

Thygesen K, Alpert JS, Jaffe AS, Simoons ML, Chaitman BR, White HD, et al. Third universal definition of myocardial infarction. *Circulation* 2012;**126**(16):2020-35. [PUBMED: 22923432]

TSA 2011 [Computer program]

Copenhagen Trial Unit. TSA - Trial Sequential Analysis. Version 0.9 Beta. Copenhagen: Copenhagen Trial Unit, 2011.

Turner 2013

Turner RM, Bird SM, Higgins JP. The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews. *PLOS ONE* 2013;**8**(3):e59202. [PUBMED: 23544056]

Wagner 1977

Wagner PD, Dantzker DR, Dueck R, Clausen JL, West JB. Ventilation-perfusion inequality in chronic obstructive pulmonary disease. *Journal of Clinical Investigation* 1977;**59**(2):203-16. [PUBMED: 833271]

Watson 2000

Watson NA, Beards SC, Altaf N, Kassner A, Jackson A. The effect of hyperoxia on cerebral blood flow: a study in healthy volunteers using magnetic resonance phase-contrast angiography. *European Journal of Anaesthesiology* 2000;**17**(3):152-9. [PUBMED: 10758463]

Wetterslev 2008

Wetterslev J, Thorlund K, Brok J, Gluud C. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. *Journal of Clinical Epidemiology* 2008;**61**(1):64-75. [PUBMED: 18083463]

Wetterslev 2009

Wetterslev J, Thorlund K, Brok J, Gluud C. Estimating required information size by quantifying diversity in random-effects model meta-analyses. *BMC Medical Research Methodology* 2009;**9**:86. [PUBMED: 20042080]

Wetterslev 2015

Wetterslev J, Meyhoff CS, Jørgensen LN, Gluud C, Lindschou J, Rasmussen LS. The effects of high perioperative inspiratory oxygen fraction for adult surgical patients. *Cochrane Database of Systematic Reviews* 2015, Issue 6. [DOI: 10.1002/14651858.CD008884.pub2]

Whitehead 2002

Whitehead T, Slutsky AS. The pulmonary physician in critical care · 7: ventilator induced lung injury. *Thorax* 2002;**57**(7):635-42. [PUBMED: 12096209]

Woods 2002

Woods KL, Abrams K. The importance of effect mechanism in the design and interpretation of clinical trials: the role of magnesium in acute myocardial infarction. *Progress*

in Cardiovascular Diseases 2002;**44**(4):267-74. [PUBMED: 12007082]

Wunsch 2010

Wunsch H, Linde-Zwirble WT, Angus DC, Hartman ME, Milbrandt EB, Kahn JM. The epidemiology of mechanical ventilation use in the United States. *Critical Care Medicine* 2010;**38**(10):1947-53. [PUBMED: 20639743]

You 2018

You J, Fan X, Bi X, Xian Y, Xie D, Fan M, et al. Association between arterial hyperoxia and mortality in critically ill patients: a systematic review and meta-analysis. *Journal of Critical Care* 2018;**47**:260-8. [PUBMED: 30077082]

Young 2012

Young P, Beasley R, Bailey M, Bellomo R, Eastwood GM, Nichol A, et al. The association between early arterial oxygenation and mortality in ventilated patients with acute ischaemic stroke. *Critical Care and Resuscitation* 2012;**14**(1):14-9. [PUBMED: 22404056]

Zaher 2007

Zaher TE, Miller EJ, Morrow DM, Javdan M, Mantell LL. Hyperoxia-induced signal transduction pathways in pulmonary epithelial cells. *Free Radical Biology and Medicine* 2007;**42**(7):897-908. [PUBMED: 17349918]

Zhang 2016

Zhang Z, Ji X. Quadratic function between arterial partial oxygen pressure and mortality risk in sepsis patients: an interaction with simplified acute physiology score. *Scientific Reports* 2016;**6**:35133. [PUBMED: 27734905]

References to other published versions of this review

Barbateskovic 2017

Barbateskovic M, Schjørring OL, Jakobsen JC, Meyhoff CS, Dahl RM, Rasmussen BS, et al. Higher versus lower inspiratory oxygen fraction or targets of arterial oxygenation for adult intensive care patients. *Cochrane Database of Systematic Reviews* 2017, Issue 4. [DOI: 10.1002/14651858.CD012631]

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Asfar 2017

Methods	RCT 2-by-2 factorial trial randomizing to 4 groups. 2 groups were included in our analysis.
Participants	**Sample size**: 442 randomized (219 experimental, 223 control) **Sex (male)**: experimental 63%, control 65% **Age (mean)**: experimental 67.8, control 66.3 **Country:** France **Setting**: multidisciplinary ICU **Disease severity score:** SAPS III median 71 **Inclusion criteria** 1. Patients aged 18 years and older if they were mechanically ventilated and exhibited septic shock refractory to fluid resuscitation as defined by an absence of response to 20 mL/kg of crystalloids or colloids and requiring vasopressor (norepinephrine or epinephrine, at a minimum infusion rate of 0.1 µg/kg per min); they also had to have been assessed within 6 hours after the initiation of vasopressors. Septic shock was defined by the presence of 2 or more diagnostic criteria of systemic inflammatory response syndrome, proven or suspected infection, and sudden dysfunction of at least 1 organ. **Exclusion criteria** 1. Severe hypoxaemia defined as PaO_2: FiO_2 ratio of less than 100 mmHg for a minimum positive end-expiratory pressure of 5 cm H_2O 2. Plasma sodium concentration of less than 130 mmol/L or more than 145 mmol/L 3. Intracranial hypertension 4. Patient admitted for cardiac arrest 5. Overt cardiac failure

Asfar 2017 *(Continued)*

	6. Under legal guardianship 7. No affiliation with the French healthcare system 8. Pregnancy 9. Recent participation in another biomedical study or another interventional study with mortality as the primary endpoint 10. An investigator's decision not to resuscitate
Interventions	**Experimental**: hyperoxia group (mechanical ventilation with FiO_2 of 1.0 for 24 hours after inclusion; thereafter FiO_2 as in the normoxia group). Categorized by us as using a high target in the experimental group. **Control**: target SaO_2 of 88% to 95% using mechanical ventilation **Co-intervention**: not specified **Duration**: 24 hours
Outcomes	**Primary outcome** 1. Death from any cause at day 28 after inclusion **Secondary outcomes** 1. 90-day mortality 2. Daily SOFA from inclusion to day 7 3. 19 days alive and free from organ dysfunction at day 28 4. Length of stay in the ICU 5. Alive at day 28 without organ support was defined as days alive without vasopressor infusion, mechanical ventilation, or renal replacement treatment 6. Safety data (as specified in protocol (NCT01722422) **Outcomes not prespecified** 1. Participants with at least 1 serious adverse event 2. Chest radiograph scores 3. Atelectasis 4. Pneumothorax 5. Ventricular arrhythmias 6. Mesenteric ischaemia 7. Digital ischaemia 8. ICU-acquired weakness 9. Participants with ≥ 1 nosocomial infection during ICU stay 10. Participants with ≥ 1 nosocomial pneumonia during ICU stay
Notes	Email sent to Dr Asfar 5 December 2018 and reply was received. The trial was funded by public grants (the French ministry of health).

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated randomization list stratified by site and presence or absence of ARDS using permuted blocks of random sizes (nQuery Advisor 6.0)
Allocation concealment (selection bias)	Low risk	The pharmacists assigned a random number to each therapeutic package. The attribution of a given therapeutic package to a participant in accordance to

Asfar 2017 *(Continued)*

		the randomization list was done with a web-based secured randomization system (Clinsight software).
Blinding of participants and personnel (performance bias) All outcomes	High risk	Unblinded
Blinding of outcome assessment (detection bias) All outcomes	High risk	Unblinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	2.7% in the experimental group and 0.9% in the control were excluded from analysis.
Selective reporting (reporting bias)	Low risk	The trial was registered prior to randomisation (NCT01722422), and all prespecified outcomes were reported on.
Other bias	High risk	Early stopping bias: the trial was stopped after a pre-planned interim analysis, criteria for stopping not specified

Girardis 2016

Methods	RCT
Participants	**Sample size**: 480 (experimental 244, control 236) **Sex (male %)**: experimental 57%, control 56% **Age (median)**: experimental 65, control 63 **Country**: Italy **Setting**: multidisciplinary ICU **Disease severity score**: SAPS II score median 38 **Inclusion criteria** 1. All patients aged 18 years or older and admitted to the ICU with an expected length of stay of 72 hours or longer **Exclusion criteria** 1. Age younger than 18 years 2. Pregnancy 3. ICU readmission 4. A decision to withhold life-sustaining treatment 5. Immunosuppression or neutropenia 6. Enrolment in another study 7. Patients with acute decompensation of COPD and ARDS with a PaO_2:FiO_2 ratio of less than 150
Interventions	**Experimental**: oxygen therapy was administered according to standard ICU practice; FiO_2 of at least 0.4, allowing PaO_2 values up to 150 mmHg and an SpO_2 between 97% and 100%. If the SpO_2 decreased below 95% to 97%, the FiO_2 was increased to reach the target value of SpO_2. Participants received FiO_2 of 1.0 during intubation, airway suction, or hospital transfer.

Girardis 2016 *(Continued)*

	Categorized by us as using a high target in the experimental group. **Control**: oxygen therapy was administered at the lowest possible FiO_2 to maintain the PaO_2 between 70 and 100 mmHg or SpO_2 values between 94% and 98%. FiO_2 was gradually reduced or oxygen supplementation discontinued whenever the PaO_2 or SpO_2 exceeded 100 mmHg or 98%. Supplemental oxygen was administered only if SpO_2 decreased below 94%. Categorized by us as using a high target in the control group. **Co-intervention**: not specified **Duration**: until ICU discharge
Outcomes	1. ICU mortality 2. New-onset respiratory, cardiovascular, liver, and renal failure (defined as a SOFA score ≥ 3 for the corresponding organ) occurring 48 hours or more after ICU admission 3. Need for reoperation in surgical patients 4. Bloodstream, respiratory, and surgical site infections (according to Centers for Disease Control and Prevention definitions). Only microbiologically documented bloodstream and respiratory tract infections were considered. **Secondary outcomes not prespecified** 1. Hospital mortality 2. Ventilation-free hours during the ICU stay
Notes	Email sent 6 December 2018 to Dr Girardis and reply was received. The trial was funded by public grants.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated random numbers
Allocation concealment (selection bias)	Low risk	The randomization sequence was concealed from the researchers by use of sequentially numbered, closed, opaque envelopes that were opened after patient study inclusion.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Open-label
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Not described; however, blinding of outcome assessment was clarified by email
Incomplete outcome data (attrition bias) All outcomes	Low risk	Results from intention-to-treat analyses are provided in the supplementary. 2 participants withdrew consent, randomization groups for these 2 participants were not reported, thus they could not be included in the sensitivity analysis on losses to follow-up. Outcome respiratory failure: 18 in experimental and 15 in control group were lost to follow-up

Girardis 2016 *(Continued)*

Selective reporting (reporting bias)	High risk	The trial was registered retrospectively (NCT01319643)
Other bias	High risk	Early stopping bias: the trial was stopped after an interim analysis that was not pre-planned

Gomersall 2002

Methods	RCT
Participants	**Sample size**: 36 (experimental 19, control 17) **Sex (male %)**: experimental 82%, control 76% **Age (mean)**: experimental 68, control 69 **Country**: Hong Kong **Setting**: multidisciplinary ICU **Disease severity score:** not reported **Inclusion criteria** 1. Patients admitted with a clinical diagnosis of an acute exacerbation of COPD and a PaO_2 < 6.6 kPa (50 mmHg), and $PaCO_2$ > 6.6 kPa (50 mmHg) on air. **Exclusion criteria** 1. Chest radiologic signs of pulmonary oedema, lung cancer, pneumothorax, or pneumonia 2. If the patient already met study criteria for mechanical ventilation 3. Mechanical ventilation for respiratory failure twice in the preceding 6 months 4. Inability to walk more than 20 yards on flat ground 5. Co-existing terminal disease
Interventions	Oxygen therapy was provided via a Venturi-type mask and adjusted according to the results of arterial blood samples with the aim of reaching the desired target oxygen tension within 1 hour of trial entry. **Experimental**: target PaO_2 above 9.0 kPa (70 mmHg) (categorized by us as using a low target in the experimental group) **Control:** target PaO_2 of > 6.6 kPa (50 mmHg) (categorized by us as using a low target in the control group) **Co-intervention**: participants in the low-oxygen tension group also received doxapram if they developed an acidosis with pH < 7.2, whereas those in the high-oxygen tension group received doxapram if they developed symptomatic acidosis. Bronchodilator, steroid, and antibiotic therapy was standardized. **Duration**: treatment protocols, including oxygen therapy, were continued after discharge from the ICU until oxygen therapy was no longer considered necessary
Outcomes	1. Need for mechanical ventilation 2. Duration of hospital stay 3. Cardiac arrhythmia 4. Mortality 5. Coma

Gomersall 2002 *(Continued)*

Notes	Email sent to Dr Gomersall 6 December 2018 but no reply was received. The trial was funded by public grants.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated numbers
Allocation concealment (selection bias)	Low risk	Unmarked, sealed envelopes
Blinding of participants and personnel (performance bias) All outcomes	High risk	Only participants were blinded.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	High risk	2/19 (11%) of participants in the experimental group were excluded from analysis due to protocol violation.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	High risk	Doxapram co-intervention differed between groups.

Ishii 2018

Methods	RCT
Participants	**Sample size**: 44 (experimental 21, control 23) **Sex**: not specified **Age**: not specified **Country**: Japen **Setting**: surgical ICU **Disease severity score:** not reported **Inclusion criteria**: mechanically ventilated patients admitted to surgical ICU for more than 12 hours **Exclusion criteria**: not specified
Interventions	**Experimental**: FiO_2 of 1.0 using high-flow nasal cannula. Categorized by us as using a high target in the experimental group **Control**: expected FiO_2 to achieve a PaO_2 of 100 mmHg (13.3 kPa) using high-flow nasal cannula

Ishii 2018 *(Continued)*

	The interventions are 'non-invasive', as they are initiated after extubation (of the mechanical ventilated), whereas after oxygen they are administered via high-flow nasal cannula. Categorized by us as using a low target in the control group **Co-intervention**: not specified **Duration**: 1 hour
Outcomes	1. Atelectasis
Notes	Email sent 6 December 2018 to Dr Ishii but no reply was received. It was unclear how the trial was funded.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	It was stated that the trial was randomized, but method of sequence generation not described.
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The trial was described as double-blinded, but it was unclear who was blinded and how blinding was maintained.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Radiologist was blinded.
Incomplete outcome data (attrition bias) All outcomes	High risk	14% were lost to follow-up; randomization groups were not reported.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Low risk	The trial appeared to be free of other issues that could put it at risk of bias.

Jakkula 2018

Methods	RCT with a 2-by-3 factorial design. We only extracted data from the normoxia and moderate-hyperoxia groups.
Participants	**Sample size**: 123 (experimental 60, control 63) **Sex (male %)**: experimental 48%, control 50% **Age**: experimental 60, control 59 **Country**: Finland **Setting**: adults admitted to the ICU after OHCA **Disease severity score:** APACHE II score median 28

Jakkula 2018 *(Continued)*

Inclusion criteria

1. Adults resuscitated from witnessed OHCA with VF or VT as the initial rhythm. In addition, all of the following inclusion criteria had to be met:
   a. ROSC 10 to 45 minutes from the onset of cardiac arrest;
   b. confirmed or suspected cardiac origin of the arrest;
   c. mechanical ventilation upon ICU arrival;
   d. markedly impaired level of consciousness defined as no response to verbal commands and GCS motor score < 5 (withdrawal to painful stimuli at best);
   e. deferred consent from next of kin possible or likely; and
   f. active intensive care and TTM initiated.

Exclusion criteria

1. Adults with confirmed or suspected acute or pre-existing intracranial pathology or suspicion of increased intracranial pressure, or both
2. Adults with severe oxygenation failure defined as PaO_2/FiO_2 < 100 mmHg upon arrival to ICU and no improvement in oxygenation after adding sufficient PEEP level
3. Severe COPD
4. Age < 18 or > 80 years
5. Pregnancy

Interventions	**Experimental**: target $PaO_2$of 20 to 25 kPa (150 to 187.5 mmHg). Categorized by us as using a high target in the experimental group **Control**: target PaO_2 of 10 to 15 kPa (75 to 112.5 mmHg) or target SpO_2 of 95% to 98%. Categorized by us as using a high target in the control group **Co-intervention**: all adults received TTM at 33 °C or 36 °C and were sedated according to the treating clinicians' instructions. All adults received standard care, monitoring and assessments based on the protocol of the ICU, including direct blood pressure monitoring via an arterial catheter. **Duration**: 36 hours
Outcomes	**Primary outcome** 1. NSE serum concentration at 48 hours after cardiac arrest **Secondary outcomes** 1. NSE serum concentration at 24 and 72 hours after cardiac arrest 2. S100 protein serum concentration at 24, 48, and 72 hours after cardiac arrest 3. TnT concentration at 24, 48, and 72 hours after cardiac arrest 4. Results of NIRS monitoring during the first 48 hours after admission to the ICU 5. Results of continuous EEG monitoring for 48 hours after arrival at the ICU and a statement of the findings by an experienced senior neurologist or neurophysiologist 6. CPC at 6 months after cardiac arrest 7. Total duration of intensive care 8. Total duration of mechanical ventilation 9. Length of hospital stay 10. Discharge destination 11. Vital status at hospital discharge (dead or alive) **Feasibility outcomes** 1. Difference in $PaCO_2$ between groups targeting low to normal (4.5 to 4.7 kPa) and high to normal (5.8 to 6.0 kPa) $PaCO_2$

Jakkula 2018 *(Continued)*

2. Difference in PaO_2 between groups targeting low to normal (10 to 15 kPa) and high to normal (20 to 25 kPa) PaO_2
3. Difference in MAP between groups targeting low to normal (65 to 75 mmHg) and high to normal (80 to 100 mmHg) MAP
4. Distribution of values for primary and secondary outcomes
5. Randomized or screened participant ratio
6. Consent rate
7. Data completion rate
8. Recruitment duration

Notes	Email sent 6 December 2018 to Dr Jakkula but no reply was received. The trial was funded by public and private funds. The funding bodies had no input regarding the design, management, or reporting of the trial.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated random numbers
Allocation concealment (selection bias)	Low risk	Web-based system
Blinding of participants and personnel (performance bias) All outcomes	High risk	The treating personnel were not blinded to treatment targets.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	The neurophysiologist analysing the EEG results and the neurologist evaluating the neurologic recovery of the participants were blinded to the study group allocations.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Less than 5% were lost to follow-up.
Selective reporting (reporting bias)	Low risk	The trial was registered prior to randomization (NCT02698917).
Other bias	Low risk	The trial appeared to be free of other issues that could put it at risk of bias.

Lång 2018

Methods	RCT
Participants	**Sample size**: 65 (experimental 38, control 27) **Sex (male)**: experimental 82%, control 85% **Age:** experimental 45, control 43 **Country**: Finland **Setting**: mechanically ventilated adults with traumatic brain disease admitted to the ICU

Lång 2018 *(Continued)*

	Disease severity score: APACHE II score median 22 **Inclusion criteria** 1. Isolated non-penetrating TBI or adults with multiple trauma with TBI with GCS 8 or less (inclusive), expected need for intubation and mechanical ventilation > 24 hours 2. Recruitment within 18 hours after admission to ICU 3. Time from TBI < 36 hours 4. Informed consent from next of kin **Exclusion criteria** 1. Age < 18 or > 65 years 2. Anticipated brain death in 12 hours or otherwise moribund adults expected to die in 24 hours 3. Expected need for mechanical ventilation < 24 hours 4. Insufficient oxygenation assessed by a clinician 5. Adults with multiple trauma with brain injury and severe abdominal, thoracic, or pelvic injury possibly affecting oxygenation 6. No consent 7. Insufficient oxygenation with the treatment modality of the lower oxygenation group (PaO_2 < 13 kPa or SpO_2 < 95% with FiO_2 of 0.40 and PEEP of 10) 8. Oxygenation failure probable during ICU care 9. Penetrating TBI 10. Suspected pregnancy (perform urinary or serological pregnancy test if suspected)
Interventions	**Experimental:** FiO_2 of 0.70. Categorized by us as using a high target in the experimental group **Control**: FiO_2 of 0.40. Categorized by us as using a low target in the control group **Co-intervention**: not specified **Duration**: maximum 14 days
Outcomes	1. Laboratory markers during the first 3 days 2. Pulmonary function (PaO_2/FiO_2 ratio, ARDS, atelectasis, pneumonia) 3. Length of mechanical ventilation 4. Length of ICU stay 5. Length of hospital stay 6. Death 7. Extended Glasgow Outcome Scale
Notes	Email sent 6 December 2018 to Dr Lång and a reply was received. It was unclear how the trial was funded. According to protocol, the trial was supported by Kuopio University Hospital.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomized, but the method of sequence generation was not described
Allocation concealment (selection bias)	Low risk	Sealed, opaque envelopes

Lång 2018 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	High risk	Open-label
Blinding of outcome assessment (detection bias) All outcomes	High risk	Only the neurologist assessing the neurological outcomes was blinded.
Incomplete outcome data (attrition bias) All outcomes	High risk	8% were lost to follow-up, and allocation groups were not specified in the publication. The number of participants lost to follow-up in each group was clarified by email.
Selective reporting (reporting bias)	Low risk	The trial was registered prior to randomization (NCT01201291), however quality of life is not reported; however trial authors are planning to publish these results.
Other bias	High risk	Unplanned trial stop

Mazdeh 2015

Methods	RCT
Participants	**Sample size**: 51 (experimental 26, control 25) **Sex (male %)**: experimental 54%, control 56% **Age**: not specified **Country**: Iran **Setting**: adults with stroke initially referred to the Department of Neurology, but admitted to the ICU **Disease severity score:** not reported **Inclusion criteria** 1. Age between 40 and 70 years 2. GCS > 12 and adults with isolated brain damage and intact airway control 3. Ischaemic and haemorrhagic stroke with no need for surgical intervention 4. Less than 12 hours have passed since the accident 5. NIHSS square between 7 and 9 Quote: "Participants were selected from adults referred to the Department of Neurology of Farshchian Hospital, an affiliated hospital of Hamadan University of Medical Sciences. The participants were admitted to the ICU and monitored by nurses." Due to participants being transferred from the Department of Neurology to the ICU to be monitored, we do not regard these adults as typical adults admitted to the ICU. **Exclusion criteria** 1. Adults under 40 and older than 70 years 2. Adults with diabetes mellitus and ischaemic heart disease, renal failure, acute pulmonary oedema, history of massive myocardial infarction, and heart failure 3. Adults who need intubation on arrival to the hospital 4. Adults with a baseline blood pressure of less than 90/60, or hypoxia 5. Adults requiring surgical intervention (i.e. acute subdural haematoma and cerebral haemorrhage)

Mazdeh 2015 *(Continued)*

	6. Adults with blood pressure greater than 170/90 in the first 12 hours of the incident 7. Adults with successful CPR within 12 hours 8. History of previous stroke or unconsciousness resulting in the need for intubation and mechanical ventilation 9. Death or lost to follow-up 10. Adults in the control group for whom oxygen therapy was inevitable
Interventions	**Experimental**: FiO_2 of 0.5 - oxygen therapy with Venturi mask (categorized by us as using a low target in the experimental group) **Control**: no supplemental oxygen was administered (categorized by us as using a low target in the control group) **Co-intervention**: routine medication (as stated in protocol) **Duration**: oxygen administration was given in the first 12 hours of admission
Outcomes	1. Good recovery and lower number of complications in the first day of admission, before discharge, and 6 months after discharge using ranking scale and Barthel Index (as stated in protocol) 2. Outcome not prespecified: mortality
Notes	Email sent 6 December 2018 to Dr Seifirad, who forwarded the email on to Dr Mazdeh, however no reply was received. The trial was funded by a public hospital (Vice Chancellor of Research and Technology, Hamadan Medical University). Overall poor reporting quality.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomized, but the method of sequence generation was not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	High risk	Unblinded
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Not described
Incomplete outcome data (attrition bias) All outcomes	High risk	1 out of 52 (2%) randomized participants was lost to follow-up, and this was not described in the manuscript. It is not stated to which group this person was allocated. Participants in the control group for whom oxygen therapy was inevitable were excluded.
Selective reporting (reporting bias)	High risk	We judged the trial to be registered retrospectively (IRCT201212199647N2). It was registered 3 November 2013 and submitted to journal 30 December 2013.
Other bias	Low risk	The trial appeared to be free of other issues that could put it at risk of bias.

Panwar 2016

Methods	RCT
Participants	**Sample size**: 104 (experimental 51, control 53 (1 lost to follow-up)) **Sex (male %)**: experimental 65%, control 62% **Age**: experimental 62, control 62 **Country**: Australia, New Zealand, and France **Setting**: mechanically ventilated adults admitted to a multidisciplinary ICU **Disease severity score:** APACHE III score median 80 (control) and 70 (experimental) **Inclusion criteria** 1. People admitted to the ICU 2. Aged ≥ 18 years 3. Receiving invasive MV for < 24 hours, and their treating clinician expected MV to continue for at least the next 24 hours The reason for the inclusion criterion of receiving invasive MV for < 24 hours was to ensure that participants who would be assigned to the conservative oxygen group were not exposed to standard liberal oxygen therapy for prolonged periods prior to randomization. **Exclusion criteria** 1. Known pregnancy 2. Imminent risk of death 3. If the treating clinician lacked equipoise for the patient to be enrolled in this trial The exclusion criterion "lacked equipoise" included those clinical situations where the most appropriate approach (conservative versus liberal) to oxygen therapy is well established. For example, in hypercapnic patients with chronic respiratory failure or exacerbation of COPD, there is level I evidence supporting a conservative approach to oxygen therapy (1), and in patients with carbon monoxide poisoning or necrotizing fasciitis a liberal approach is preferred. However, amongst patients who had COPD listed as 1 of the prior comorbid conditions, the treating clinicians could permit enrolment of those adults who were admitted for reasons unrelated to COPD.
Interventions	**Experimental**: SpO_2 target ≥ 96%. Categorized by us as using a high target in the experimental group **Control**: target SpO_2 of 88% to 92%. When FiO_2 requirement was < 0.50, an SpO_2 of 90% to 92% was recommended, and when FiO_2 requirement was ≥ 0.50, an SpO_2 of 88% to 90% was recommended. Categorized by us as using a low target in the control group **Co-intervention**: participating sites were requested to adhere to best practice guidelines in relation to other potentially confounding co-interventions such as adjustment of tidal volume, PEEP, fluid management, blood transfusion, muscle relaxation, sedation interruption, ventilator weaning, nutrition, use of steroids, early mobilization, and physiotherapy. **Duration**: entire duration of mechanical ventilation
Outcomes	**Primary outcomes** 1. Proportion of time spent in the assigned SpO_2 range in each arm 2. Area under the curve for PaO_2, FiO_2, and SpO_2 on day 0 to day 7 in each arm **Secondary outcomes** 1. Incidence of circulation-related events

Panwar 2016 *(Continued)*

2. Incidence of respiration-related events
3. Incidence of acute kidney injury
4. Incidence of outcomes related to other organ systems
5. Time to successful extubation (alive and extubated for > 48 hours)
6. MV-free days
7. ICU mortality
8. Hospital mortality
9. All-cause mortality

Notes	Email sent to Dr Panwar 5 December 2018. Reminder sent 10 December 2018; reply was received. The trial was supported by public grants.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated randomization list
Allocation concealment (selection bias)	Low risk	Opaque, sealed envelopes
Blinding of participants and personnel (performance bias) All outcomes	High risk	Participants were unaware of their assigned group, but blinding of treating clinicians was not considered feasible.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Not described; however, Dr Panwar clarified in an email that outcome assessment was not blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	Only 1 (1/104) participant was lost to follow-up.
Selective reporting (reporting bias)	Low risk	A study protocol was registered prior to randomization (ACTRN12613000505707), and all outcomes were reported on.
Other bias	Low risk	The trial appeared to be free of other issues that could put it at risk of bias.

Taher 2016

Methods	RCT
Participants	**Sample size**: 68 (experimental 34, control 34) **Sex (male %)**: experimental 74%, control 68% **Age**: experimental 40, control 46 **Country**: Iran **Setting**: adults with traumatic brain injury initially referred to the emergency department, but who were admitted to the ICU **Disease severity score:** GCS score mean 7.4

Taher 2016 *(Continued)*

	Inclusion criteria 1. Age between 18 and 65 years 2. Less than 6 hours passed since the accident; haemodynamic stability; and GCS between 3 and 8 **Exclusion criteria** 1. Pregnancy 2. People under 18 or older than 65 years 3. GCS under 3 or more than 8 4. People with chronic disease such as diabetes mellitus, ischaemic heart disease, renal failure, acute pulmonary oedema, history of massive myocardial infarction, and heart failure 5. People with a baseline blood pressure of less than 90/60 6. People with successful CPR 7. Death or loss to follow-up Participants in the control group for whom oxygen therapy was inevitable were also excluded from this study.
Interventions	**Experimental**: FiO_2 of 80% oxygen by mechanical ventilator in the first 6 hours after the traumatic accident. Categorized by us as using a high target in the experimental group **Control**: FiO_2 of 0.5 using mechanical ventilator in the first 6 hours after the traumatic accident. Categorized by us as using a low target in the control group **Co-intervention**: not specified **Duration**: 6 hours
Outcomes	1. Glasgow Coma Scale 2. Barthel Index 3. mRS neurologic disability scoring system at the time of discharge from hospital and at 6-month follow-up
Notes	No relevant outcomes reported. Participants who died were excluded (from analyses). Email sent 6 December 2018 to Dr Pilehvari but no reply was received. The trial was funded by public funds.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomized, but sequence generation was not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	The trial was described as double-blind; however, it was unclear who was blinded and how blinding was maintained.
Blinding of outcome assessment (detection bias)	Unclear risk	Not described

Taher 2016 *(Continued)*

All outcomes		
Incomplete outcome data (attrition bias) All outcomes	High risk	Participants who died or were lost to follow-up were excluded.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Low risk	The trial appeared to be free of other issues that could put it at risk of bias.

Young 2017

Methods	RCT
Participants	**Sample size**: 100 (experimental 51, control 49 (48 analysed)) **Sex (male %)**: experimental 67%, control 65% **Age**: experimental 60, control 61 **Country**: New Zealand **Setting**: mechanically ventilated adults admitted to a multidisciplinary ICU **Disease severity score:** APACHE II score median 22.1 **Inclusion criteria** 1. People at least 18 years of age who require invasive mechanical ventilation in the ICU and are expected to be receiving mechanical ventilation beyond the next calendar day **Exclusion criteria** 1. Greater than 2 hours of invasive mechanical ventilation or non-invasive ventilation, or both, in an ICU during this hospital admission (includes time ventilated in another hospital's ICU) 2. In the view of the treating clinician, hyperoxia is clinically indicated for reasons including (but not limited to) carbon monoxide poisoning or a requirement for hyperbaric oxygen therapy 3. In the view of the treating clinician, avoidance of hyperoxia is clinically indicated for reasons including (but not limited to) COPD, paraquat poisoning, previous exposure to bleomycin, or chronic hypercapnic respiratory failure 4. Pregnancy 5. Death is deemed to be inevitable as a result of the current acute illness, and either the treating clinician, the participant, or the substitute decision-maker is not committed to full active treatment 6. Adults with a life expectancy of less than 90 days due to a chronic or underlying medical condition 7. Admitted following a drug overdose (including alcohol intoxication) 8. Long-term dependence on invasive ventilation prior to this acute illness 9. Confirmed or suspected diagnosis of any of the following: Guillain-Barré syndrome, cervical cord injury above C5, muscular dystrophy, or motor neuron disease 10. Enrolment not considered to be in the patient's best interest 11. Enrolled in any other trial of targeted oxygen therapy 12. Previously enrolled in the ICU-ROX study
Interventions	**Experimental**: no specific measures taken to avoid high FiO_2 or SpO_2, FiO_2< 0.30 discouraged (thus we could not categorize the experimental group as using either a low or high target). Participants assigned to the 'higher group' received 'standard care' both whilst ventilated and after extubation with no specific measures taken to avoid high FiO_2 or high SpO_2. The use of upper alarm limits for SpO_2 in

Young 2017 *(Continued)*

	the higher group was prohibited, as upper alarm limits for SpO_2 were not used as part of standard care. The lower limit alarm for SpO_2 was set at 90% (or lower if clinically appropriate). If the PaO_2 or SaO_2 was lower than the acceptable limit, inspired oxygen might be increased if clinically appropriate, irrespective of the SpO_2 reading. The use of an FiO_2 of less than 0.3 whilst ventilated was discouraged. **Control**: target SaO_2/SpO_2 91% to 96%. When a participant was allocated to conservative oxygen therapy, the inspired oxygen concentration was decreased to room air as rapidly as possible provided that the SpO_2 measured by peripheral pulse oximetry was greater than the acceptable lower limit. SpO_2 levels of greater than 96% were strictly avoided, and an upper SpO_2 alarm limit of 97% applied whenever supplemental oxygen was administered in the ICU to minimize the risk of hyperoxaemia. After extubation, in the conservative oxygen group, the upper monitored alarm limit of acceptable SpO_2 of 97% was applied whenever supplemental oxygen was being administered. In the event that the SpO_2 exceeded the acceptable upper limit, downward titration of supplemental oxygen was undertaken as a high priority and supplemental oxygen was discontinued as soon possible. The lower limit alarm for SpO_2 was set at 90% (or lower if clinically appropriate). If the PaO_2 or SaO_2 was lower than the acceptable limit, inspired oxygen might be increased if clinically appropriate, irrespective of the SpO_2 reading. Categorized by us as using a low target in the control group **Co-intervention**: there were no restrictions on concomitant treatments provided to participants. If an increase in FiO_2 for procedures performed in the ICU included (but were not limited to) bronchoscopy, suctioning, tracheostomy, or preparation for extubation, this was permitted in both groups. **Duration**: until death or discharge from the ICU, or day 28 postrandomization
Outcomes	*Outcomes that will be reported in the final trial report: 1. Ventilator-free days 2. All-cause mortality (day 90 and day 180) 3. Duration of survival 4. Quality of life 5. Functional outcome assessed by the extended Glasgow Outcome Scale 6. Proportion of participants in paid employment at baseline who are unemployed at 180 days 7. Cognitive function
Notes	*The trial report included data from a pilot phase of the ICU-ROX trial. It included the first 100 patients of an overall sample of 1000, which was to examine the feasibility. Only feasibility outcomes were reported, and outcomes prespecified in protocol will be reported in final trial report including 1000 participants, thus no relevant outcomes were reported. Email sent 6 December 2018 to Dr Young and reply was received. The trial was supported by public funds.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Encrypted web-based system
Allocation concealment (selection bias)	Low risk	Central randomization
Blinding of participants and personnel (performance bias) All outcomes	High risk	Not blinded

Young 2017 *(Continued)*

Blinding of outcome assessment (detection bias) All outcomes	Low risk	Not described; however, blinding of outcome assessment was clarified by email
Incomplete outcome data (attrition bias) All outcomes	Low risk	Less than 5% were lost to follow-up.
Selective reporting (reporting bias)	Low risk	The trial was registered prior to randomization (ACTRN12615000957594). Only feasibility outcomes were reported, and outcomes prespecified in the protocol will be reported in the final trial report including 1000 participants. However, mortality is reported in total (30.3%), but is not specified according to treatment group.
Other bias	Low risk	The trial appeared to be free of other issues that could put it at risk of bias.

APACHE II**:** Acute Physiology, Age, Chronic Health Evaluation II**; ARDS**: acute respiratory distress syndrome; **AUC**: area under the curve; **C5**: cervical spine vertebral level 5; **COPD**: chronic obstructive pulmonary disease; **CPC**: cerebral performance category; **CPR**: cardiopulmonary resuscitation; **EEG**: electroencephalogram; **FiO_2:** fraction of inspired oxygen; **GCS**: Glasgow Coma Scale; **H_2O**: dihydrogen monoxide (water); **ICU**: intensive care unit; **MAP**: mean arterial pressure; **mRS**: modified ranking scale; **MV**: mechanical ventilation; **NIRS**: cerebral near-infrared spectroscopy; **NIHSS**: National Institutes of Health Stroke Scale; **NSE**: neuron-specific enolase; **OHCA**: out-of-hospital cardiac arrest; **$PaCO_2$**: partial pressure of arterial carbon dioxide; **PaO_2**: partial pressure of arterial oxygen; **PEEP**: positive end-expiratory pressure; **PaO_2/FiO_2 ratio**: ratio of arterial oxygen partial pressure to fractional inspired oxygen; **RCT**: randomized controlled trial; **ROSC**: return of spontaneous circulation; **SaO_2:** arterial oxygen saturation of haemoglobin; **SAPS**: simplified acute physiology score; **SOFA**: sequential organ failure assessment; **SpO_2**: peripheral oxygen saturation; **TBI**: traumatic brain injury; **TnT**: cardiac troponin; **TTM**: targeted temperature management; **VF**: ventricular fibrillation; **VT**: ventricular tachycardia

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Ali 2013	Wrong population
Amar 1994	Wrong population
Austin 2010	Wrong population
Bickel 2011	Wrong population
Bray 2018	Wrong population
Hofmann 2017	Wrong population
Huynh Ky 2017	Wrong population
Khoshnood 2018	Wrong population
Khosnood 2017	Wrong population
Kuisma 2006	Wrong population
Meyhoff 2009	Wrong population
Padma 2010	Wrong population

Study	Reason for exclusion
Perrin 2011	Wrong population
Ranchord 2012	Wrong population
Rawles 1976	Wrong population
Rodrigo 2003	Wrong population
Roffe 2010	Wrong population
Roffe 2017	Wrong population
Sills 2003	Wrong population
Singhal 2005	Wrong population
Singhal 2013	Wrong population
Stub 2014	Wrong population
Ukholkina 2005	Wrong population
Wu 2014	Wrong population
Young 2014	Wrong population
Zughaft 2013	Wrong population

Characteristics of studies awaiting assessment *[ordered by study ID]*

ICU-ROX 2019

Methods	RCT
Participants	**Sample size**: 1000 (experimental 501, control 499) **Country**: New Zealand **Setting**: mechanically ventilated adults admitted to a multidisciplinary ICU
Interventions	**Experimental**: no specific measures taken to avoid high FiO_2 or SpO_2, FiO_2<0.30 discouraged (thus, we could not categorize the experimental group as either using a low or a high target). Patients assigned to the 'higher group' received 'standard care' both while ventilated and after extubation with no specific measures taken to avoid high FiO_2 or high SpO_2. The use of upper alarm limits for SpO_2 in the 'higher group' was prohibited as upper alarm limits for SpO_2 were not used as part of standard care. The lower limit alarm for SpO_2 was set at 90% (or lower if clinically appropriate). If the PaO_2 or the SaO_2 were lower than the acceptable limit, inspired oxygen might be increased if clinically appropriate, irrespective of the SpO_2 reading. The use of an FiO_2 of less than 0.3 whilst ventilated was discouraged. **Control**: target SaO_2/SpO_2 91% to 96%. When a participant was allocated to conservative oxygen therapy, the inspired oxygen concentration was decreased to room air as rapidly as possible provided that the SpO_2 measured by peripheral pulse oximetry was greater than the acceptable lower limit. SpO_2 levels of greater than 96% were strictly avoided and an upper SpO_2 alarm limit of 97%

ICU-ROX 2019 *(Continued)*

	applied whenever supplemental oxygen was administered in the ICU to minimise the risk of hyperoxaemia. After extubation, in the conservative oxygen group, the upper monitored alarm limit of acceptable SpO_2 of 97% was applied whenever supplemental oxygen was being administered. In the event that the SpO_2 exceeded the acceptable upper limit, downward titration of supplemental oxygen was undertaken as a high priority and supplemental oxygen was discontinued as soon possible. The lower limit alarm for SpO_2 was set at 90% (or lower if clinically appropriate). If the PaO_2 or the SaO_2 were lower than the acceptable limit, inspired oxygen might be increased if clinically appropriate, irrespective of the SpO_2 reading. Categorized by us as using a low target in the control group. **Duration**: until death or discharge from the ICU, or day 28 post randomization
Outcomes	**Primary outcome:** 1. Ventilator free days to day 28 **Secondary outcomes:** 1. All-cause mortality (day 90 and 180) 2. Duration of survival 3. Proportion of participants in paid employment at baseline who were unemployed at 180 days 4. Cognitive function at day 180 5. Quality of life at day 180 6. Cause-specific mortality Functional outcome assessed by the extended Glasgow outcome scale (in patients with acute brain pathologi)
Notes	The ICU-ROX trial was published post our literature search and thus was not included in this review. The ICU-ROX trial will be included in a review update.

Characteristics of ongoing studies *[ordered by study ID]*

NCT02321072

Trial name or title	The effects of hyperoxia on organ dysfunction and outcome in critically ill patients with SIRS (O_2-ICU)
Methods	RCT
Participants	Patients admitted to the ICU with ≥ 2 positive SIRS criteria and an expected ICU stay of more than 48 hours
Interventions	**Active comparator: high-normal PaO_2** In participants requiring respiratory monitoring, supplemental oxygen is titrated to achieve a PaO_2 of 120 mmHg (16 kPa), range 105 to 135 mmHg (14 to 18 kPa). **Active comparator: low-normal PaO_2** In participants requiring respiratory monitoring, supplemental oxygen is titrated to achieve a target PaO_2 of 75 mmHg (10 kPa), range 60 to 90 mmHg (8 to 18 kPa).
Outcomes	**Primary outcome** 1. Daily delta SOFA score (time frame: 14 days)

NCT02321072 *(Continued)*

	Secondary outcomes: 1. Total maximum SOFA score minus SOFA score on admission (time frame: 14 days) 2. SOFA rate of decline (time frame: 14 days) 3. Total maximum SOFA score, total maximum SOFA score minus SOFA score on admission, SOFA rate of decline (time frame: 14 days) 4. Mortality (time frame: 14 days, in-ICU (max 90 days), in-hospital (max 90 days) 5. Hypoxic events (PaO_2 < 55 mmHg) (time frame: 14 days) 6. Vasopressor or inotrope requirements (time frame: 14 days) 7. Renal function, fluid balance (time frame: 14 days) 8. Oxidative stress (F2-isoprostanes) (time frame: days 1, 3, 7) 9. Duration of mechanical ventilation and ventilator-free days (time frame: 14 days) 10. Length of stay (ICU) (time frame: average expected 2 to 28 days) 11. Length of stay (hospital) (time frame: average expected 10 to 28 days) 12. Systemic vascular resistance index (time frame: 14 days) in a random subpopulation 13. Cardiac index (time frame: 14 days) in a random subpopulation 14. Microcirculatory flow index and perfused vessel density (time frame: 14 days) in a random sub-population. Composite endpoint for 2 sidestream dark-field microcirculatory measurements
Starting date	February 2015
Contact information	Dr HJS de Grooth
Notes	

NCT02713451

Trial name or title	Liberal oxygenation versus conservative oxygenation in ARDS ($LOCO_2$)
Methods	RCT
Participants	Patients with ARDS
Interventions	**Active comparator: liberal oxygenation (LO) group** A modulation of inspired fraction of oxygen will be performed with an objective of PaO_2 between 90 to 105 mmHg, which will be checked on ABG. Between these measurements, SpO_2 will be kept at ≥ 96%. Alarms will be set at 95% for SpO_2. **Experimental: conservative oxygenation (CO) group** A modulation of inspired fraction of oxygen will be performed with an objective of PaO_2 between 55 to 70 mmHg, which will be checked on ABG. Between these measurements, SpO_2 will be kept between 88% and 92%. Alarms will be set between 87% and 93% for SpO_2.
Outcomes	**Primary outcome** 1. Death (time frame: day 28) **Secondary outcomes** 1. Death (time frame: day 90) 2. Days free of mechanical ventilation in ICU (time frame: day 28) 3. SOFA score (time frame: days 0, 3, and 7)

NCT02713451 *(Continued)*

	4. Score of morbidity (time frame: day 28). This score is based on 3 points: need for mechanical ventilation, need for renal replacement therapy, need for catecholamine. 5. Ventilator-associated pneumonia (time frame: day 28) 6. Septicaemia (time frame: day 28) 7. Antibiotic consumption (time frame: day 28) 8. Cardiovascular complications (time frame: day 28 and day 90). New onset of rhythm disorders, cardiac ischaemia, and dose of catecholamine at days 28 and 90 9. Neurological evolution (time frame: day 28). Neurological evolution measured with daily Richmond Agitation Sedation Scale score, seizures, new stroke, daily sedation doses, neuroleptic administration. 10. Respiratory autonomy (time frame: days 28 and 90). Need for oxygen or mechanical ventilation support
Starting date	June 2016
Contact information	Loïc Barrot
Notes	

NCT03141099

Trial name or title	Blood pressure and oxygenation targets in post-resuscitation care (BOX)
Methods	RCT
Participants	Comatose OHCA patients
Interventions	**Active comparator: low normal MAP and low normal PaO_2** MAP 63 mmHg and PaO_2 9 to 10 kPa during targeted temperature management (36 hours) after OHCA **Active comparator: high normal MAP and low normal PaO_2** MAP 77 mmHg and PaO_2 9 to 10 kPa during targeted temperature management (36 hours) after OHCA **Active comparator: low normal MAP and high normal PaO_2** MAP 63 mmHg and PaO_2 13 to 14 kPa during targeted temperature management (36 hours) after OHCA **Active comparator: high normal MAP and high normal PaO_2** MAP 77 mmHg and PaO_2 13 to 14 kPa during targeted temperature management (36 hours) after OHCA
Outcomes	**Primary outcome** 1. All-cause mortality or severe anoxic brain injury (time frame: 3 months after OHCA) **Secondary outcomes** 1. Renal replacement therapy (time frame: 3 months) 2. Time to death (time frame: 180 days) 3. Neuron-specific enolase (time frame: 48 hours) 4. MOCA score (time frame: 3 months)

NCT03141099 *(Continued)*

	5. Modified Ranking Scale (time frame: 3 months) 6. NT-pro-BNP (time frame: 3 months) 7. eGFR (time frame: 3 months) 8. LVEF (time frame: 3 months) 9. Vasopressor use (time frame: first week after cardiac arrest) 10. Renal function (time frame: 96 hours) **Other outcome measures** 1. Vital status at 180 days post-cardiac arrest (time frame: 180 days post-cardiac arrest) 2. CPC at 180 days post-cardiac arrest (time frame: 180 days post-cardiac arrest)
Starting date	March 2017
Contact information	Dr Jesper Kjaergaard
Notes	

NCT03174002

Trial name or title	Handling oxygenation targets in adults with acute hypoxaemic respiratory failure in the intensive care unit (HOT-ICU)
Methods	RCT
Participants	ICU patients
Interventions	**Experimental: low oxygenation target** Partial pressure of oxygen in arterial blood (PaO_2) 8 kPa (60 mmHg) **Active comparator: high oxygenation target** Partial pressure of oxygen in arterial blood (PaO_2) 12 kPa (90 mmHg)
Outcomes	**Primary outcome** 1. 90-day mortality (time frame: 90 days) **Secondary outcomes** 1. Days alive without organ support (time frame: within 90 days) 2. Days alive out of the hospital (time frame: within 90 days) 3. Number of participants with 1 or more serious adverse events (time frame: until ICU discharge, maximum 90 days) 4. 1-year mortality (time frame: 1 year) 5. Quality of life assessment using the EQ-5D-5L telephone interview in selected sites (time frame: 1 year) 6. Cognitive function 1-year after randomization as assessed using the RBANS score in selected sites (time frame: 1 year) 7. Pulmonary function (time frame: 1 year) 8. A health economic analysis (time frame: 90 days)
Starting date	June 2017

NCT03174002 *(Continued)*

Contact information	Dr Bodil Steen Rasmussen
Notes	

NCT03287466

Trial name or title	A randomised controlled trial of targeted oxygen therapy in mechanically ventilated critically ill patients (TOXYC)
Methods	RCT
Participants	Mechanically ventilated adults
Interventions	**Experimental: SpO_2 88% to 92%** The intervention is TO2T to achieve an arterial haemoglobin oxygen saturation (SpO_2) of 88% to 92%. **Active comparator: SpO_2 96% or above** The control group will also receive TO2T, but to achieve an SpO_2 of 96% or above (standard care).
Outcomes	**Primary outcome measures** 1. Feasibility (time frame: 15 months) **Secondary outcomes** 1. Measurement of ABG (time frame: up to 21 days) 2. Measurement of oxygen saturation (time frame: up to 21 days) 3. Measurement of fraction of inspired oxygen (time frame: up to 21 days) 4. Time to extubation or detachment from mechanical ventilation (time frame: up to 21 days) 5. Mechanical ventilation-free days on ICU (time frame: up to 21 days) 6. Measurement of blood pressure (time frame: up to 21 days) 7. Measurement of heart rate (time frame: up to 21 days) 8. Measurement of cardiac rhythm (time frame: up to 21 days) 9. Measurement of cardiac output and stroke volume (if measured) (time frame: up to 21 days) 10. Measurement of vasopressor doses (time frame: up to 21 days) 11. Measurement of inotrope doses (time frame: up to 21 days) 12. Measurement of daily fluid balance (time frame: up to 21 days) 13. Measurement of inotrope-free days on ICU (time frame: up to 21 days) 14. Measurement of vasopressor-free days on ICU (time frame: up to 21 days) 15. Measurement of urea (time frame: up to 21 days) 16. Measurement of creatinine (time frame: up to 21 days) 17. Measurement of urine output (time frame: up to 21 days) 18. Need for renal replacement therapy (time frame: up to 21 days) 19. Renal replacement therapy-free days on ICU (time frame: up to 21 days) 20. Measurement of transaminases (time frame: up to 21 days) 21. Measurement of blood clotting values (time frame: up to 21 days) 22. Measurement of bilirubin (time frame: up to 21 days) 23. Measurement of blood lactate (time frame: up to 21 days) 24. Measurement of troponin (time frame: up to 21 days) 25. Adverse events (time frame: 90 days)

NCT03287466 *(Continued)*	26.SOFA score change (time frame: up to 21 days) 27.APACHE II score change (time frame: up to 21 days) 28.Length of ICU stay (time frame: up to 21 days) 29.Length of hospital stay (time frame: 90 days) 30.Mortality rates (time frame: 90 days) 31.Days alive out of hospital (time frame: 90 days)
Starting date	January 2018
Contact information	Dr Jack D Grierson
Notes	

ABG: arterial blood gases; **APACHE**: Acute Physiology, Age, Chronic Health Evaluation; **ARDS**: acute respiratory distress syndrome; **CO**: conservative oxygenation; **CPC**: cerebral performance category; **eGFR**: estimated glomerular filtration rate; **EQ-5D-5L:** an instrument for measuring quality of life; **FiO_2**: fraction of inspired oxygen; **ICU**: intensive care unit; **LVEF**: left ventricular ejection fraction; **LO**: liberal oxygenation; **MAP**: mean arterial pressure; **MOCA**: Montreal Cognitive Assessment; **NT-pro-BNP**: cardiac biomarker; **OHCA**: out-of-hospital cardiac arrest; **PaO_2**: partial pressure of arterial oxygen; **RBANS**: Repeatable Battery for the Assessment of Neuropsychological Status; **RCT**: randomized controlled trial; **SaO_2**: arterial oxygen saturation of haemoglobin; **SIRS**: systemic inflammatory response syndrome; **SOFA**: sequential organ failure assessment; **SpO_2**: peripheral oxygen saturation; **TO2T**: targeted oxygen therapy

DATA AND ANALYSES

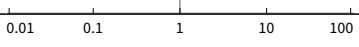
Comparison 1. All-cause mortality - at time point closest to 3 months follow-up

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 All-cause mortality - at time point closest to 3 months	4	1135	Risk Ratio (M-H, Random, 95% CI)	1.18 [1.01, 1.37]
2 Sensitivity analysis: all-cause mortality - at time point closest to 3 months - high vs high and low vs low targets excluded	2	537	Risk Ratio (M-H, Fixed, 95% CI)	1.11 [0.92, 1.35]
3 Sensitivity analysis: all-cause mortality - at time point closest to 3 months - best-worst-case scenario	4	1149	Risk Ratio (M-H, Random, 95% CI)	1.13 [0.97, 1.31]
3.1 All-cause mortality	4	1149	Risk Ratio (M-H, Random, 95% CI)	1.13 [0.97, 1.31]
4 Sensitivity analysis: all-cause mortality - at time point closest to 3 months - worst-best-case scenario	4	1149	Risk Ratio (M-H, Random, 95% CI)	1.21 [1.04, 1.41]
4.1 All-cause mortality	4	1149	Risk Ratio (M-H, Random, 95% CI)	1.21 [1.04, 1.41]
5 All-cause mortality - at time point closest to 3 months - types of oxygen interventions	4	1135	Risk Ratio (M-H, Random, 95% CI)	1.18 [1.01, 1.37]
5.1 PaO_2 (SaO_2 or SpO_2)	3	701	Risk Ratio (M-H, Random, 95% CI)	1.20 [0.96, 1.50]
5.2 Difference between groups	1	434	Risk Ratio (M-H, Random, 95% CI)	1.16 [0.94, 1.43]

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
6 All-cause mortality - at time point closest to 3 months - level of FiO_2/target in higher group	4	1135	Risk Ratio (M-H, Random, 95% CI)	1.18 [1.01, 1.37]
6.1 Higher	4	1135	Risk Ratio (M-H, Random, 95% CI)	1.18 [1.01, 1.37]
7 All-cause mortality - at time point closest to 3 months - level of FiO_2/target in lower group	4	1135	Risk Ratio (M-H, Random, 95% CI)	1.18 [1.01, 1.37]
7.1 Lower	2	537	Risk Ratio (M-H, Random, 95% CI)	1.12 [0.92, 1.35]
7.2 Higher	2	598	Risk Ratio (M-H, Random, 95% CI)	1.29 [1.00, 1.66]
8 All-cause mortality - at time point closest to 3 months - ICU population	4	1135	Risk Ratio (M-H, Fixed, 95% CI)	1.19 [1.02, 1.38]
8.1 Mixed ICU	3	1015	Risk Ratio (M-H, Fixed, 95% CI)	1.19 [1.01, 1.40]
8.2 Any cerebral disease	1	120	Risk Ratio (M-H, Fixed, 95% CI)	1.15 [0.68, 1.95]
9 Mortality - at time point closest to 3 months - oxygen delivery system	4	1135	Risk Ratio (M-H, Fixed, 95% CI)	1.19 [1.02, 1.38]
9.1 Invasive mechanical ventilation	3	657	Risk Ratio (M-H, Fixed, 95% CI)	1.12 [0.93, 1.34]
9.2 Mixed	1	478	Risk Ratio (M-H, Fixed, 95% CI)	1.33 [1.00, 1.78]

Analysis 1.1. Comparison 1 All-cause mortality - at time point closest to 3 months follow-up, Outcome 1 All-cause mortality - at time point closest to 3 months.

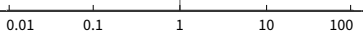
Study or subgroup	Higher n/N	Lower n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Asfar 2017	104/217	90/217		53.04%	1.16[0.94,1.43]
Girardis 2016	80/243	58/235		28.57%	1.33[1,1.78]
Jakkula 2018	20/59	18/61		8.45%	1.15[0.68,1.95]
Panwar 2016	19/51	21/52		9.94%	0.92[0.57,1.5]
Total (95% CI)	**570**	**565**		**100%**	**1.18[1.01,1.37]**
Total events: 223 (Higher), 187 (Lower)					
Heterogeneity: Tau^2=0; Chi^2=1.75, df=3(P=0.63); I^2=0%					
Test for overall effect: Z=2.08(P=0.04)					



Favours Higher — Favours Lower

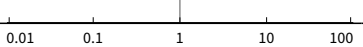
Analysis 1.2. Comparison 1 All-cause mortality - at time point closest to 3 months follow-up, Outcome 2 Sensitivity analysis: all-cause mortality - at time point closest to 3 months - high vs high and low vs low targets excluded.

Study or subgroup	Higher n/N	Lower n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Asfar 2017	104/217	90/217		81.23%	1.16[0.94,1.43]
Panwar 2016	19/51	21/52		18.77%	0.92[0.57,1.5]
Total (95% CI)	**268**	**269**		**100%**	**1.11[0.92,1.35]**
Total events: 123 (Higher), 111 (Lower)					
Heterogeneity: Tau^2=0; Chi^2=0.7, df=1(P=0.4); I^2=0%					
Test for overall effect: Z=1.08(P=0.28)					



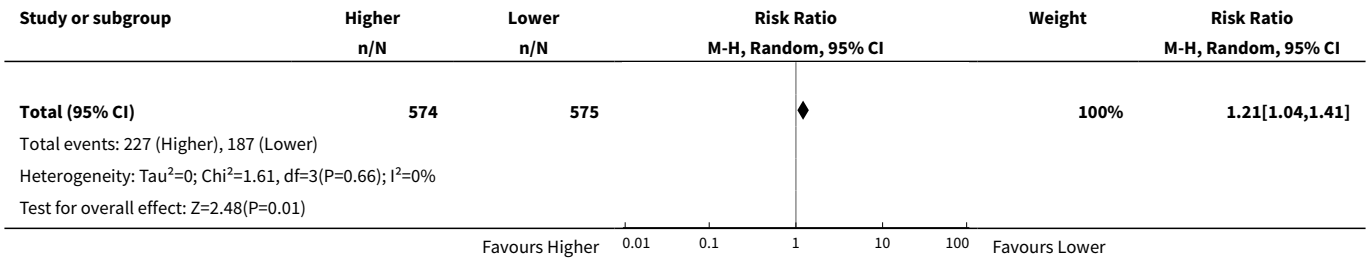
Analysis 1.3. Comparison 1 All-cause mortality - at time point closest to 3 months follow-up, Outcome 3 Sensitivity analysis: all-cause mortality - at time point closest to 3 months - best-worst-case scenario.

Study or subgroup	Higher n/N	Lower n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
1.3.1 All-cause mortality					
Asfar 2017	104/219	96/223		53.56%	1.1[0.9,1.35]
Girardis 2016	80/244	59/236		27.86%	1.31[0.99,1.74]
Jakkula 2018	20/60	20/63		8.72%	1.05[0.63,1.75]
Panwar 2016	19/51	22/53		9.87%	0.9[0.56,1.45]
Subtotal (95% CI)	**574**	**575**		**100%**	**1.13[0.97,1.31]**
Total events: 223 (Higher), 197 (Lower)					
Heterogeneity: Tau^2=0; Chi^2=2.09, df=3(P=0.55); I^2=0%					
Test for overall effect: Z=1.59(P=0.11)					
Total (95% CI)	**574**	**575**		**100%**	**1.13[0.97,1.31]**
Total events: 223 (Higher), 197 (Lower)					
Heterogeneity: Tau^2=0; Chi^2=2.09, df=3(P=0.55); I^2=0%					
Test for overall effect: Z=1.59(P=0.11)					

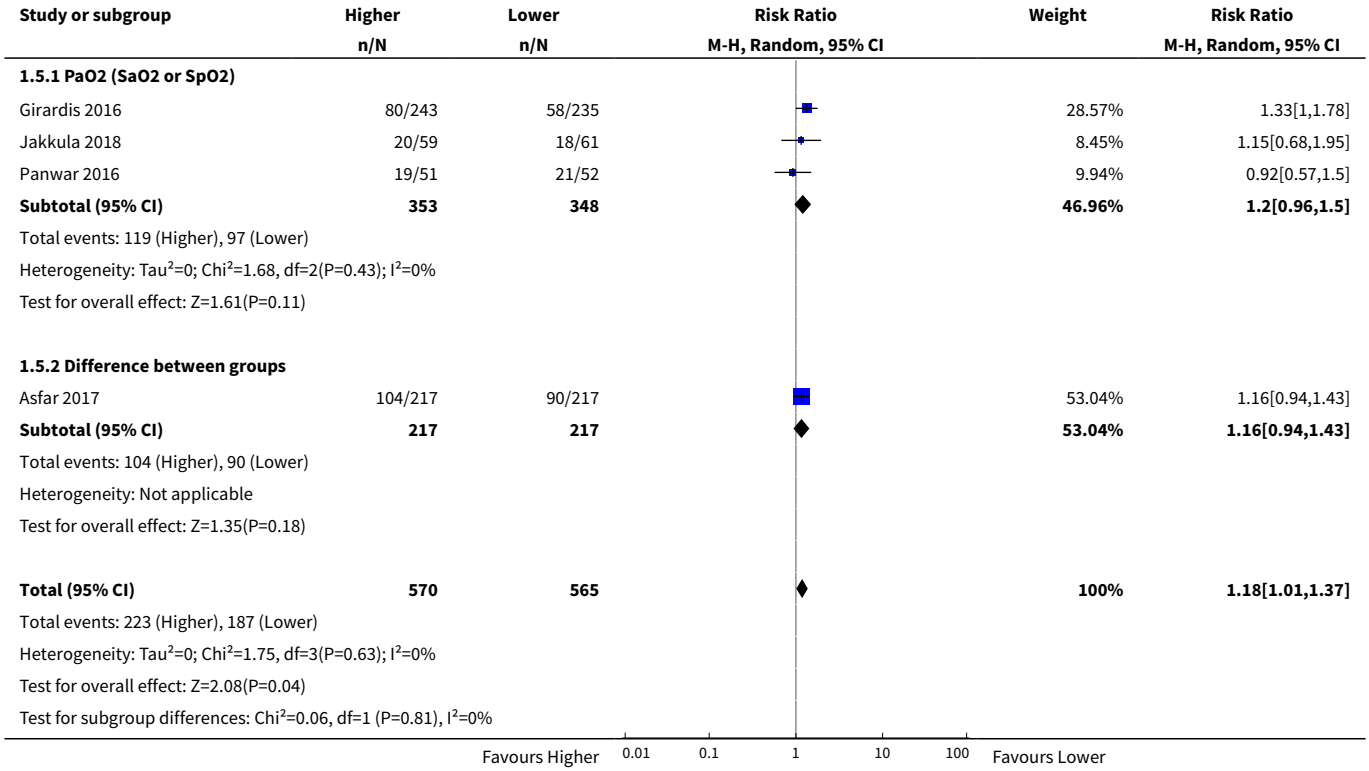


Analysis 1.4. Comparison 1 All-cause mortality - at time point closest to 3 months follow-up, Outcome 4 Sensitivity analysis: all-cause mortality - at time point closest to 3 months - worst-best-case scenario.

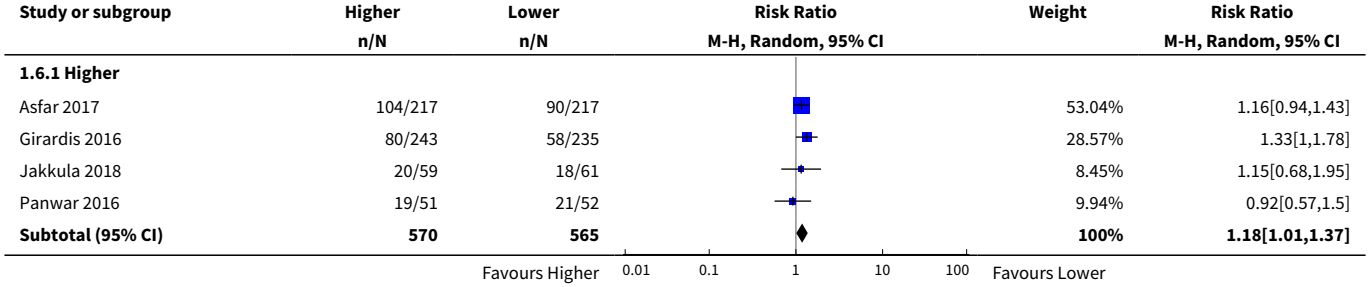
Study or subgroup	Higher n/N	Lower n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
1.4.1 All-cause mortality					
Asfar 2017	106/219	90/223		52.92%	1.2[0.97,1.48]
Girardis 2016	81/244	58/236		28.62%	1.35[1.02,1.8]
Jakkula 2018	21/60	18/63		8.61%	1.23[0.73,2.06]
Panwar 2016	19/51	21/53		9.85%	0.94[0.58,1.53]
Subtotal (95% CI)	**574**	**575**		**100%**	**1.21[1.04,1.41]**
Total events: 227 (Higher), 187 (Lower)					
Heterogeneity: Tau^2=0; Chi^2=1.61, df=3(P=0.66); I^2=0%					
Test for overall effect: Z=2.48(P=0.01)					

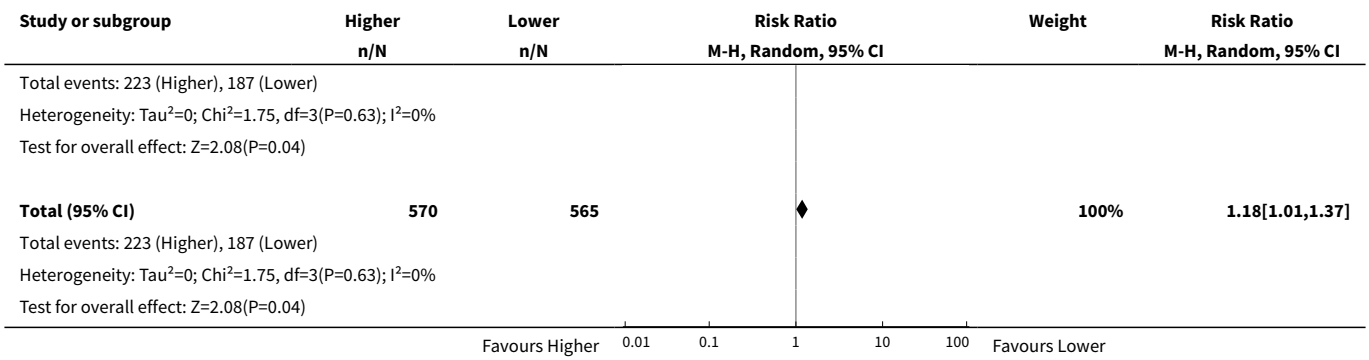


Analysis 1.5. Comparison 1 All-cause mortality - at time point closest to 3 months follow-up, Outcome 5 All-cause mortality - at time point closest to 3 months - types of oxygen interventions.

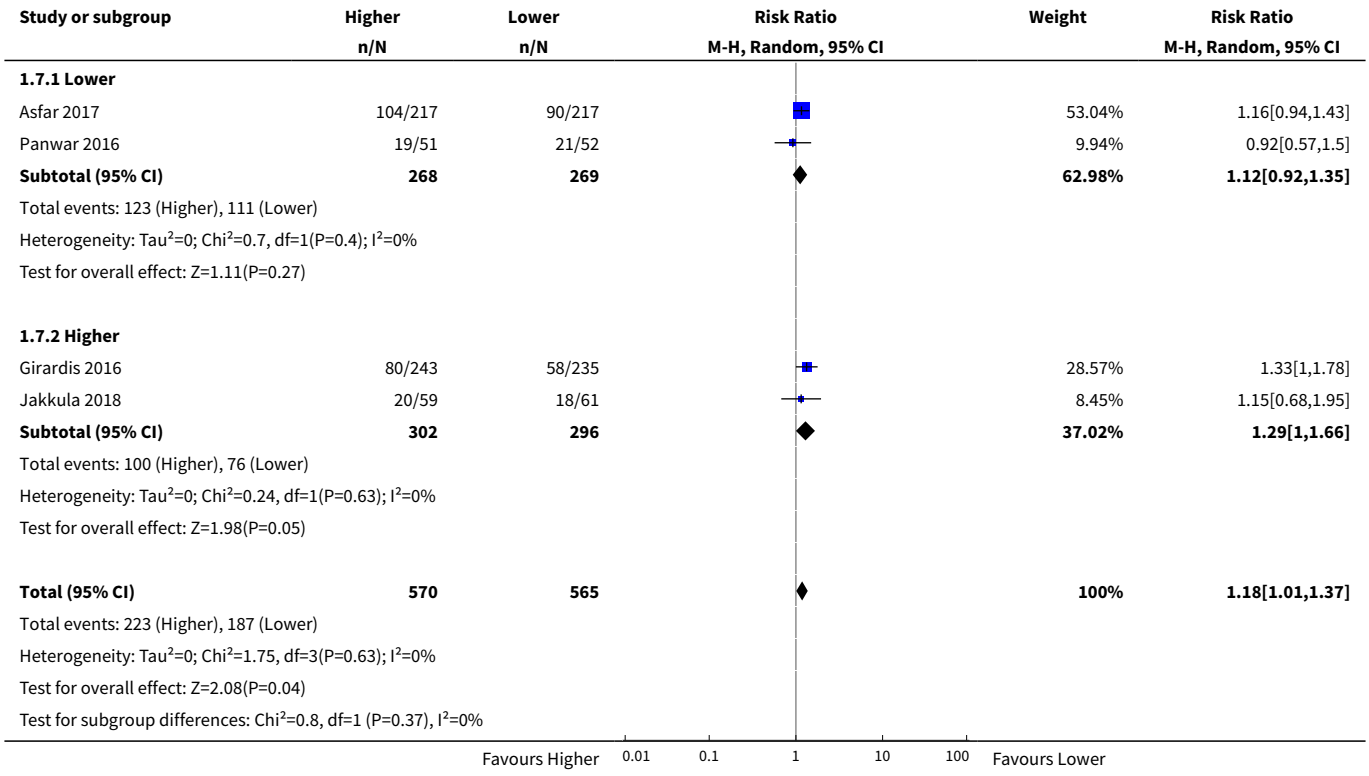


Analysis 1.6. Comparison 1 All-cause mortality - at time point closest to 3 months follow-up, Outcome 6 All-cause mortality - at time point closest to 3 months - level of FiO_2/target in higher group.

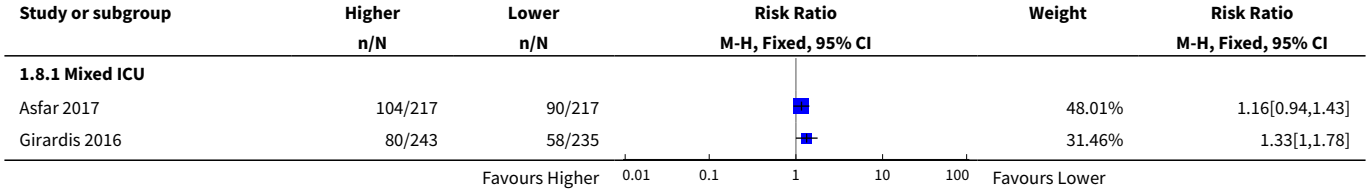


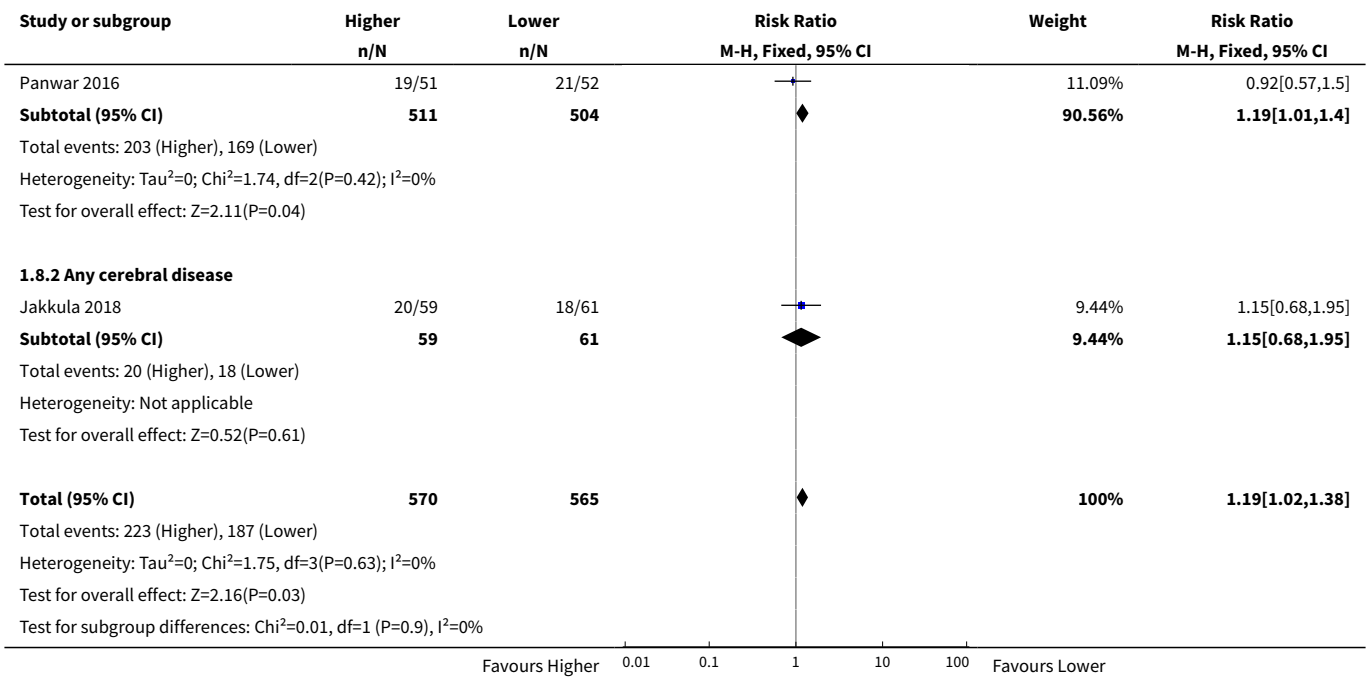


Analysis 1.7. Comparison 1 All-cause mortality - at time point closest to 3 months follow-up, Outcome 7 All-cause mortality - at time point closest to 3 months - level of FiO_2/target in lower group.

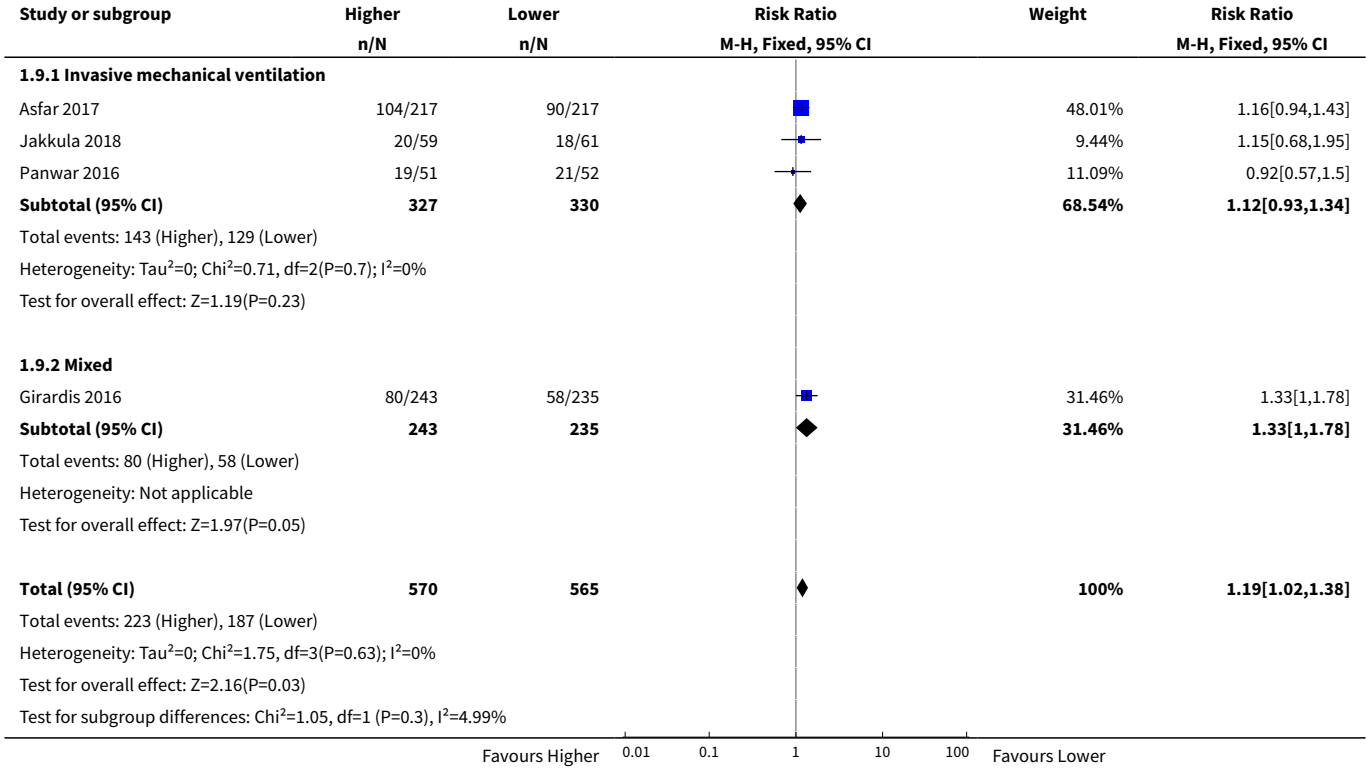


Analysis 1.8. Comparison 1 All-cause mortality - at time point closest to 3 months follow-up, Outcome 8 All-cause mortality - at time point closest to 3 months - ICU population.





Analysis 1.9. Comparison 1 All-cause mortality - at time point closest to 3 months follow-up, Outcome 9 Mortality - at time point closest to 3 months - oxygen delivery system.



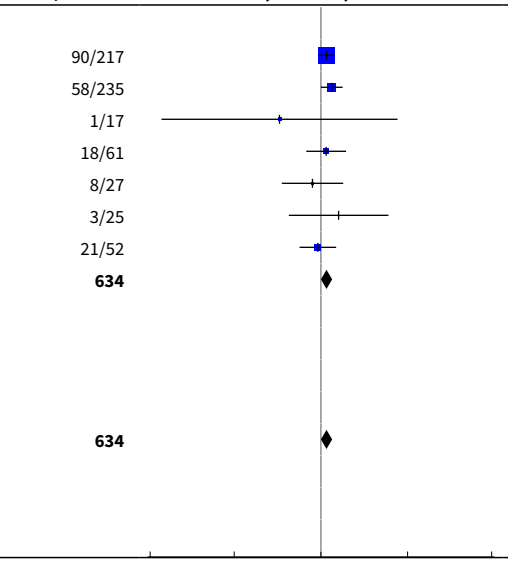
Comparison 2. Sensitivity analysis: all-cause mortality - at maximum follow-up

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 All-cause mortality - at maximum follow-up	7	1285	Risk Ratio (M-H, Random, 95% CI)	1.16 [1.00, 1.35]
1.1 All-cause mortality	7	1285	Risk Ratio (M-H, Random, 95% CI)	1.16 [1.00, 1.35]
2 Sensitivity analysis: all-cause mortality - at maximum follow-up - high vs high and low vs low excluded	2	537	Risk Ratio (M-H, Fixed, 95% CI)	1.11 [0.92, 1.35]
3 Sensitivity analysis: all-cause mortality - at maximum follow-up - best-worst-case scenario	7	1306	Risk Ratio (M-H, Random, 95% CI)	1.11 [0.96, 1.28]
3.1 All-cause mortality	7	1306	Risk Ratio (M-H, Random, 95% CI)	1.11 [0.96, 1.28]
4 Sensitivity analysis: all-cause mortality - at maximum follow-up - worst-best-case scenario	7	1306	Risk Ratio (M-H, Random, 95% CI)	1.21 [1.05, 1.41]
4.1 All-cause mortality	7	1306	Risk Ratio (M-H, Random, 95% CI)	1.21 [1.05, 1.41]
5 All-cause mortality - at maximum follow-up - types of oxygen interventions	7	1285	Risk Ratio (M-H, Fixed, 95% CI)	1.17 [1.00, 1.36]
5.1 PaO_2 (SaO_2 or SpO_2)	4	735	Risk Ratio (M-H, Fixed, 95% CI)	1.20 [0.96, 1.50]
5.2 FiO_2	2	116	Risk Ratio (M-H, Fixed, 95% CI)	1.00 [0.50, 1.98]
5.3 Difference between groups	1	434	Risk Ratio (M-H, Fixed, 95% CI)	1.16 [0.94, 1.43]
6 All-cause mortality - at maximum follow-up - level of FiO_2/target in higher group	7	1285	Risk Ratio (M-H, Fixed, 95% CI)	1.17 [1.00, 1.36]
6.1 Lower	2	85	Risk Ratio (M-H, Fixed, 95% CI)	1.18 [0.37, 3.81]
6.2 Higher	5	1200	Risk Ratio (M-H, Fixed, 95% CI)	1.17 [1.00, 1.36]
7 All-cause mortality - at maximum follow-up - level of FiO_2/target in lower group	7	1285	Risk Ratio (M-H, Fixed, 95% CI)	1.17 [1.00, 1.36]
7.1 Lower	4	622	Risk Ratio (M-H, Fixed, 95% CI)	1.11 [0.92, 1.35]
7.2 Higher	3	663	Risk Ratio (M-H, Fixed, 95% CI)	1.24 [0.97, 1.57]
8 All-cause mortality - at maximum follow-up - ICU population	7	1350	Risk Ratio (M-H, Fixed, 95% CI)	1.15 [0.99, 1.33]
8.1 Mixed ICU	3	1015	Risk Ratio (M-H, Fixed, 95% CI)	1.19 [1.01, 1.40]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
8.2 Medical ICU	1	34	Risk Ratio (M-H, Fixed, 95% CI)	0.33 [0.01, 7.65]
8.3 Any trauma	1	65	Risk Ratio (M-H, Fixed, 95% CI)	0.80 [0.35, 1.81]
8.4 Any cerebral disease	3	236	Risk Ratio (M-H, Fixed, 95% CI)	1.09 [0.71, 1.65]
9 Mortality - at maximum follow-up - oxygen delivery system	7	1285	Risk Ratio (M-H, Fixed, 95% CI)	1.17 [1.00, 1.36]
9.1 Invasive mechanical ventilation	4	722	Risk Ratio (M-H, Fixed, 95% CI)	1.10 [0.92, 1.31]
9.2 Any non-invasive oxygen administration	2	85	Risk Ratio (M-H, Fixed, 95% CI)	1.18 [0.37, 3.81]
9.3 Mixed	1	478	Risk Ratio (M-H, Fixed, 95% CI)	1.33 [1.00, 1.78]

Analysis 2.1. Comparison 2 Sensitivity analysis: all-cause mortality - at maximum follow-up, Outcome 1 All-cause mortality - at maximum follow-up.

Study or subgroup	Higher n/N	Lower n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
2.1.1 All-cause mortality					
Asfar 2017	104/217	90/217		50.46%	1.16[0.94,1.43]
Girardis 2016	80/243	58/235		27.18%	1.33[1,1.78]
Gomersall 2002	0/17	1/17		0.23%	0.33[0.01,7.65]
Jakkula 2018	20/59	18/61		8.04%	1.15[0.68,1.95]
Lång 2018	9/38	8/27		3.36%	0.8[0.35,1.81]
Mazdeh 2015	5/26	3/25		1.28%	1.6[0.43,6.01]
Panwar 2016	19/51	21/52		9.46%	0.92[0.57,1.5]
Subtotal (95% CI)	**651**	**634**		**100%**	**1.16[1,1.35]**
Total events: 237 (Higher), 199 (Lower)					
Heterogeneity: Tau^2=0; Chi^2=3.41, df=6(P=0.76); I^2=0%					
Test for overall effect: Z=1.98(P=0.05)					
Total (95% CI)	**651**	**634**		**100%**	**1.16[1,1.35]**
Total events: 237 (Higher), 199 (Lower)					
Heterogeneity: Tau^2=0; Chi^2=3.41, df=6(P=0.76); I^2=0%					
Test for overall effect: Z=1.98(P=0.05)					

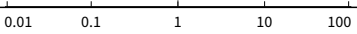


Analysis 2.2. Comparison 2 Sensitivity analysis: all-cause mortality - at maximum follow-up, Outcome 2 Sensitivity analysis: all-cause mortality - at maximum follow-up - high vs high and low vs low excluded.

Study or subgroup	Higher n/N	Lower n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Asfar 2017	104/217	90/217		81.23%	1.16[0.94,1.43]
Panwar 2016	19/51	21/52		18.77%	0.92[0.57,1.5]
Total (95% CI)	**268**	**269**		**100%**	**1.11[0.92,1.35]**
Total events: 123 (Higher), 111 (Lower)					
Heterogeneity: Tau^2=0; Chi^2=0.7, df=1(P=0.4); I^2=0%					
Test for overall effect: Z=1.08(P=0.28)					

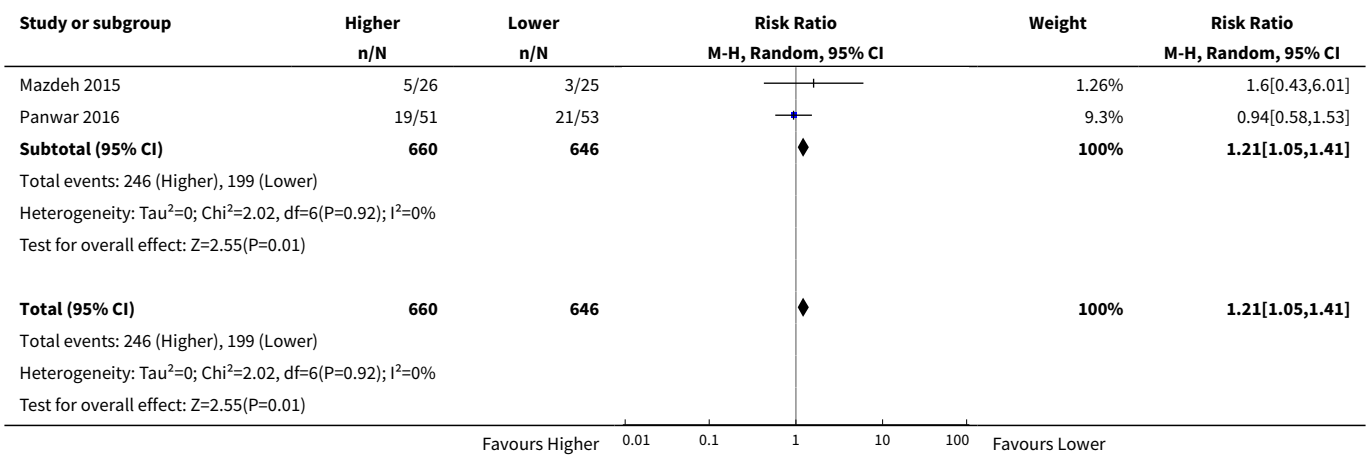
Analysis 2.3. Comparison 2 Sensitivity analysis: all-cause mortality - at maximum follow-up, Outcome 3 Sensitivity analysis: all-cause mortality - at maximum follow-up - best-worst-case scenario.

Study or subgroup	Higher n/N	Lower n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
2.3.1 All-cause mortality					
Asfar 2017	104/219	96/223		50.82%	1.1[0.9,1.35]
Girardis 2016	80/244	59/236		26.43%	1.31[0.99,1.74]
Gomersall 2002	0/19	1/17		0.22%	0.3[0.01,6.91]
Jakkula 2018	20/60	20/63		8.27%	1.05[0.63,1.75]
Lång 2018	9/41	10/29		3.67%	0.64[0.3,1.37]
Mazdeh 2015	5/26	3/25		1.23%	1.6[0.43,6.01]
Panwar 2016	19/51	22/53		9.36%	0.9[0.56,1.45]
Subtotal (95% CI)	**660**	**646**		**100%**	**1.11[0.96,1.28]**
Total events: 237 (Higher), 211 (Lower)					
Heterogeneity: Tau^2=0; Chi^2=5.12, df=6(P=0.53); I^2=0%					
Test for overall effect: Z=1.37(P=0.17)					
Total (95% CI)	**660**	**646**		**100%**	**1.11[0.96,1.28]**
Total events: 237 (Higher), 211 (Lower)					
Heterogeneity: Tau^2=0; Chi^2=5.12, df=6(P=0.53); I^2=0%					
Test for overall effect: Z=1.37(P=0.17)					

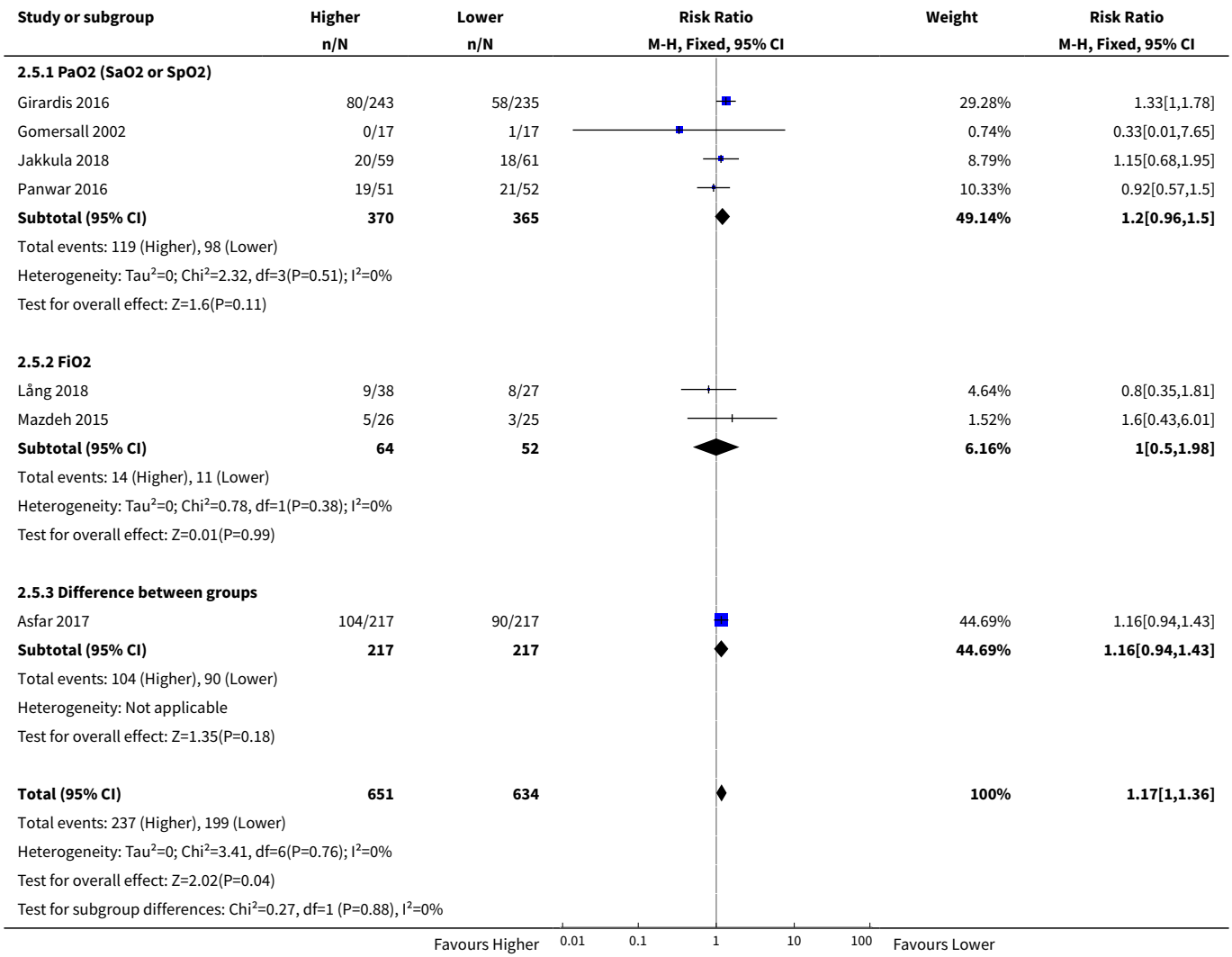


Analysis 2.4. Comparison 2 Sensitivity analysis: all-cause mortality - at maximum follow-up, Outcome 4 Sensitivity analysis: all-cause mortality - at maximum follow-up - worst-best-case scenario.

Study or subgroup	Higher n/N	Lower n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
2.4.1 All-cause mortality					
Asfar 2017	106/219	90/223		50%	1.2[0.97,1.48]
Girardis 2016	81/244	58/236		27.04%	1.35[1.02,1.8]
Gomersall 2002	2/19	1/17		0.41%	1.79[0.18,18.02]
Jakkula 2018	21/60	18/63		8.14%	1.23[0.73,2.06]
Lång 2018	12/41	8/29		3.85%	1.06[0.5,2.26]

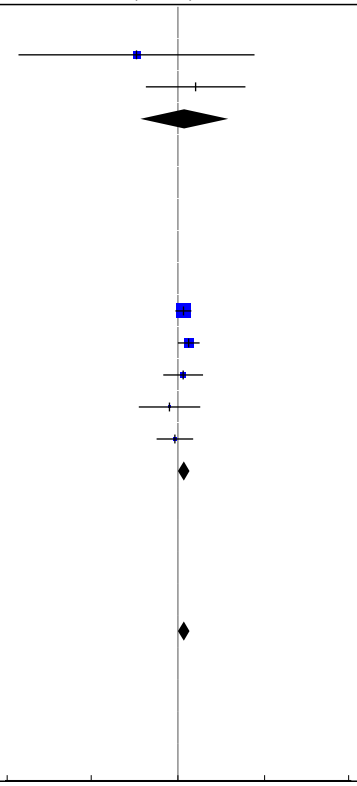


Analysis 2.5. Comparison 2 Sensitivity analysis: all-cause mortality - at maximum follow-up, Outcome 5 All-cause mortality - at maximum follow-up - types of oxygen interventions.



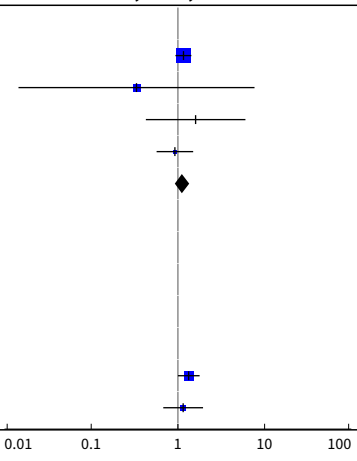
Analysis 2.6. Comparison 2 Sensitivity analysis: all-cause mortality - at maximum follow-up, Outcome 6 All-cause mortality - at maximum follow-up - level of FiO_2/target in higher group.

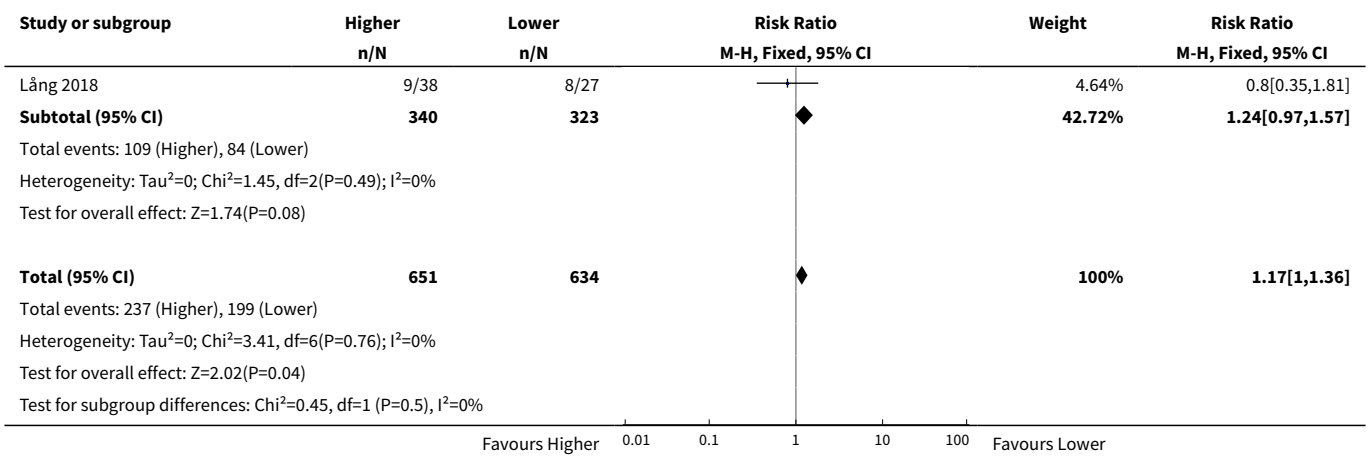
Study or subgroup	Higher n/N	Lower n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
2.6.1 Lower					
Gomersall 2002	0/17	1/17		0.74%	0.33[0.01,7.65]
Mazdeh 2015	5/26	3/25		1.52%	1.6[0.43,6.01]
Subtotal (95% CI)	**43**	**42**		**2.26%**	**1.18[0.37,3.81]**
Total events: 5 (Higher), 4 (Lower)					
Heterogeneity: Tau^2=0; Chi^2=0.83, df=1(P=0.36); I^2=0%					
Test for overall effect: Z=0.28(P=0.78)					
2.6.2 Higher					
Asfar 2017	104/217	90/217		44.69%	1.16[0.94,1.43]
Girardis 2016	80/243	58/235		29.28%	1.33[1,1.78]
Jakkula 2018	20/59	18/61		8.79%	1.15[0.68,1.95]
Lång 2018	9/38	8/27		4.64%	0.8[0.35,1.81]
Panwar 2016	19/51	21/52		10.33%	0.92[0.57,1.5]
Subtotal (95% CI)	**608**	**592**		**97.74%**	**1.17[1,1.36]**
Total events: 232 (Higher), 195 (Lower)					
Heterogeneity: Tau^2=0; Chi^2=2.58, df=4(P=0.63); I^2=0%					
Test for overall effect: Z=2(P=0.05)					
Total (95% CI)	**651**	**634**		**100%**	**1.17[1,1.36]**
Total events: 237 (Higher), 199 (Lower)					
Heterogeneity: Tau^2=0; Chi^2=3.41, df=6(P=0.76); I^2=0%					
Test for overall effect: Z=2.02(P=0.04)					
Test for subgroup differences: Chi^2=0, df=1 (P=0.98), I^2=0%					



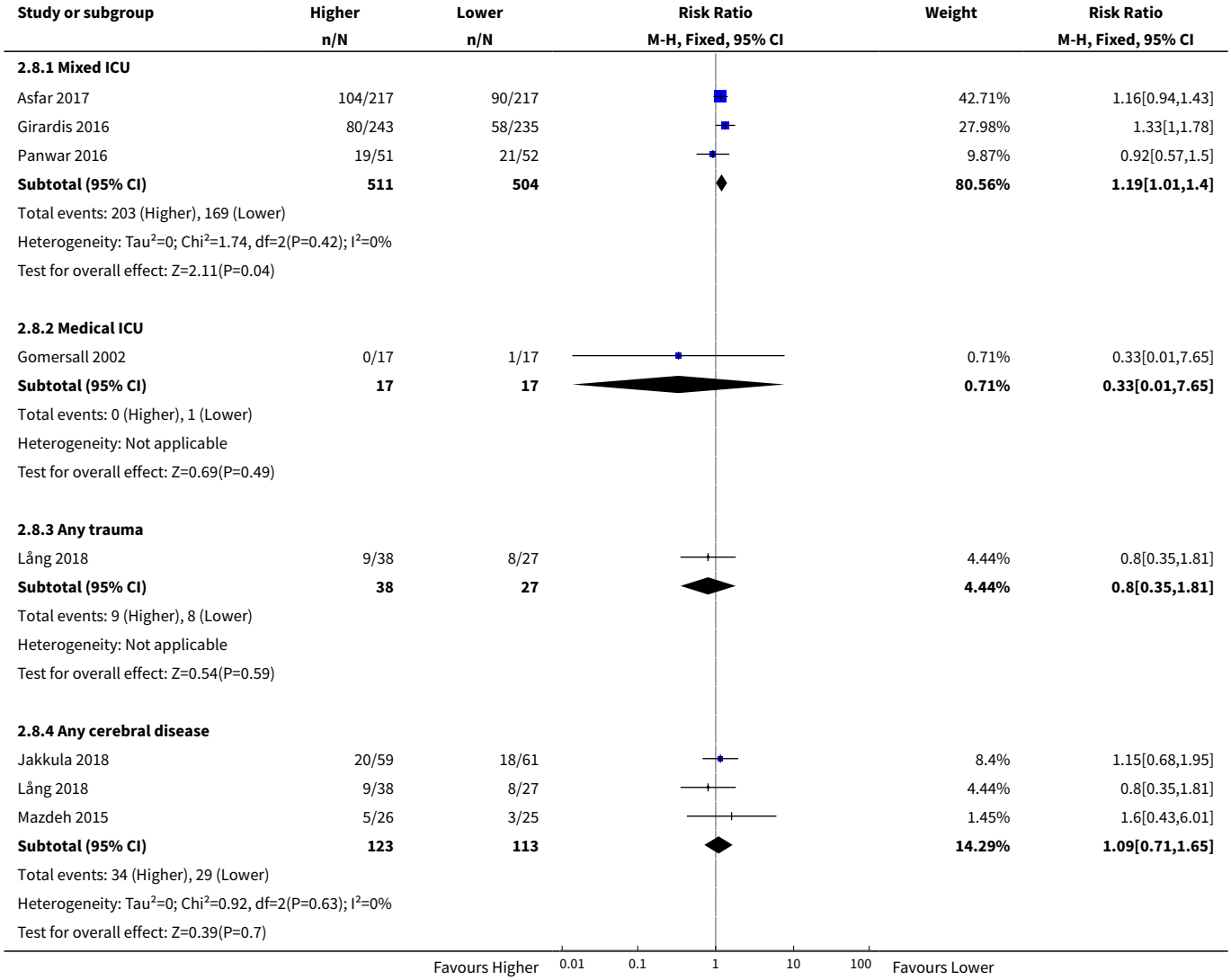
Analysis 2.7. Comparison 2 Sensitivity analysis: all-cause mortality - at maximum follow-up, Outcome 7 All-cause mortality - at maximum follow-up - level of FiO_2/target in lower group.

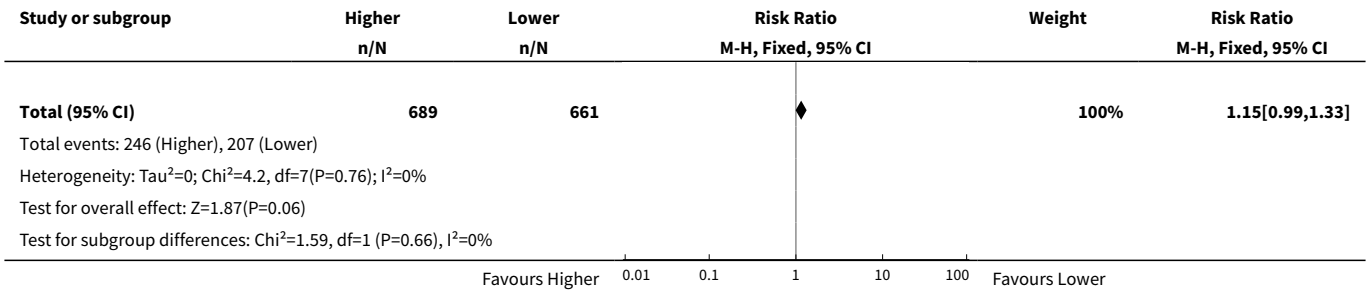
Study or subgroup	Higher n/N	Lower n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
2.7.1 Lower					
Asfar 2017	104/217	90/217		44.69%	1.16[0.94,1.43]
Gomersall 2002	0/17	1/17		0.74%	0.33[0.01,7.65]
Mazdeh 2015	5/26	3/25		1.52%	1.6[0.43,6.01]
Panwar 2016	19/51	21/52		10.33%	0.92[0.57,1.5]
Subtotal (95% CI)	**311**	**311**		**57.28%**	**1.11[0.92,1.35]**
Total events: 128 (Higher), 115 (Lower)					
Heterogeneity: Tau^2=0; Chi^2=1.56, df=3(P=0.67); I^2=0%					
Test for overall effect: Z=1.11(P=0.27)					
2.7.2 Higher					
Girardis 2016	80/243	58/235		29.28%	1.33[1,1.78]
Jakkula 2018	20/59	18/61		8.79%	1.15[0.68,1.95]



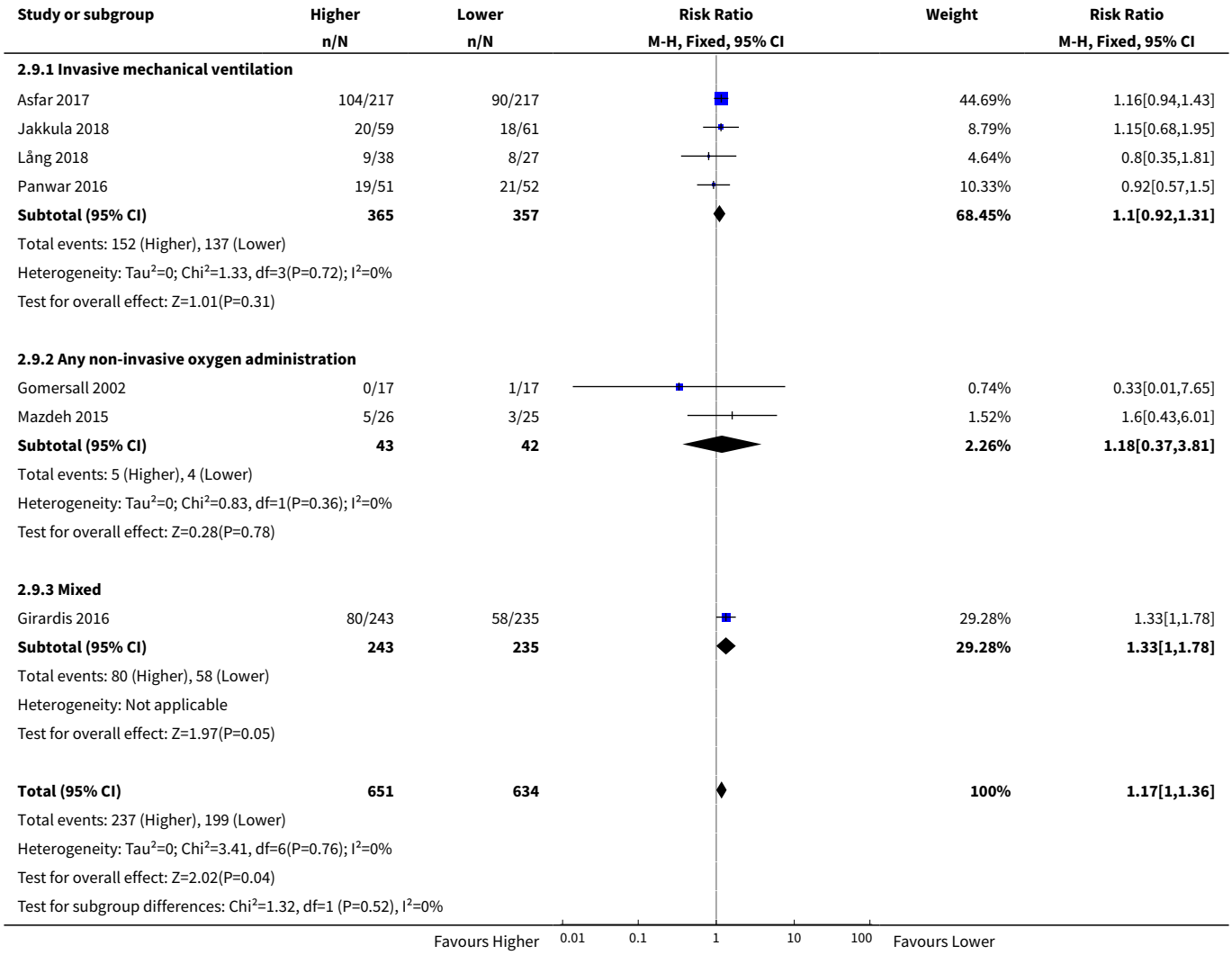


Analysis 2.8. Comparison 2 Sensitivity analysis: all-cause mortality - at maximum follow-up, Outcome 8 All-cause mortality - at maximum follow-up - ICU population.





Analysis 2.9. Comparison 2 Sensitivity analysis: all-cause mortality - at maximum follow-up, Outcome 9 Mortality - at maximum follow-up - oxygen delivery system.

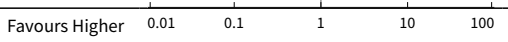


Comparison 3. Serious adverse events - at time point closest to 3 months

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Serious adverse events - at time point closest to three months - highest proportion	6	1234	Risk Ratio (M-H, Random, 95% CI)	1.13 [1.04, 1.23]
2 Serious adverse events - at time point closest to three months - cumulated	6	1234	Risk Ratio (M-H, Random, 95% CI)	1.08 [0.99, 1.18]

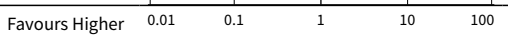
Analysis 3.1. Comparison 3 Serious adverse events - at time point closest to 3 months, Outcome 1 Serious adverse events - at time point closest to three months - highest proportion.

Study or subgroup	Higher n/N	Lower n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Asfar 2017	185/217	165/217		84.47%	1.12[1.02,1.23]
Girardis 2016	80/243	58/235		8.89%	1.33[1,1.78]
Gomersall 2002	2/17	2/17		0.22%	1[0.16,6.3]
Jakkula 2018	20/59	18/61		2.63%	1.15[0.68,1.95]
Lång 2018	6/38	6/27		0.7%	0.71[0.26,1.97]
Panwar 2016	19/51	21/52		3.09%	0.92[0.57,1.5]
Total (95% CI)	**625**	**609**		**100%**	**1.13[1.04,1.23]**
Total events: 312 (Higher), 270 (Lower)					
Heterogeneity: Tau^2=0; Chi^2=2.88, df=5(P=0.72); I^2=0%					
Test for overall effect: Z=2.78(P=0.01)					



Analysis 3.2. Comparison 3 Serious adverse events - at time point closest to 3 months, Outcome 2 Serious adverse events - at time point closest to three months - cumulated.

Study or subgroup	Higher n/N	Lower n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Asfar 2017	196/217	186/217		38.06%	1.05[0.98,1.13]
Girardis 2016	243/243	204/235		43.1%	1.15[1.1,1.21]
Gomersall 2002	2/17	2/17		0.22%	1[0.16,6.3]
Jakkula 2018	21/59	21/61		2.97%	1.03[0.63,1.68]
Lång 2018	6/38	9/27		0.89%	0.47[0.19,1.17]
Panwar 2016	42/51	41/52		14.76%	1.04[0.86,1.26]
Total (95% CI)	**625**	**609**		**100%**	**1.08[0.99,1.18]**
Total events: 510 (Higher), 463 (Lower)					
Heterogeneity: Tau^2=0; Chi^2=9.76, df=5(P=0.08); I^2=48.78%					
Test for overall effect: Z=1.85(P=0.07)					

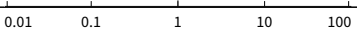


Comparison 4. Sensitivity analysis: serious adverse events - at maximum follow-up

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Serious adverse events - at maximum follow-up - highest proportion	7	1285	Risk Ratio (M-H, Random, 95% CI)	1.13 [1.04, 1.23]
2 Serious adverse events - at maximum follow-up - cumulated	7	1285	Risk Ratio (M-H, Random, 95% CI)	1.07 [0.97, 1.18]

Analysis 4.1. Comparison 4 Sensitivity analysis: serious adverse events - at maximum follow-up, Outcome 1 Serious adverse events - at maximum follow-up - highest proportion.

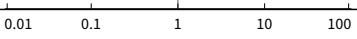
Study or subgroup	Higher n/N	Lower n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Asfar 2017	185/217	165/217		83.78%	1.12[1.02,1.23]
Girardis 2016	80/243	58/235		8.82%	1.33[1,1.78]
Gomersall 2002	2/17	2/17		0.21%	1[0.16,6.3]
Jakkula 2018	20/59	18/61		2.61%	1.15[0.68,1.95]
Lång 2018	9/38	8/27		1.09%	0.8[0.35,1.81]
Mazdeh 2015	5/26	3/25		0.41%	1.6[0.43,6.01]
Panwar 2016	19/51	21/52		3.07%	0.92[0.57,1.5]
Total (95% CI)	**651**	**634**		**100%**	**1.13[1.04,1.23]**
Total events: 320 (Higher), 275 (Lower)					
Heterogeneity: Tau^2=0; Chi^2=3.09, df=6(P=0.8); I^2=0%					
Test for overall effect: Z=2.81(P=0)					



Favours Higher — Favours Lower

Analysis 4.2. Comparison 4 Sensitivity analysis: serious adverse events - at maximum follow-up, Outcome 2 Serious adverse events - at maximum follow-up - cumulated.

Study or subgroup	Higher n/N	Lower n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Asfar 2017	196/217	186/217		36.18%	1.05[0.98,1.13]
Girardis 2016	243/243	204/235		39.86%	1.15[1.1,1.21]
Gomersall 2002	2/17	3/17		0.33%	0.67[0.13,3.5]
Jakkula 2018	21/59	21/61		3.51%	1.03[0.63,1.68]
Lång 2018	15/38	17/27		3.49%	0.63[0.38,1.02]
Mazdeh 2015	5/26	3/25		0.51%	1.6[0.43,6.01]
Panwar 2016	42/51	41/52		16.12%	1.04[0.86,1.26]
Total (95% CI)	**651**	**634**		**100%**	**1.07[0.97,1.18]**
Total events: 524 (Higher), 475 (Lower)					
Heterogeneity: Tau^2=0.01; Chi^2=12.57, df=6(P=0.05); I^2=52.25%					
Test for overall effect: Z=1.41(P=0.16)					



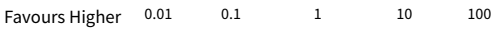
Favours Higher — Favours Lower

Comparison 5. Lung injury - at time point closest to 3 months

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Lung injury - at time point closest to three months - highest proportion	5	1167	Risk Ratio (M-H, Fixed, 95% CI)	1.03 [0.78, 1.36]
2 Lung injury - at time point closest to three months - cumulated	5	1167	Risk Ratio (M-H, Fixed, 95% CI)	0.99 [0.75, 1.30]
3 Lung injury - at time point closest to three months - ARDS	3	288	Risk Ratio (M-H, Random, 95% CI)	0.79 [0.28, 2.20]
4 Lung injury - at time point closest to three months - pneumonia	3	944	Risk Ratio (M-H, Fixed, 95% CI)	1.03 [0.76, 1.40]

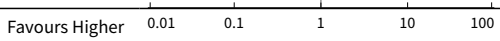
Analysis 5.1. Comparison 5 Lung injury - at time point closest to 3 months, Outcome 1 Lung injury - at time point closest to three months - highest proportion.

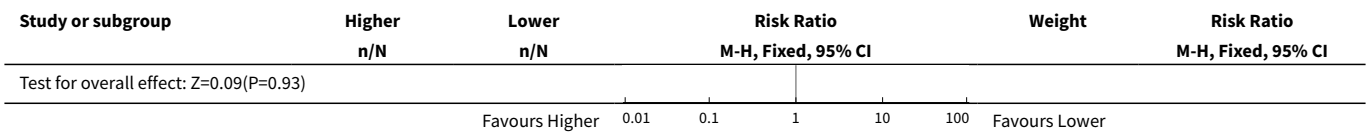
Study or subgroup	Higher n/N	Lower n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Asfar 2017	30/217	32/217		39.39%	0.94[0.59,1.49]
Girardis 2016	37/225	30/220		37.35%	1.21[0.77,1.88]
Jakkula 2018	1/59	1/61		1.21%	1.03[0.07,16.15]
Lång 2018	6/38	6/27		8.64%	0.71[0.26,1.97]
Panwar 2016	11/51	11/52		13.41%	1.02[0.49,2.14]
Total (95% CI)	**590**	**577**		**100%**	**1.03[0.78,1.36]**
Total events: 85 (Higher), 80 (Lower)					
Heterogeneity: Tau^2=0; Chi^2=1.16, df=4(P=0.89); I^2=0%					
Test for overall effect: Z=0.21(P=0.83)					



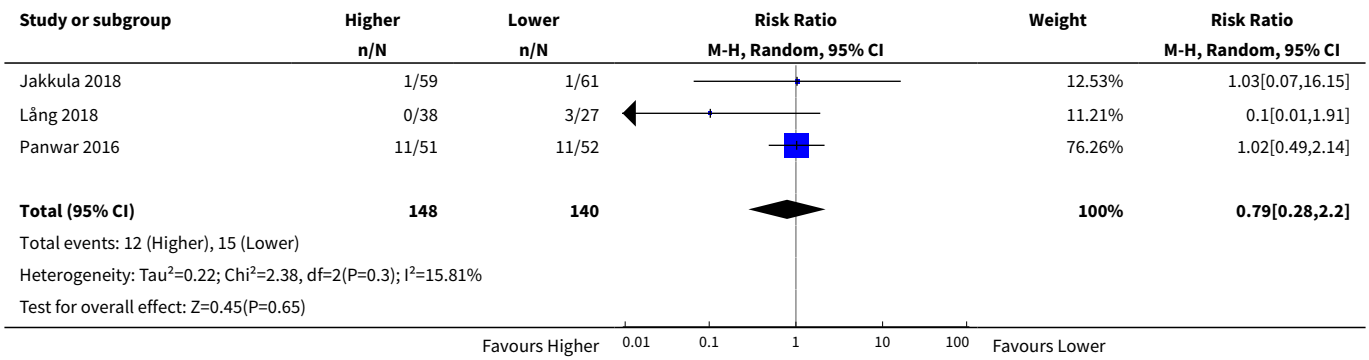
Analysis 5.2. Comparison 5 Lung injury - at time point closest to 3 months, Outcome 2 Lung injury - at time point closest to three months - cumulated.

Study or subgroup	Higher n/N	Lower n/N	Risk Ratio M-H, Fixed, 95% CI	Weight	Risk Ratio M-H, Fixed, 95% CI
Asfar 2017	30/217	32/217		37.76%	0.94[0.59,1.49]
Girardis 2016	37/225	30/220		35.8%	1.21[0.77,1.88]
Jakkula 2018	1/59	1/61		1.16%	1.03[0.07,16.15]
Lång 2018	6/38	9/27		12.42%	0.47[0.19,1.17]
Panwar 2016	11/51	11/52		12.86%	1.02[0.49,2.14]
Total (95% CI)	**590**	**577**		**100%**	**0.99[0.75,1.3]**
Total events: 85 (Higher), 83 (Lower)					
Heterogeneity: Tau^2=0; Chi^2=3.35, df=4(P=0.5); I^2=0%					

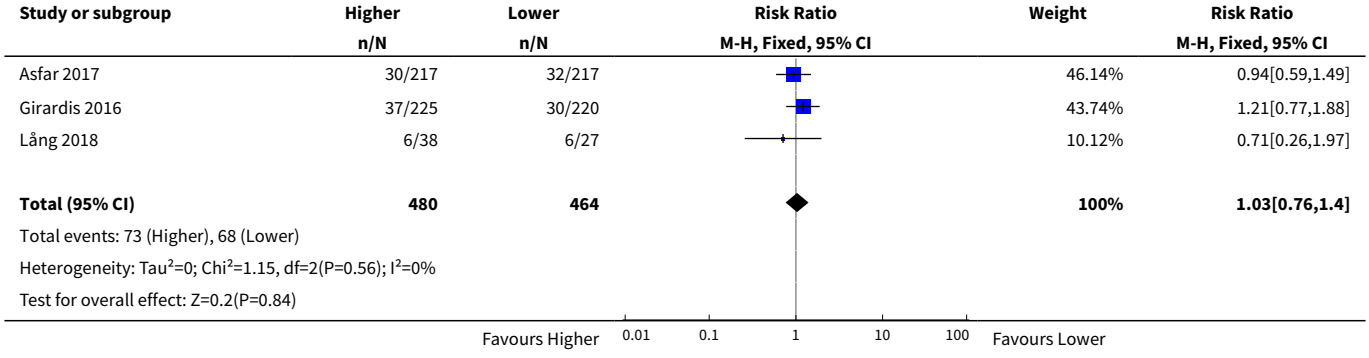




Analysis 5.3. Comparison 5 Lung injury - at time point closest to 3 months, Outcome 3 Lung injury - at time point closest to three months - ARDS.



Analysis 5.4. Comparison 5 Lung injury - at time point closest to 3 months, Outcome 4 Lung injury - at time point closest to three months - pneumonia.



ADDITIONAL TABLES

Table 1. Interventions used in the higher and lower group *(Continued)*

	Higher group			Lower group		
	FiO_2	PaO_2	SaO_2/SpO_2	FiO_2	PaO_2	SaO_2/SpO_2
Asfar 2017	1.00	-	-	-	-	88% to 95%
Girardis 2016	≥ 0.40	≤ 20 kPa (150 mmHg)	97% to 100%	-	9.3 to 13.3 kPa (70 to 100 mmHg)	94% to 98%

Table 1. Interventions used in the higher and lower group *(Continued)*

Gomersall 2002	-	> 9.0 kPa (67.5 mmHg)	-	-	> 6.6 kPa (50 mmHg)	-
Ishii 2018	1.00	-	-	-	100 mmHg (13.3 kPa)	-
Jakkula 2018	-	20 to 25 kPa (150 to 187.5 mmHg)	-	-	10 to 15 kPa (75 to 112.5 mmHg)	95% to 98%
Lång 2018	0.70	-	-	0.40	-	-
Mazdeh 2015	0.50	-	-	Supplemental oxygen not used		
Panwar 2016	-	-	≥ 96%	-	-	88% to 92%
Taher 2016	0.80	-	-	0.50	-	-
Young 2017	No specific measures taken to avoid high FiO_2 or SpO_2, FiO_2 < 0.30 discouraged.			-	-	91% to 96%

FiO_2: fraction of inspired oxygen; **PaO_2**: partial pressure of arterial oxygen; **SaO_2**: arterial oxygen saturation of haemoglobin; **SpO_2**: peripheral oxygen saturation

Table 2. Calculated Bayes factors for the primary outcomes

Outcome	Intervention effect hypothesised	Intervention effect shown by the meta-analysis	Bayes factor (BF)	Interpretation
Mortality *Time point closest to 3 months*	RR 0.80	RR 1.18	18078	*
Mortality *Time point closest to 3 months*	RR 1.20	RR 1.18	0.12 (BF^{-1} = 8.3)	**
Mortality *Maximum follow-up*	RR 0.80	RR 1.16	12867	*
Mortality *Maximum follow-up*	RR 1.20	RR 1.16	0.18 (BF^{-1} = 5.6)	**
Estimated highest reported proportion of serious adverse events *Time point closest to 3 months*	RR 0.80	RR 1.13	2114269	*
Estimated highest reported proportion of serious adverse events *Time point closest to 3 months*	RR 1.20	RR 1.13	0.21 (BF^{-1} = 4.8)	**
Estimated cumulated number of serious adverse events	RR 0.80	RR 1.08	$6.2*10^{20}$	*

Table 2. Calculated Bayes factors for the primary outcomes *(Continued)*

Time point closest to 3 months				
Estimated cumulated number of serious adverse events *Time point closest to 3 months*	RR 1.20	RR 1.08	19 (BF^{-1} = 0.05)	**
Estimated highest reported proportion of serious adverse events *Maximum follow-up*	RR 0.80	RR 1.13	1624463	*
Estimated highest reported proportion of serious adverse events *Maximum follow-up*	RR 1.20	RR 1.13	0.21 (BF^{-1} = 4.8)	**
Estimated cumulated number of serious adverse events *Maximum follow-up*	RR 0.80	RR 1.07	$1.96*10^{19}$	*
Estimated cumulated number of serious adverse events *Maximum follow-up*	RR 1.20	RR 1.07	117 (BF^{-1} = 0.01)	**

Abbreviations: **RR:** risk ratio
*The result is likely BF times more compatible with the null-hypothesis of a relative risk reduction of 0% than the alternative hypothesis of a relative risk reduction of 20% for an effect of higher versus lower supplemental oxygen on all-cause mortality.
**The result is likely BF^{-1} times more compatible with the alternative hypothesis of a relative risk increase of 20% than the null-hypothesis of a relative risk increase of 0% for an effect of higher versus lower supplemental oxygen on all-cause mortality.

APPENDICES

Appendix 1. CENTRAL search strategy

#1 MeSH descriptor: [Hyperoxia] explode all trees
#2 MeSH descriptor: [Anoxia] explode all trees
#3 MeSH descriptor: [Oxygen Inhalation Therapy] explode all trees
#4 MeSH descriptor: [Oxygen] explode all trees
#5 (inspir* or inhal* or fraction* or concentrat* or arterial* or saturation or level* or tension* or supply* or supplement* or supplie* or therap* or administr* or dosag* or dose* or dosing*) near/3 (oxygen):ti,ab,kw
#6 (hyperoxia or hyperoxemia or hyperoxaemia or hypoxia or hypoxemia or hypoxaemia or anoxia or anoxemia or anoxaemia or arterial oxygen or high oxygen or oxygenat* or blood gas or oxygen saturation or pao2 or sao2 or spo2 or fio2):ti,ab,kw
#7 (#1 or #2 or #3 or #4 or #5 or #6)
#8 MeSH descriptor: [Critical Illness] explode all trees
#9 MeSH descriptor: [Critical Care] explode all trees
#10 MeSH descriptor: [Intensive Care Units] explode all trees
#11 MeSH descriptor: [Emergency Medicine] explode all trees
#12 MeSH descriptor: [Emergency Service, Hospital] explode all trees
#13 (emergency department* or ED or emergency room* or ER or high dependency unit* or HDU or prehospital* or critically ill or acutely ill or intensive care or critical care or ICU* or coronary care unit or neurological intermediate care unit):ti,ab,kw
#14 MeSH descriptor: [Heart Arrest] explode all trees
#15 MeSH descriptor: [Myocardial Ischemia] explode all trees
#16 (cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct* or myocardial ischemia or acute coronary syndrome):ti,ab,kw
#17 MeSH descriptor: [Shock] explode all trees
#18 (shock):ti,ab,kw
#19 MeSH descriptor: [Meningitis] explode all trees
#20 (meningitis):ti,ab,kw
#21 MeSH descriptor: [Pneumonia] explode all trees
#22 (pneumonia):ti,ab,kw

#23 MeSH descriptor: [Pulmonary Disease, Chronic Obstructive] explode all trees
#24 (COPD or chronic obstructive pulmonary disease):ti,ab,kw
#25 MeSH descriptor: [Acute Lung Injury] explode all trees
#26 (acute lung injury):ti,ab,kw
#27 MeSH descriptor: [Respiratory Distress Syndrome, Adult] explode all trees
#28 (adult respiratory distress syndrome or ARDS):ti,ab,kw
#29 MeSH descriptor: [Pulmonary Embolism] explode all trees
#30 (pulmonary embolism or pulmonary infarct*):ti,ab,kw
#31 MeSH descriptor: [Multiple Trauma] explode all trees
#32 (severe trauma or multiple trauma):ti,ab,kw
#33 MeSH descriptor: [Craniocerebral Trauma] explode all trees
#34 (traumatic brain injury or TBI or head trauma or craniocerebral trauma):ti,ab,kw
#35 MeSH descriptor: [Stroke] explode all trees
#36 (stroke):ti,ab,kw
#37 MeSH descriptor: [Sepsis] explode all trees
#38 MeSH descriptor: [Shock, Septic] explode all trees
#39 (sepsis or septic shock):ti,ab,kw
#40 MeSH descriptor: [Intracranial Hemorrhages] explode all trees
#41 intracranial hemorrhage or subarachnoid hemorrhage or cerebral hemorrhage or intracranial bleeding or life-threatening bleeding:ti,ab,kw
#42 MeSH descriptor: [Poisoning] explode all trees
#43 (severe poisoning):ti,ab,kw
#44 MeSH descriptor: [Diabetic Ketoacidosis] explode all trees
#45 (diabetic ketoacidosis):ti,ab,kw
#46 MeSH descriptor: [Liver Failure, Acute] explode all trees
#47 (acute hepatic failure or fulminating hepatic failure):ti,ab,kw
#48 MeSH descriptor: [Acute Kidney Injury] explode all trees
#49 (acute kidney failure or acute renal injuries):ti,ab,kw
#50 MeSH descriptor: [Intestinal Perforation] explode all trees
#51 MeSH descriptor: [Appendicitis] explode all
#52 (intestinal perforation or appendicitis):ti,ab,kw
#53 (acute or emergency) near/2 (surgery or operat* or resection):ti,ab,kw
#54 (#8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39 or #40 or #41 or #42 or #43 or #44 or #45 or #46 or #47 or #48 or #49 or #50 or #51 or #52 or #53)
#55 (#7 and #54)

Appendix 2. MEDLINE (OvidSP) search strategy

1. exp Hyperoxia/
2. exp Anoxia/
3. exp Oxygen Inhalation Therapy/
4. exp Oxygen/
5. ((inspir* or inhal* or fraction* or concentrat* or arterial* or saturation or level* or tension* or supply* or supplement* or supplie* or therap* or administr* or dosag* or dose* or dosing*) adj3 oxygen).tw.
6. (hyperoxia or hyperoxemia or hyperoxaemia or hypoxia or hypoxemia or hypoxaemia or anoxia or anoxemia or anoxaemia or arterial oxygen or high oxygen or oxygenat* or blood gas or oxygen saturation or pao2 or sao2 or spo2 or fio2).tw.
7. (1 or 2 or 3 or 4 or 5 or 6)
8. exp Critical Illness/
9. exp Critical Care/
10. exp Intensive Care Units/
11. exp Emergency Medicine/
12. exp Emergency Service, Hospital/
13. (emergency department* or ED or emergency room* or ER or high dependency unit* or HDU or prehospital* or critically ill or acutely ill or intensive care or critical care or ICU* or coronary care unit or neurological intermediate care unit).tw.
14. exp Heart Arrest/
15. exp Myocardial Ischemia/
16. (cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct* or myocardial ischemia or acute coronary syndrome).tw.
17. exp Shock/
18. shock.tw.
19. exp Meningitis/

20. meningitis.tw.
21. exp Pneumonia/
22. pneumonia.tw.
23. exp Pulmonary Disease, Chronic Obstructive/
24. (COPD or chronic obstructive pulmonary disease).tw.
25. exp Acute Lung Injury/
26. acute lung injury.tw.
27. exp Respiratory Distress Syndrome, Adult/
28. (adult respiratory distress syndrome or ARDS).tw.
29. exp Pulmonary Embolism/
30. (pulmonary embolism or pulmonary infarct*).tw.
31. exp Multiple Trauma/
32. (severe trauma or multiple trauma).tw.
33. exp Craniocerebral Trauma/
34. (traumatic brain injury or TBI or head trauma or craniocerebral trauma).tw.
35. exp Stroke/
36. stroke.tw.
37. exp Sepsis/
38. exp Shock, Septic/
39. (sepsis or septic shock).tw.
40. exp Intracranial Hemorrhages/
41. (intracranial hemorrhage or subarachnoid hemorrhage or cerebral hemorrhage or intracranial bleeding or life-threatening bleeding).tw.
42. exp Poisoning/
43. severe poisoning.tw.
44. exp Diabetic Ketoacidosis/
45. diabetic ketoacidosis.tw.
46. exp Liver Failure, Acute/
47. (acute hepatic failure or fulminating hepatic failure).tw.
48. exp Acute Kidney Injury/
49. (acute kidney failure or acute renal injuries).tw.
50. exp Intestinal Perforation/
51. exp Appendicitis/
52. (intestinal perforation or appendicitis).tw.
53. ((acute or emergency) adj2 (surgery or operat* or resection)).tw.
54. (8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53)
55. (7 and 54)
56. randomized controlled trial.pt.
57. controlled clinical trial.pt.
58. randomized.ab.
59. placebo.ab.
60. clinical trial.sh.
61. randomly.ab.
62. trial.ti.
63. (56 or 57 or 58 or 59 or 60 or 61 or 62)
64. exp animals/not humans.sh.
65. (63 not 64)
66. (55 and 65)

Appendix 3. Embase (OvidSP) search strategy

1. *hyperoxia/
2. *hypoxia/
3. *oxygen therapy/
4. *oxygen/
5. *arterial oxygen saturation/
6. *oxygen blood level/
7. *arterial oxygen tension/
8. *blood oxygen tension/
9. ((inspir* or inhal* or fraction* or concentrat* or arterial* or saturation or level* or tension* or supply* or supplement* or supplie* or therap* or administr* or dosag* or dose* or dosing*) adj3 oxygen).tw.

10. (hyperoxia or hyperoxemia or hyperoxaemia or hypoxia or hypoxemia or hypoxaemia or anoxia or anoxemia or anoxaemia or arterial oxygen or high oxygen or oxygenat* or blood gas or oxygen saturation or pao2 or sao2 or spo2 or fio2).tw.
11. (1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10)
12. *critical illness/
13. *intensive care/
14. *intensive care unit/
15. *emergency medicine/
16. *emergency health service/
17. *coronary care unit/
18. (emergency department* or ED or emergency room* or ER or high dependency unit* or HDU or prehospital* or critically ill or acutely ill or intensive care or critical care or ICU* or coronary care unit or neurological intermediate care unit).tw.
19. *heart arrest/
20. *acute heart infarction/
21. (cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct* or myocardial ischemia or acute coronary syndrome).tw.
22. *shock/
23. shock.tw.
24. *meningitis/
25. meningitis.tw.
26. *pneumonia/
27. pneumonia.tw.
28. *chronic obstructive lung disease/
29. (COPD or chronic obstructive pulmonary disease).tw.
30. *acute lung injury/
31. acute lung injury.tw.
32. *adult respiratory distress syndrome/
33. (adult respiratory distress syndrome or ARDS).tw.
34. *lung embolism/
35. (pulmonary embolism or pulmonary infarct*).tw.
36. *multiple trauma/
37. (severe trauma or multiple trauma).tw.
38. *head injury/
39. *brain injury/
40. (traumatic brain injury or TBI or head trauma or craniocerebral trauma).tw.
41. *cerebrovascular accident/
42. *stroke unit/
43. stroke.tw.
44. *sepsis/
45. *septic shock/
46. (sepsis or septic shock).tw.
47. *brain hemorrhage/
48. (intracranial hemorrhage or subarachnoid hemorrhage or cerebral hemorrhage or intracranial bleeding or life-threatening bleeding).tw.
49. *intoxication/
50. severe poisoning.tw.
51. *diabetic ketoacidosis/
52. diabetic ketoacidosis.tw.
53. *acute liver failure/
54. (acute hepatic failure or fulminating hepatic failure).tw.
55. *acute kidney failure/
56. (acute kidney failure or acute renal injuries).tw.
57. *intestine perforation/
58. *appendicitis/
59. (intestinal perforation or appendicitis).tw.
60. ((acute or emergency) adj2 (surgery or operat* or resection)).tw.
61. (12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57 or 58 or 59 or 60)
62. (11 and 61)
63. CROSSOVER PROCEDURE.sh.
64. DOUBLE-BLIND PROCEDURE.sh.
65. SINGLE-BLIND PROCEDURE.sh.
66. (crossover* or cross over*).ti,ab.

67. placebo*.ti,ab.
68. (doubl* adj blind*).ti,ab.
69. allocat*.ti,ab.
70. trial.ti.
71. RANDOMIZED CONTROLLED TRIAL.sh.
72. random*.ti,ab.
73. (63 or 64 or 65 or 66 or 67 or 68 or 69 or 70 or 71 or 72)
74. (exp animal/ or exp invertebrate/ or animal.hw. or nonhuman/) not (exp human/ or human cell/ or (human or humans or man or men or wom?n).ti.)
75. (73 not 74)
76. (62 and 75)

Appendix 4. Science Citation Index - Expanded search strategy

#27 (#26 AND #25)
#26 TOPIC: (((random* OR control* OR RCT OR placebo OR group* OR trial*)))
#25 (#24 AND #3)
#24 (#23 OR #22 OR #21 OR #20 OR #19 OR #18 OR #17 OR #16 OR #15 OR #14 OR #13 OR #12 OR #11 OR #10 OR #9 OR #8 OR #7 OR #6 OR #5 OR #4)
#23 TITLE: (((acute or emergency) and (surgery or operat* or resection)))
#22 TOPIC: ((intestinal perforation or appendicitis))
#21 TOPIC: ((acute kidney failure or acute renal injuries))
#20 TOPIC: ((acute hepatic failure or fulminating hepatic failure))
#19 TOPIC: ((diabetic ketoacidosis))
#18 TOPIC: ((severe poisoning))
#17 TOPIC: ((intracranial hemorrhage or subarachnoid hemorrhage or cerebral hemorrhage or intracranial bleeding or life-threatening bleeding))
#16 TOPIC: ((sepsis or septic shock))
#15 TOPIC: (stroke)
#14 TOPIC: ((traumatic brain injury or TBI or head trauma or craniocerebral trauma))
#13 TOPIC: ((severe trauma or multiple trauma))
#12 TOPIC: ((pulmonary embolism or pulmonary infarct*))
#11 TOPIC: ((adult respiratory distress syndrome or ARDS))
#10 TOPIC: (acute lung injury)
#9 TOPIC: ((COPD or chronic obstructive pulmonary disease))
#8 TOPIC: (pneumonia)
#7 TOPIC: (meningitis)
#6 TOPIC: (shock)
#5 TOPIC: ((cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct* or myocardial ischemia or acute coronary syndrome))
#4 TOPIC: ((emergency department* or ED or emergency room* or ER or high dependency unit* or HDU or prehospital* or critically ill or acutely ill or intensive care or critical care or ICU* or coronary care unit or neurological intermediate care unit))
#3 (#2 OR #1)
#2 TITLE: (((hyperoxia or hyperoxemia or hyperoxaemia or hypoxia or hypoxemia or hypoxaemia or anoxia or anoxemia or anoxaemia or arterial oxygen or high oxygen or oxygenat* or blood gas or oxygen saturation or pao2 or sao2 or spo2 or fio2)))
#1 TITLE: ((((inspir* or inhal* or fraction* or concentrat* or arterial* or saturation or level* or tension* or supply* or supplement* or supplie* or therap* or administr* or dosag* or dose* or dosing*) and oxygen)))

Appendix 5. BIOSIS Previews search strategy

#27 (#26 AND #25)
#26 TOPIC: ((random* OR control* OR RCT OR placebo OR group* OR trial*))
#25 (#24 AND #3)
#24 (#23 OR #22 OR #21 OR #20 OR #19 OR #18 OR #17 OR #16 OR #15 OR #14 OR #13 OR #12 OR #11 OR #10 OR #9 OR #8 OR #7 OR #6 OR #5 OR #4)
#23 TITLE: ((((acute or emergency) and (surgery or operat* or resection))))
#22 TOPIC: (((intestinal perforation or appendicitis)))
#21 TOPIC: (((acute kidney failure or acute renal injuries)))
#20 TOPIC: (((acute hepatic failure or fulminating hepatic failure)))
#19 TOPIC: (((diabetic ketoacidosis)))
#18 TOPIC: (((severe poisoning)))
#17 TOPIC: (((intracranial hemorrhage or subarachnoid hemorrhage or cerebral hemorrhage or intracranial bleeding or life-threatening bleeding)))

#16 TOPIC: (((sepsis or septic shock)))
#15 TOPIC: ((stroke))
#14 TOPIC: (((traumatic brain injury or TBI or head trauma or craniocerebral trauma)))
#13 TOPIC: (((severe trauma or multiple trauma)))
#12 TOPIC: (((pulmonary embolism or pulmonary infarct*)))
#11 TOPIC: (((adult respiratory distress syndrome or ARDS)))
#10 TOPIC: ((acute lung injury))
#9 TOPIC: (((COPD or chronic obstructive pulmonary disease)))
#8 TOPIC: ((pneumonia))
#7 TOPIC: ((meningitis))
#6 TOPIC: ((shock))
#5 TOPIC: (((cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct* or myocardial ischemia or acute coronary syndrome)))
#4 TOPIC: (((emergency department* or ED or emergency room* or ER or high dependency unit* or HDU or prehospital* or critically ill or acutely ill or intensive care or critical care or ICU* or coronary care unit or neurological intermediate care unit)))
#3 (#2 OR #1)
#2 TITLE: (((hyperoxia or hyperoxemia or hyperoxaemia or hypoxia or hypoxemia or hypoxaemia or anoxia or anoxemia or anoxaemia or arterial oxygen or high oxygen or oxygenat* or blood gas or oxygen saturation or pao2 or sao2 or spo2 or fio2)))
#1 TITLE: (((inspir* or inhal* or fraction* or concentrat* or arterial* or saturation or level* or tension* or supply* or supplement* or supplie* or therap* or administr* or dosag* or dose* or dosing*) and oxygen))

Appendix 6. CINAHL search strategy

S66 (S53 AND S65)
S65 (S54 or S55 or S56 or S57 or S58 or S59 or S60 or S61 or S62 or S63 or S64)
S64 TX allocat* random*
S63 (MH "Quantitative Studies")
S62 (MH "Placebos")
S61 TX placebo*
S60 TX random* allocat*
S59 (MH "Random Assignment")
S58 TX randomi* control* trial*
S57 TX ((singl* n1 blind*) or (singl* n1 mask*)) or TX ((doubl* n1 blind*) or (doubl* n1 mask*)) or TX ((tripl* n1 blind*) or (tripl* n1 mask*)) or TX ((trebl* n1 blind*) or (trebl* n1 mask*))
S56 TX clinic* n1 trial*
S55 PT Clinical trial
S54 (MH "Clinical Trials+")
S53 (S7 AND S52)
S52 (S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35 OR S36 OR S37 OR S38 OR S39 OR S40 OR S41 OR S42 OR S43 OR S44 OR S45 OR S46 OR S47 OR S48 OR S49 OR S50 OR S51)
S51 AB ((acute or emergency)) AND AB ((surgery or operat* or resection))
S50 AB (intestinal perforation or appendicitis)
S49 MW Appendicitis
S48 MW Intestinal Perforation
S47 AB (acute kidney failure or acute renal injuries)
S46 MW acute kidney failure
S45 AB (acute hepatic failure or fulminating hepatic failure)
S44 MW Liver Failure, Acute
S43 AB diabetic ketoacidosis
S42 MW Diabetic Ketoacidosis
S41 AB severe poisoning
S40 MW Poisoning
S39 AB (intracranial hemorrhage or subarachnoid hemorrhage or cerebral hemorrhage or intracranial bleeding or life-threatening bleeding)
S38 MW Intracranial Hemorrhage
S37 AB (sepsis or septic shock)
S36 MW Shock, Septic
S35 MW Sepsis
S34 AB stroke
S33 MW Stroke
S32 AB (traumatic brain injury or TBI or head trauma or craniocerebral trauma)

S31 AB (severe trauma or multiple trauma)
S30 MW Multiple Trauma
S29 AB (pulmonary embolism or pulmonary infarct*)
S28 MW Pulmonary Embolism
S27 AB (adult respiratory distress syndrome or ARDS)
S26 MW Respiratory Distress Syndrome
S25 AB acute lung injury
S24 MW Acute Lung Injury
S23 MW (COPD or chronic obstructive pulmonary disease)
S22 MW Pulmonary Disease, Chronic Obstructive
S21 AB pneumonia
S20 MW Pneumonia
S19 AB meningitis
S18 MW Meningitis
S17 AB shock
S16 MW Shock
S15 AB (cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct* or myocardial ischemia or acute coronary syndrome)
S14 MW Myocardial Ischemia
S13 MW heart arrest
S12 AB (emergency department*) or (ED) or (emergency room*) or (ER) or (high dependency unit*) or (HDU) or (prehospital*) or (critically ill) or (acutely ill) or (intensive care) or (critical care) or (ICU*) or (coronary care unit) or (neurological intermediate care unit)
S11 MW emergency medicine
S10 MW intensive care units
S9 MW critical care
S8 MW critical illness
S7 (S1 OR S2 OR S3 OR S4 OR S5 OR S6)
S6 AB (hyperoxia) or (hyperoxemia) or (hyperoxaemia) or (hypoxia) or (hypoxemia) or (hypoxaemia) or (anoxia) or (anoxemia) or (anoxaemia) or (arterial oxygen) or (high oxygen) or (oxygenat*) or (blood gas) or (oxygen saturation) or (pao2) or (sao2) or (spo2) or (fio2)
S5 AB (((inspir*) or (inhal*) or (fraction*) or (concentrat*) or (arterial*) or (saturation) or (level*) or (tension*) or (supply*) or (supplement*) or (supplie*) or (therap*) or (administr*) or (dosag*) or (dose*) or (dosing*))) AND AB (oxygen)
S4 MW oxygen
S3 MW oxygen therapy
S2 MW anoxia
S1 MW hyperoxia

Appendix 7. LILACS search strategy

(tw:((hyperoxia OR hyperoxemia OR hyperoxaemia OR hypoxia OR hypoxemia OR hypoxaemia OR anoxia OR anoxemia OR anoxaemia OR oxygenation OR oxygen OR pao2 OR sao2 OR spo2 OR fio2))) AND (tw:((acute surgery OR acute operation OR acute resection OR emergency surgery OR emergency operation OR emergency resection) OR (intestinal perforation OR appendicitis) OR (acute kidney failure OR acute renal injuries) OR (acute hepatic failure OR fulminating hepatic failure) OR (diabetic ketoacidosis) OR (severe poisoning) OR (intracranial hemorrhage OR subarachnoid hemorrhage OR cerebral hemorrhage OR intracranial bleeding OR life-threatening bleeding) OR (sepsis OR septic shock) OR (stroke) OR (traumatic brain injury OR tbi OR head trauma OR craniocerebral trauma) OR (severe trauma OR multiple trauma) OR (pulmonary embolism OR pulmonary infarction) OR (adult respiratory distress syndrome OR ards) OR (acute lung injury) OR (copd OR chronic obstructive pulmonary disease) OR (pneumonia) OR (meningitis) OR (shock) OR (cardiac arrest OR cardiac failure OR cpr OR heart arrest OR heart failure OR myocardial infarction OR myocardial ischemia OR acute coronary syndrome) OR (emergency department OR ed OR emergency room OR er OR high dependency unit OR hdu OR prehospital OR critically ill OR acutely ill OR intensive care OR critical care OR icu OR coronary care unit OR neurological intermediate care unit))) AND (tw:((randomized OR randomised OR random OR randomly OR controlled OR rct OR placebo OR group OR trial))) AND (instance:"regional") AND (db:("LILACS"))

Appendix 8. Data collection form

TRIAL IDENTIFICATION		
Author and year		
Publication type	**Lead trial:**	**Secondary publ.:** Name of primary publication of the same trial

STUDY ELIGIBILITY												
RCT			Relevant participants			Relevant intervention			Relevant outcomes			
Yes	No	Unclear	Yes	No	Unclear	Yes	No	Unclear	Yes	No*	Unclear	

*Issue relates to selective reporting when study authors may have taken measurements for particular outcomes but did not report these within the paper(s). Review authors should contact trialists for information on possible non-reported outcomes and reasons for exclusion from publication. Study should be listed in 'Studies awaiting assessment' until clarified. If no clarification is received after three attempts, study should be excluded.

DO NOT PROCEED IF ANY OF THE ABOVE ANSWERS IS 'NO'

Include	Exclude
	Record reason for exclusion, which is to be inserted into the 'Table of excluded studies'

PARTICIPANTS	
Eligibility	How was participant eligibility defined?
Age (mean, median, range, etc.)	
Sex of participants (numbers/%, etc.)	
Disease status/type, etc. (if applicable)	
Notes	

INTERVENTIONS	
Experimental intervention	Describe experimental intervention (incl. oxygenation target, oxygen administration system, duration)
Control intervention	Describe control intervention (incl. oxygenation target, oxygen administration system, duration)
Co-interventions (any intervention given equally in both interventions)	Specify any other co-interventions

OTHER TRIAL INFORMATION	
Aim of trial	
Country/Countries	

(Continued)

Trial design (parallel/cross-over)	
Trial duration (intervention and follow-up)	Weeks, months, years, not stated
The trial included only participants admitted to ICU?	
Which targets did the participants actually achieve?	
Withdrawals	Were these described?
Study funding source (Incl. role of funders)	
Possible conflicts of interest (for study authors)	
Other	
Notes	

RISK OF BIAS ASSESSMENT

L: low risk of bias, U: unclear risk of bias, H: high risk of bias

Random sequence generation	
Low risk: if sequence generation is achieved using computer, random number generator or a random numbers table. Drawing lots, tossing a coin, shuffling cards and throwing dice are also adequate if performed by an independent adjudicator. **Unclear risk**: if the method of randomization is not specified. **High risk**: if the allocation sequence is not random.	**Grade** L / U / H
Support for judgement	

Allocation sequence concealment*	
Low risk: if the allocation of participants is performed by a central independent unit, on-site locked computer, identically looking numbered sealed opaque envelopes, drug bottles or containers prepared by an independent investigator. There must be no risk of the investigator knowing the sequence.	**Grade** L / U / H

(Continued)

Unclear risk: if the trial is classified as randomized but the allocation concealment process is not described.

High risk: if the allocation sequence is known to the investigators who assigned participants.

Support for judgement

*Process used to prevent foreknowledge of group assignment in a RCT, which should be seen as distinct from blinding

Blinding of participants and personnel	
Person responsible for participant care	Yes / No
Participant	Yes / No
Outcome assessor	Yes / No
Other (please specify)	Yes / No
Low risk: if the participants and the personnel are blinded to treatment allocation and this is described. **Unclear risk**: if the procedure of blinding is insufficiently described or not described at all. **High risk**: if blinding of participants and personnel is not performed.	**Grade** L / U / H
Support for judgement	

Blinding of outcome assessment	
Low risk: if the trial investigators performing the outcome assessments, analyses and calculations are blinded to the intervention. **Unclear risk**: if the procedure of blinding is insufficiently described or not described at all. **High risk**: if blinding of outcome assessment is not performed.	**Grade** L / U / H
Support for judgement	

Incomplete outcome data	
Low risk: there are no dropouts or withdrawals for all outcomes, or the numbers and reasons for the withdrawals and dropouts for all outcomes are clearly stated and can be described as being similar in both groups. As a general rule the trial is judged as at a low risk of bias due to incomplete outcome data if the number of dropouts is less than 5%. However, the 5% cut-off is not definitive.	**Grade** L / U / H

(Continued)

Unclear risk: the numbers and reasons for withdrawals and dropouts are not clearly stated.
High risk: the pattern of dropouts can be described as being different in the two intervention groups or the trial uses improper methodology in dealing with the missing data, e.g. last observation carried forward.

Support for judgement

Selective outcome reporting

Low risk: a protocol is published before or at the time the trial is begun and the outcome called for in the protocol is reported on.
Unclear risk: if there is no protocol and the outcome is not reported on.
High risk: if the outcomes which are called on in a protocol are not reported on.

Grade

L / U / H

Support for judgement

Baseline imbalance

Low risk: no baseline imbalance in important characteristics was noted.

Unclear risk: baseline characteristics were not reported.

High risk: baseline imbalance was due to chance or was due to imbalanced exclusion after randomization.

Grade

L / U / H

Support for judgement

Early stopping

Low risk: sample size calculation was reported and the trial was not stopped, or if the trial was stopped early by formal stopping rules at a point at which the likelihood of observing an extreme intervention effect due to chance was low.

Unclear risk: sample size calculation was not reported, and if it is not clear whether or not the trial was stopped early.

High risk: the trial was stopped early because of informal stopping rules, or if the trial was stopped early by a formal stopping rule at a point at which the likelihood of observing an extreme intervention effect due to chance was high.

Grade

L / U / H

Support for judgement

Other bias risk	
Low risk: the trial appears to be free of other components (e.g. academic bias or for-profit bias) that could put it at risk of bias. **Unclear risk**: the trial may or may not be free of other components that could put it at risk of bias. **High risk**: there are other factors in the trial that could put it at risk of bias (e.g. authors have conducted trials on the same topic, for-profit bias, etc.)	**Grade** L / U / H
Support for judgement	

Overall risk of bias	
Low risk: each outcome result will be classified as overall 'low risk of bias' only if all of the bias domains described in the above paragraphs are classified as low risk of bias. **High risk**: the outcome result will be classified 'high risk of bias' if any of the bias risk domains described in the above are classified as 'unclear' or 'high risk of bias'. In addition, if one or more of the bias domains described in the above paragraphs are classified as 'unclear' or at 'high risk of bias'.	**Grade** L / H
Support for judgement	

OUTCOMES

PRIMARY OUTCOMES	Available for the trial
All-cause mortality	Yes / No
Number of participants with one or more serious adverse events (dichotomous outcome)	Yes / No
Quality of life	Yes / No

*We used the International Conference on Harmonisation (ICH) Guidelines for Good Clinical Practice's definition of a serious adverse event (ICH-GCP 1997), that is, any untoward medical occurrence that results in death, is life-threatening, requires hospitalization or prolongation of existing hospitalization, or results in persistent or significant disability or incapacity. We will consider all other adverse events as non-serious.

SECONDARY OUTCOMES	Available for the trial
Lung injury*	Yes / No
Acute myocardial infarction**	Yes / No

(Continued)

Stroke**	Yes / No
Severe sepsis**	Yes / No

* Diagnosed after randomization (composite outcome) defined as either ARDS, lung fibrosis, or pulmonary embolism.
** Diagnosed after randomization.

OTHER OUTCOMES OF THE TRIAL	
Additional outcomes	List additional reported outcomes

SUBGROUPS	
Overall risk of bias	High risk of bias
	Low or uncertain risk of bias
According to ICU population	Medical
	Surgical
According to different definitions of oxygen target	Oxygen level measured using FiO_2
	Oxygen level measured using PaO_2
	Oxygen level measured using SaO_2 or SpO_2
	Oxygen level measured using PaO_2 or SaO_2 or SpO_2
According to oxygen delivery system	Invasive mechanical ventilation with endotracheal tube
	Any non-invasive oxygen administration

OUTCOMES						
Follow-up periods	List all follow-up periods given in report					
Total no. of randomized participants	Participants in experimental group			Participants in control group		
Primary outcomes						
(dichotomous 'end point' outcome)		Participants analysed		Number of events in the groups: E = experimental C = control		**Bias of the outcome**
All-cause mortality	Maximum follow-up	E (n)		E (n)		L / U / H
		C (n)		C (n)		
	End of trial intervention period	E (n)		E (n)		L / U / H
		C (n)		C (n)		
Serious adverse events: Nb. Number of counts. If SAE is reported, list them individually	Maximum follow-up	E (n)		E (n)		L / U / H
		C (n)		C (n)		
	End of trial intervention period	E (n)		E (n)		L / U / H
		C (n)		C (n)		
(continuous outcome)		Participants analysed	Mean (endpoint or change)		SD	**Bias of the outcome**
Quality of life: **Type of QoL scale**:	Maximum follow-up	E (n)	E		E	L / U / H
		C (n)	C		C	L / U / H
	End of trial intervention period	E (n)	E		E	L / U / H
		C (n)	C		C	L / U / H

Secondary outcomes				
(dichotomous outcome)		Participants analysed	Number of events in the groups: E = experimental C = control	**Bias of the outcome**
Lung injury	Maximum follow-up	E (n)	E (n)	L / U / H
		C (n)	C (n)	
	End of trial intervention period	E (n)	E (n)	L / U / H
		C (n)	C (n)	
Acute myocardial infarction	Maximum follow-up	E (n)	E (n)	L / U / H
		C (n)	C (n)	
	End of trial intervention period	E (n)	E (n)	L / U / H
		C (n)	C (n)	
Stroke	Maximum follow-up	E (n)	E (n)	L / U / H
		C (n)	C (n)	
	End of trial intervention period	E (n)	E (n)	L / U / H
		C (n)	C (n)	
Severe sepsis	Maximum follow-up	E (n)	E (n)	L / U / H
		C (n)	C (n)	
	End of trial intervention period	E (n)	E (n)	L / U / H
		C (n)	C (n)	

OTHER INFORMATION

Key conclusion of study authors as stated in paper

Information relevant to the results

(Continued)

Indicate if any data were obtained from the primary author; if results were estimated from graphs, etc. or were calculated by you using a formula (should be stated and the formula given). In general, if results not reported in paper(s) are not obtained, this should be made clear here to be cited in the review.

Appendix 9. Criteria for 'Risk of bias' evaluation

Random sequence generation

1. Low risk: if sequence generation is achieved using computer, random number generator, or a random numbers table. Drawing lots, tossing a coin, shuffling cards, and throwing dice are also adequate if performed by an independent adjudicator.
2. Unclear risk: if the method of randomization is not specified.
3. High risk: if the allocation sequence is not random.

Allocation sequence concealment

1. Low risk: if the allocation of participants is performed by a central, independent unit; on-site locked computer; identically appearing, numbered, sealed, opaque envelopes; or drug bottles or containers prepared by an independent investigator. There must be no risk of the investigator knowing the sequence.
2. Unclear risk: if the trial is classified as randomized but the allocation concealment process is not described.
3. High risk: if the allocation sequence is known to the investigators who assigned participants.

Blinding of participants and personnel

1. Low risk: if the participants and personnel are blinded to treatment allocation and this is described.
2. Unclear risk: if the description of the blinding procedure is insufficient or absent.
3. High risk: if blinding of participants and personnel is not performed.

Blinding of outcome assessment

1. Low risk: if the trial investigators performing the outcome assessments, analyses, and calculations are blinded to the intervention.
2. Unclear risk: if the description of the blinding procedure is insufficient or absent.
3. High risk: if blinding of outcome assessment is not performed.

Incomplete outcome data

1. Low risk: there are no dropouts or withdrawals for all outcomes, or the numbers and reasons for withdrawals and dropouts for all outcomes are clearly stated and are described as being similar in both groups. As a general rule, a judgement of low risk of bias is made if the number of dropouts is less than 5%; however, the 5% cut-off is not definitive.
2. Unclear risk: the numbers and reasons for withdrawals and dropouts are not clearly stated.
3. High risk: the pattern of dropouts is described as being different in the two intervention groups, or the trial uses improper methodology in dealing with the missing data, e.g. last observation carried forward.

Selective outcome reporting

1. Low risk: a protocol is published before or at the time the trial is begun, and the outcome called for in the protocol is reported on.
2. Unclear risk: if there is no protocol and the outcome is not reported on.
3. High risk: if the outcomes called for in the protocol are not reported on.

Other bias

1. Low risk: the trial appears to be free of other issues (e.g. academic bias or for-profit bias) that could put it at risk of bias.
2. Unclear risk: the trial may or may not be free of other components that could put it at risk of bias.
3. High risk: there are other factors in the trial that could put it at risk of bias (e.g. authors have conducted trials on the same topic, for-profit bias, etc.).

Overall risk of bias

1. Low risk: the trial will be classified as overall 'low risk of bias' only if all of the 'Risk of bias' domains described above are classified as low risk of bias.
2. High risk: the trial will be classified as overall 'high risk of bias' if any of the 'Risk of bias' domains described above are classified as 'unclear' or 'high risk of bias'.

WHAT'S NEW

Date	Event	Description
27 November 2019	Amended	The ICU-ROX trial (ICU-ROX 2019) was added as a reference awaiting classification. ICU-ROX was published after our literature search was run and thus was not included in this review. The ICU-ROX trial will be included in a review update.

HISTORY

Protocol first published: Issue 4, 2017
Review first published: Issue 11, 2019

Date	Event	Description
27 November 2019	Amended	Author affiliations updated
8 January 2019	Amended	Editorial team changed to Cochrane Emergency and Critical Care
20 September 2017	Amended	We have cited the systematic review Permissive hypoxaemia versus normoxaemia for mechanically ventilated critically ill patients (Gilbert-Kawai 2014).

CONTRIBUTIONS OF AUTHORS

Marija Barbateskovic (MB), Olav L Schjørring (OLS), Sara Russo Krauss, (SRK), Janus C Jakobsen (JJ), Christian S Meyhoff (CM), Rikke M Dahl (RD), Bodil S Rasmussen (BR), Anders Perner (AP), Jørn Wetterslev (JW).

Writing first draft protocol and co-ordinating the protocol: MB

Performing search strategies, searches, and analyses: MB

Literature screening and data extraction: MB, OLS, SRK

Writing first draft review: MB

Writing the review: MB, OLS, SRK, JJ, CM, RD, BR, AP, JW

Person responsible for reading and checking the review before submission: MB

DECLARATIONS OF INTEREST

Marija Barbateskovic: Innovation Fund Denmark provided a grant to Center for Research in Intensive Care (CRIC), which made it possible for Copenhagen Trial Unit as a partner of CRIC to write the review during Marija Barbateskovic's PhD study.

Olav L Schjørring: Oliver's PhD study is funded through a grant from the Innovation Fund Denmark. Furthermore, he is the co-ordinating investigator of the Handling Oxygenation Targets in the Intensive Care Unit (HOT-ICU) trial, a randomized clinical trial comparing a higher versus lower oxygenation target in adult patients with hypoxaemic respiratory failure acutely admitted to the intensive care unit.

Sara Russo Krauss: None known.

Janus C Jakobsen: None known.

Christian S Meyhoff: Dr Meyhoff is the chief investigator for the VitamIn and oXygen Interventions and Cardiovascular Events (VIXIE) trial (a randomized controlled trial comparing perioperative oxygen fractions); site investigator in the HOT-ICU trial (a randomized controlled

trial investigating oxygenation targets in the intensive care unit); co-author of several Cochrane Reviews about oxygen therapy; and was the primary investigator of the PROXI trial (a randomized controlled trial comparing perioperative oxygen fractions).

Rikke M Dahl: None known.

Bodil S Rasmussen: Bodil is the sponsor and primary investigator of a randomized clinical trial comparing a higher versus lower oxygenation target in adult patients with hypoxaemic respiratory failure acutely admitted to the intensive care unit (the Handling Oxygenation Targets in the Intensive Care Unit (HOT-ICU) trial (NCT03174002)).

Anders Perner: Anders's institution receives money for research from Ferring Pharmaceuticals and the Novo Nordisk Foundation

Jørn Wetterslev: Jørn is a member of the task force on Trial Sequential Analysis (TSA) at the Copenhagen Trial Unit, developing and programming TSA (see www.ctu.dk/tsa). I am a supervisor for PhD student Marija Barbateskovic, and the work concerning this review was paid for in part by a grant from Innovation Fund Denmark.

SOURCES OF SUPPORT

Internal sources

- No sources of support supplied

External sources

- Innovation Fund Denmark, Denmark.

  Marija Barbateskovic and Olav L Schjørring are supported by grants from Innovation Fund Denmark, which is a public fund.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

1. We changed the title from 'Higher versus lower inspiratory oxygen fraction or targets of arterial oxygenation for adult intensive care patients' to 'Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the intensive care unit'.
2. We used a power of 90%, not 80% as reported in the protocol (Barbateskovic 2017), as a meta-analysis should use higher (or same) power as its included trials in order to communicate the best available evidence.
3. We changed the wording in the Types of interventions section from "the aim of which was exposure to hyperoxaemia" to "the aim of which was exposure to hyperoxia in the lungs".
4. We added the subgroup 'mixed ICU' to the subgroup analysis (including five trials) of ICU setting, as only one trial included adults admitted to a medical ICU and none to a surgical ICU.
5. In our protocol we stated that we would search the Allied and Complementary Medicine Database (AMED) for eligible trials. We had no access to AMED, and so this search was not conducted.
6. We stated in the 'Types of outcome measures' section of the protocol that we would estimate all continuous and dichotomous outcomes at two time points: the time point closest to three months, which was our assessment time point of primary interest; and at maximum follow-up, as reported by trialists. We realized that this information was confusing. We intended for the assessments at maximum follow-up to be considered as a sensitivity analyses, thus we have specified this in the Sensitivity analysis section.
7. We have now precisely defined the analyses estimating the effect of higher versus lower fraction of inspired oxygen or targets of arterial oxygenation on the proportion of participants with one or more serious adverse events. As the reporting of serious adverse events as a combined outcome was not carried out strictly according to the ICH-GCP recommendation, we estimated the proportion of participants with one or more serious adverse events in a primary analysis: highest proportion of specific serious adverse event reported in each trial. We estimated the effect of higher versus lower inspired fraction or target of oxygen in a sensitivity analysis: the proportion estimated as cumulated number of serious adverse events reported in each trial divided by the number of participants in each intervention group.
8. We have now precisely defined the analyses estimating the effect of higher versus lower fraction of inspired oxygen or targets of arterial oxygenation on the proportion of participants with lung injuries. No trial reported on lung injury as a composite outcome, however some trials reported on ARDS and pneumonia. We estimated the proportion of participants with one or more lung injuries in a primary analysis: highest proportion of specific lung injuries reported in each trial. We estimated the effect of higher versus lower inspired fraction or target of oxygen in a sensitivity analysis: the proportion estimated as cumulated number of lung injuries reported in each trial divided by the number of participants in each intervention group.
9. We changed the wording of the second co-primary outcome (proportion of participants with one or more serious adverse events), without changing the content and implication of the definition.
10. We added a paragraph on Bayes factors in the Methods section. In our protocol, we did not explicitly state that we would present Bayes factors, however we did state that the review would be conducted following the recommendations by Jakobsen and colleagues (Jakobsen 2014a), which include an eight-step assessment involving Bayes factors. In addition, we specified in the Methods section that TSA and calculation of Bayes factors are included in the eight-step assessment.

PAPER V

Higher versus lower levels of oxygen supplementation in critically ill adults. A systematic review with meta-analysis and Trial Sequential Analysis

Marija Barbateskovic[1,2] (ORCID 0000-0001-8566-3660), Olav L. Schjørring[2,3,9] (0000-0002-7749-6003), Sara Russo Krauss[1], Christian S. Meyhoff[4,5] (ORCID 0000-0003-4885-4609), Janus J. Jakobsen[1,6,7,8], Bodil S. Rasmussen[2,3,9], Anders Perner[2,10] (0000-0002-4668-0123), Jørn Wetterslev[1,2] (0000-0001-7778-1771)

[1] Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, University of Copenhagen, Denmark
[2] Centre for Research in Intensive Care, Rigshospitalet, University of Copenhagen, Denmark
[3] Department of Anaesthesia and Intensive Care, Aalborg University Hospital, Denmark
[4] Department of Anaesthesia and Intensive Care, Bispebjerg and Frederiksberg Hospital, University of Copenhagen, Denmark
[5] Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark
[6] Cochrane Hepato-Biliary Group, Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, University of Copenhagen, Denmark
[7] Department of Regional Health Research, the Faculty of Health Sciences, University of Southern Denmark, Denmark
[8] Department of Cardiology, Holbaek Hospital, Denmark
[9] Department of Clinical Medicine, Aalborg University, Aalborg, Denmark
[10] Department of Intensive Care, Rigshospitalet, University of Copenhagen, Denmark

Correspondence
Marija Barbateskovic
Copenhagen Trial Unit, Centre for Clinical Intervention Research, DK2100 Copenhagen, Denmark
Tel: +45 3545 7113
Email: marija.barbateskovic@ctu.dk

Keywords: Oxygen, Critical care, Meta-analysis; Mortality; Systematic review

Words: 3970

ABSTRACT

Background: Liberal oxygen supplementation is often used in critical care but has been associated with harm.

Methods: We conducted an updated systematic review with meta-analysis and Trial Sequential Analysis (TSA) of randomised clinical trials (RCTs) of the effects of higher versus lower levels of oxygen supplementation on all-cause mortality, serious adverse events (SAEs), quality of life, lung injury, sepsis, and cardiovascular events, at timepoint closest to three months in adult critically ill patients.

Results: We included 50 RCTs of 21,014 participants; 36 trials with a total of 20,166 participants contributed data to the analyses. Meta-analysis and TSAs showed no difference between higher and lower levels of oxygen supplementation in mortality and SAEs in trials at overall low risk of bias except for blinding (low certainty evidence): relative risk (RR) 0.98, 95% confidence interval (CI) 0.89-1.09, TSA-adjusted CI 0.86-1.12 and RR 0.99, 95% CI 0.89-1.12, TSA-adjusted CI 0.83-1.19, respectively. The corresponding summary estimates in all trials showed similar results. We did not find a difference between higher and lower levels of oxygen supplementation in meta-analyses and TSAs regarding quality of life, lung injury, sepsis, and cardiovascular events (very low certainty evidence).

Conclusion: We did not find evidence of beneficial or harmful effects of higher versus lower levels of oxygen supplementation in critically ill adults (low to very low certainty evidence). We were able to refute a relative change of 15% in mortality and 20% in SAEs.

Take-home message: There is no evidence for a beneficial or harmful effect on mortality when comparing higher with lower levels of oxygen supplementation in critically ill adults. The effect on mortality, if any, is less than 15%.

Tweet: No evidence of an effect on mortality exist when comparing higher versus lower levels of oxygen supplementation in critically ill adults.

INTRODUCTION

The mainstay treatment and prevention strategy for hypoxaemia is supplemental oxygen, which is frequently used in critical care settings. Despite lack of robust evidence regarding the balance between benefit and harm, oxygen therapy is widely recommended in international practice guidelines [1-5]. Accordingly, clinical practice of oxygen use is often liberal and often results in hyperoxaemia [6-12].

Two meta-analyses of observational studies found an association between hyperoxaemia and mortality in critically ill adults [13,14] and recently a systematic review of randomised clinical trials (RCTs) found an increase in mortality [15] resulting in a recent clinical practice guideline recommending a more restrictive use of oxygen in acutely ill adults medical patients [16].

As new trial data have been published [17], we performed a systematic review comparing the effects of higher versus lower levels of oxygen supplementation in critically ill adults. We hypothesised that higher levels of oxygen supplementation were associated with increased mortality and serious adverse events (SAEs).

METHODS

This systematic review was conducted according to the pre-planned statistical analysis plan of the published protocol [18]. We registered the protocol in the international prospective register of systematic reviews database (PROSPERO) (CRD42017058011), used the methodology of the Cochrane Collaboration supplemented with worst-best case and best-worse case scenarios for participants lost to follow-up, Trial Sequential Analysis (TSA), Bayes factor, and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Electronic Supplementary Material (ESM)) [19,20].

Eligibility criteria

We included RCTs comparing higher versus lower targets of arterial oxygenation (partial pressure of arterial oxygen (PaO_2), arterial oxygen saturation (SaO_2), peripheral oxygen saturation (SpO_2)) or fractions of inspired oxygen (FiO_2) applied by any device in critically ill adults. Both mechanically ventilated and non-mechanically ventilated adults were eligible for inclusion. We included RCTs irrespective of durations of interventions. Quasi randomised trials were excluded.

Outcomes

Predefined co-primary outcomes were all-cause mortality and the proportion of participants with one or more SAEs (composite outcome).

Co-secondary outcomes were: quality of life; severe lung injury (composite outcome) defined as either ALI/ARDS, pulmonary fibrosis or pneumonia, or as defined by trialists; sepsis; and cardiovascular events (composite outcome) defined as either myocardial infarction, stroke, peripheral arterial thrombosis, deep vein thrombosis, pulmonary embolism, or as defined by trialists. Each predefined component of the composite outcome of severe lung injury and cardiovascular events were analysed separately.

For the composite outcomes, we estimated the reported proportion of participants with one or more SAEs, lung injuries and cardiovascular events in two ways:

1. by choosing the one specific event with the highest proportion reported in each trial that addresses the lowest possible proportion of participants with one or more events.
2. by cumulating all reported events, assuming that participants only experience one event (the number of participants in each group will constitute a maximum), address the highest possible reported proportion of participants with one or more events.

For all outcomes, we used the trial results reported at time-points closest to 90 days [18].

Search methods

We searched the following electronic databases: Cochrane Central Register of Controlled Trials (CENTRAL) in the Cochrane Library; Ovid MEDLINE; Ovid Embase; Science Citation Index Expanded (Web of Science); Biosis Previews (Web of Science); and PubMed. Search strategies are presented in the ESM.

The literature search was last updated on 17 October 2019. We manually identified additional potential eligible trials by screening the reference lists of the included studies, other relevant systematic reviews, and searched trial registries.

Trial selection and data extraction

Three review authors (MB, OLS, SRK) independently and in pairs screened titles and abstracts. Reports deemed potentially relevant were obtained in full-text and assessed for inclusion. Disagreements were resolved by consensus and JW were consulted when agreement could not be met.

Three review authors (MB, OLS, SRK) independently and in pairs extracted predefined data of the included trials using a predefined data collection form (ESM).

Risk of bias assessment

MB, OLS and SRK independently and in pairs assessed the risk of systematic errors (bias) of the included trials using the Cochrane Collaboration's risk of bias tool [19]. We planned to present trials as 'overall low risk of bias' when all bias domains were adjudicated as low risk of bias except for blinding of participants and personnel [18]. We post-hoc decided to accommodate the possible challenges of blinding of outcome assessors and presented trials as overall low risk of bias when blinding was not maintained or not reported adequately. Conversely, trials were adjudicated as 'overall high risk of bias' when unclear or high risk of bias was adjudicated in domains other than blinding.

We assessed publication bias by inspecting funnel plots for signs of asymmetry when ten or more trials were included in an analysis [19,21]. We tested asymmetry with the Harbord test [22].

Data synthesis

Summary measures

Risk ratios (RRs) with 95% confidence intervals (CIs) and CIs adjusted for sparse data, multiple outcomes and testing (TSA adjusted CIs) were calculated for dichotomous outcomes. For continuous outcomes, mean-scores were used and mean difference (MD) with CIs and TSA adjusted CIs were calculated.

Meta-analysis

We calculated pooled effect estimates using Review Manager 5 [23]. We used a family wise error rate of 5% and considered a p-value of 0.05/[(2+1)/2] = 0.033 or less as statistically significant in the analyses of each co-primary outcome, and we considered a p-value of 0.05/[(4+1)/2] = 0.02 or less as statistically significant in the analyses of each co-secondary outcome to account for statistical multiplicity due to multiple outcomes [21]. We calculated Bayes factor to assess if the summary effect estimates fitted better with the null hypothesis than alternative hypotheses of the anticipated intervention effects [21].

Dealing with missing data

Corresponding authors were contacted to clarify important missing data related to the methods, data reporting, or if further trial details were needed (ESM).

We conducted a predefined sensitivity analysis by imputing missing outcome data in a best-worst case scenario and a worst-best case scenario to assess the potential impact of loss to follow-up [21,18].

Assessment of heterogeneity

We assessed heterogeneity by visual inspection of the forest plots and calculated the inconsistency statistics (I^2) and the diversity statistics (D^2) [24]. We assessed intervention effects with both random-effects model meta-analyses and fixed-effect model meta-analyses. We used the more conservative point estimate of the two, which is the point estimate clisest to no effect. If the estimates from the two models were approximately equal, we used the estimate with the widest CI [18,21].

Subgroup analyses and sensitivity analyses

We conducted the following predefined subgroup analyses: trials with overall low risk of bias except for blinding versus overall high risk of bias; oxygen level defined by FiO_2 versus oxygen level defined by targets of PaO_2, SaO_2 or SpO_2; low versus high level of oxygenation in control group; subpopulations of critically ill adults; and administration of oxygen below or above median duration of oxygen supplementation.

We conducted a post-hoc subgroup analysis of the effect of supplemental oxygen versus no supplemental oxygen.

Trial Sequential Analysis

We used TSA adjusted CI to assess the uncertainty (risk of random errors) due to sparse data, multiple outcomes, and multiple testing of accumulating data [25-34], and we calculated the required information size [24].

We used a power of 90% (beta 10%) and a diversity as suggested by the trials in the meta-analysis [24,21,34]. As anticipated intervention effects for the primary and secondary outcomes in the TSA, we used a realistic a priori relative risk reduction (RRR) or relative risk increase (RRI) of 20%, and a ¼ SD in Quality of life. We present 95% CI and TSA adjusted CI. For a more detailed description of the statistical analysis plan and TSA, we refer to the published review protocol [18].

Grading certainty of evidence

We used The Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach to assess the overall certainty of evidence for all pre-defined outcomes [35]. We appraised the certainty of the evidence and our confidence in the effect estimates based on risk of bias, inconsistency, indirectness, imprecision and publication bias. Thus, we rated the overall certainty of evidence as high, moderate, low or very low.

RESULTS

Results of the search and selection of trials

We identified 61,852 titles and assessed 327 full text for eligibility (ESM). We included 50 RCTs (of which one was a three-arm trial constituting two trials in this paper) randomising a total of 21,014 participants to higher versus lower levels of oxygen supplementation.

Characteristics of included trials

Fifteen RCTs did not report on any of our outcomes; 35 RCTs contributed with data to the meta-analyses. The number of participants in the trials ranged from 9 to 8003 and all included critically ill adults in different clinical settings (Table 1).

All trials assessed higher versus lower levels of oxygen supplementation using either FiO_2 or arterial oxygenation targets or a combination. However, the definitions of higher and lower levels of oxygen supplementation differed to a great extent between the trials. In the higher groups, FiO_2 ranged from 1.00 to 0.28. In the control groups, 23 trials did not use an FiO_2 or oxygenation target corresponding to our definition of 'low' (FiO_2 below at/or 0.21-0.30 or PaO_2 below at/or 6-8 kPa or SaO_2/SpO^2 below at/or 85%-90%), whilst 17 trials did not apply supplemental oxygen by default. Duration of oxygen administration ranged from 15 minutes to 6 days.

Risk of bias

Five trials were at overall low risk of bias except for blinding of participants and personnel [36,37,17,38,39], and another four trials were at overall low risk of bias except for blinding of participants, personnel and outcome assessors [40-43]. The remaining trials were at overall high risk of bias (ESM). Funnel plots indicated asymmetry but Harbord tests indicated no small-study effect (ESM).

Effect of interventions

All-cause mortality

Thirty-four trials including 19,439 participants reported on all-cause mortality, 8 of these trials (16,156 participants) had overall low risk of bias except for blinding. At follow-up, 1102 of 11,037 (10.0%) participants in the higher levels of oxygen supplementation group had died versus 812/8402 (9.7%) in the lower group (follow-up ranged from 1 day to 1 year). Meta-analysis of 8 trials at overall low risk of bias except for blinding showed no evidence of a difference of higher versus lower oxygen groups (RR 0.98; 95% CI 0.89-1.09; I^2= 0%; TSA-adjusted CI 0.86-1.12; Figure 1-2). The certainty of the evidence, using the GRADE approach, was low (Table 2). The corresponding summary estimate of all 34 trials regardless of risk of bias showed similar results (RR 1.04; 95% CI 0.96-1.13; I^2=2%; TSA-adjusted CI 0.96-1.13; Bayes factor for a 20% RRI=135; Bayes factor for a 20% RRR=37,517,301; Figure 1, ESM). The certainty of evidence was very low (Table 2). Incomplete outcome data had the potential to influence the results (best-worst case scenario: RR 0.80; 95% CI 0.65-0.99 and worst-best case scenario: RR 1.40; 95% CI 1.30-1.52) (ESM). The following tests of interaction showed evidence of a difference: 1) trials with overall low risk of bias except for blinding versus trials with overall high risk of bias (P=0.02). When analysing each subgroup separately, meta-analysis of trials with overall low risk of bias except for blinding showed no evidence of a difference (RR 0.98, 95% CI 0.89-1.098, whilst trials of overall high risk of bias showed evidence of a difference (RR 1.21, 95% 1.05-1.38). 2) trials randomising participants with COPD versus trials randomising participants with other diagnosis than COPD (P=0.09). When analysing each subgroup separately, meta-analyses of trials with COPD did not show a statistically significant difference (RR 2.05, 95% CI

0.94, 4.46); trials of participants with other diagnosis showed no evidence of a difference (RR 1.03, 95% CI 0.95, 1.12). A post-hoc subgroup analysis of the effect of supplemental oxygen versus no supplemental oxygen showed no interaction (P=0.92; ESM). Additional subgroup analyses were consistent with the primary analysis (Figure 3, ESM).

Serious adverse events

Six trials including 8874 participants reported data on the proportion of participants with at least one SAE, 3 of these trials (8056 participants) had overall low risk of bias except for blinding. A total of 924 of 5727 participants (16.1%) in the higher levels of oxygen supplementation group had at least one SAE versus 578 of 3147 (18.4%) in the lower group (follow-up ranged from 1 day to 1 year). Meta-analysis of 3 trials at overall low risk of bias except for blinding showed no evidence of a difference of higher versus lower oxygen groups (RR 0.99; 95% CI 0.89-1.12, I^2=0%; TSA-adjusted 0.83-1.19; Figure 4-5). The certainty of the evidence was low (Table 2). The corresponding summary estimate of all 6 trials regardless of risk of bias showed similar results (RR 1.03; 95% CI 0.95-1.13; I^2=17%, TSA-adjusted CI 0.91-1.18; Bayes factor for a 20% RRI=127; Bayes factor for a 20% RRR=785767; ESM). The certainty of the evidence was low (Table 2).

Thirty-five trials including 19,502 participants reported on single SAEs; 8 of these trials (16,156 participants) had overall low risk of bias except for blinding. Meta-analysis of 8 trials at overall low risk of bias except for blinding showed no evidence of a difference of higher versus lower oxygen groups when assessing the estimated highest reported proportion of specific SAEs in each trial (RR 1.00; 95% CI 0.93-1.08; I^2= 0%; TSA-adjusted CI 0.92-1.09; ESM). The corresponding summary estimate of all 35 trials regardless of risk of bias showed similar results (RR 1.05; 95xs% CI 0.98-1.11; I^2= 1%; TSA-adjusted CI 0.95-1.15; Bayes factor for a 20% RRI=1084; Bayes factor for a 20% RRR=$1.15x10^{14}$; Figure 4), ESM). Meta-analysis of the estimated cumulated number of SAEs showed similar results (RR 1.03; 95% CI 0.98-1.09; I^2=30%; ESM). Results of the subgroup analyses and sensitivity analyses are reportd in the ESM.

Quality of life

Six trials including 7445 participants reported on quality of life using the EuroQoL visual analogue scale (EQ-VAS). Mean scores were 66.1 in the higher levels of oxygen supplementation group versus 64.6 in the lower group (follow-up ranged from 90 to 180 days). Meta-analysis regardless of risk of bias showed no evidence of a difference of higher versus lower oxygen groups (MD 0.37; 95% CI -1.55-2.29; I^2= 57%; TSA-adjusted CI -2.41-3.16; ESM). The certainty of evidence was very low (Table 2).

Lung injury

Ten trials including 9279 participants reported on lung injury. A total of 248 of 5934 participants (4.2%) in the higher levels of oxygen supplementation group developed lung injury versus 227 of 3293 (6.9%) in the lower group (follow-up ranged from 4 to 90 days). Meta-analysis regardless of risk of bias showed no evidence of a difference of higher versus lower oxygen groups when assessing the estimated highest reported proportion of specific lung injury events in each trial (RR 0.93; 95% CI 0.76-1.12; I^2= 0%; TSA-adjusted CI 0.64-1.32; ESM). The certainty of evidence was very low (Table 2). Meta-analysis of the estimated cumulated number of lung injuries showed similar results (RR 0.92; 95% CI 0.78-1.10; I^2=0%; ESM). Meta-analysis showed no evidence of a difference of higher versus lower oxygen groups when assessing ARDS and pneumonia individually (ESM).

Sepsis

Four trials including 1307 participants reported on new onset of sepsis after randomisation. A total of 33 of 649 participaents (5.1%) in the higher levels of oxygen supplementation group developed sepsis versus 20 of 658 (3.0%) in the lower group (follow-up ranged from 6 days to 6 months). Meta-analysis regardless of risk of bias did not show a statistically significant difference of higher versus lower oxygen groups (RR 1.64; 95% CI 0.96-2.80; I^2= 0%; ESM). As only 2.89% of the required information size (n=45,241) had been reached, TSA-adjusted CI could not be calculated. The certainty of evidence was very low (Table 2).

Cardiovascular events

Sixteen trials including 16,615 participants reported on cardiovascular events. A total of 277 of 9580 participants (2.9%) in the higher levels of oxygen supplementation group had a cardiovascular event versus 225 of 7027 (3.2%) in the lower group (follow-up ranged from 1 day to 1 year). Meta-analysis regardless of risk of bias showed no evidence of a difference of higher versus lower oxygen groups when assessing the estimated highest reported proportion of specific cardiovascular events in each trial (RR 1.06; 95% CI 0.86-1.31; I^2= 11%; TSA-adjusted CI 0.45-2.51; ESM). The certainty of evidence was very low (Table 2). Meta-analysis of the estimated cumulated number of cardiovascular events showed similar results (RR 1.10; 95% CI 0.98-1.23; I^2=8%; ESM). Meta-analysis showed no evidence of a difference in myocardial infarction, stroke, deep vein thrombosis, pulmonary embolism individually between the higher versus lower oxygen groups (ESM).

DISCUSSION

In this systematic review we found no evidence of a difference in mortality or SAEs with the use of higher versus lower levels of oxygen supplementation. TSA considering multiple outcomes, sparse data, and repetitive testing, revealed that we obtained the information to reject a 15% relative change in mortality and a 20% relative change in SAEs.

There was no evidence of a 4-point difference in quality of life as measured with EQ-VAS with higher versus lower levels of oxygen supplementation, and TSA revealed that we obtained the information size required to reject such difference.

There was no evidence of a 20% relative change in lung injury, sepsis and cardiovascular events with higher versus lower levels of oxygen supplementation, but the TSA revealed that more data are required. Furthermore, duration of supplemental oxygen for 12 hours or more was not associated with harm as compared to duration of supplemental oxygen of less than 12 hours. And we found no association between the use of a predifined true low level of supplemental oxygen in the control group and the effect of supplemental oxygen.

Strengths and limitations

Our review has several strengths. We included trials regardless of publication type, publication status, language, and choice of outcomes and we contacted relevant trial investigators if additional information was needed. We used predefined, up-to-date systematic review methodology, and the few differences between protocol and review are transparently reported. We used GRADE to assess the certainty of the evidence and TSA with adjusted CI to control the risk of random errors due to multiple outcomes, sparse data, and multiple testing on accumulating data. We assessed the risk of bias of each trial to evaluate the risk of systematic errors (bias) and we used an eight-step procedure to assess if the thresholds for statistical and clinical significance were crossed [21]. We also tested the robustness of our results in sensitivity analyses.

Our review also has several limitations. The primary limitation was that trials did not use the same definition of higher versus lower levels of oxygen supplementation. Some trials used a fixed FiO_2, whilst others used a targeted oxygenation interval, resulting in a large span of oxygenation levels achieved in the trials, in both intervention groups. Furthermore, the oxygenation targets used in some trials can be considered to be overlapping, meaning that these trials might not be categorised as comparing truly high to truly low targets [44,45,17,46,47]. Clinical heterogeneity also included differences in diagnoses and settings. Nonetheless, statistical heterogeneity appeared to be low.

None of the included trials had overall low risk of bias and only two trials were fully blinded [48,49]; this was not unexpected due to the complexity and difficulties of blinding interventions of oxygen supplementation for participants and personnel, and possibly for outcome assessors. Inadequate blinding is therefore a limitation in the included trials, as it is associated with exaggeration of beneficial intervention effects and underestimation of harmful effects [50,51]. We thus cannot rule out a biased effect estimate of the included trials, and as a result, we downgraded the certainty of the evidence one level for all outcomes in GRADE.

To estimate the effects on SAEs, lung injuries and cardiovascular events reported in the included trials, we conducted two supplementary analyses to estimate the effect on the proportion of participants having one or more SAEs, lung injuries and cardiovascular events, which may be expected to lie between the effect estimates of the estimated highest reported proportion and the estimated cumulated number.

Our results in relation to previous reviews

Our systematic review includes twice as many trials as the review by Chu and colleagues indicating increased mortality with higher levels of oxygen supplementation (RR 1.14, 95% CI 1.01-1.28) and rating the evidence as high quality [15]. Our results conflict with those of Chu. First, we found no evidence of a difference on mortality. Second, we do not agree on the certainty of evidence, which we believe should be downgraded for risk of bias and inconsistency, especially the risk of outcome reporting bias seems substantial as 15 trials did not report any of our patient important outcomes. We performed TSA in order to control the risk of random errors in a cumulative meta-analysis and to prevent premature statements regarding the inferiority of higher versus lower levels of oxygen supplementation. TSA was also used by Chu and colleagues, but they did not adjust for multiple outcomes and may have used an inadequate power of 80% [15]. Including more information, we did not find a difference in mortality and were able to exclude a 15% relative change. Recently, we published a systematic review finding an increase in mortality with higher levels of oxygen supplementation in patients admitted to the ICU in the traditional meta-analysis; however, TSA showed that the required information to detect or reject a 20% RRI was not reached and the evidence was very low [52]. The findings of the current review, including results from two recently reported trials conducted in the ICU setting, could not demonstrate evidence of a difference in mortality. This highlights that care should be taken when concluding based on meta-analyses with insufficient information size.

Definitions of critical illness in systematic reviews often differs, and data are analysed and presented in different subgroups; therefore, it may be difficult to consider our assessed subpopulations in relation to other reviews. Our results regarding the lack of a 20% relative change in mortality in patients with acute myocardial infarction support the results of previous systematic reviews [53,54]. We found no effect on mortality in patients randomised prior to hospital admission, in patients admitted to the ICU, in patients with any cerebral disease, in patients with any cardiac disease, in patients with trauma, and in patients with out-of-hospital-cardiac arrest.

Clinical implications and perspectives

We found no evidence supporting the use of either higher or lower levels of oxygen supplementation in critically ill adults. Despite a lack of robust evidence of effectiveness, oxygen administration is widely recommended in international clinical practice guidelines [2,3,55,1]. However, a change towards a more restrictive approach is under way. Based on the results from the systematic reviews by Chu and colleagues [15], clinical practice guidelines are now being updated and revised and now recommend a restrictive oxygenation strategy [16].

We did not find evidence supporting a specific FiO_2 or target of PaO_2, SaO_2 or SpO_2, particularly due to the very high clinical heterogeneity in the types of interventions in the trials included in this review [56,16]. However, it may be worth noticing that almost all the point estimates in our meta-analyses favored a lower level of oxygen supplementation.

With our findings, we cannot reject that higher versus lower levels of oxygen supplementation impacts mortality, but any such effect appears to be below a relative change of 15%. We therefore need more patients randomised into trials with the lowest possible risk of bias to be able to show smaller, but still relevant differences in patient important outcomes with the use of higher versus lower levels of oxygen supplementation.

CONCLUSIONS

The evidence for the use of higher versus lower levels of oxygen supplementation in critically ill patients is of low or very low certainty. Our analyses refuted a relative change of 15% in mortality and 20% in SAEs. The evidence is inconclusive regarding smaller effects of higher versus lower levels of oxygen supplementation on mortality, SAEs, quality of life, lung injury, sepsis and cardiovascular events because too few participants have been randomised. Thus, more patients should be randomised in trials with the lowest possible risk of bias.

TABLES AND FIGURES

Table 1. Characteristics of included trials

					Interventions						
	Trial/comparison	**Country**	**Setting**	**Sample size**	**Duration, h**	**Higher group**			**Lower group**			**Maximum follow-up**
						FiO_2/O_2 flow*	**PaO_2**	**SaO_2/SpO_2**	**FiO_2/O_2 flow***	**PaO_2**	**SaO_2/SpO_2**	
1	Ali 2013 [40]	UK	Stroke	301	72	2 L/min by nasal cannula if baseline SpO_2 > 93% and 3 L/min if baseline SpO_2 ≤ 93%			No supplemental oxygen			6 months
2	Asfar 2017 [57]	France	Septic shock, invasively mechanically ventilated	442	24	1.00					88-95%	90 days
3	Austin 2010 [58]	Australia	AECOPD	405	Pre-hospital transport (mean 47 min)	8-10 L/min by non-rebreather facemask					88-92%	In-hospital
4	Baekgaard 2019 [59]	Denmark	Trauma	41	24	15 L/min by non-rebreather facemask and FiO_2 of 1.00 (or 0.80 if SpO_2 ≥ 98%) when mechanically ventilated					94%	30 days
5	Bardsley 2018 [60]	New Zealand	AECOPD	90	0.25	8 L/min by nebulisation mask			No supplemental oxygen (air 8 L/min by nebulisation mask)			-

6	Bickel 2011 [61]	Israel	Acute appendicitis	210	2	0.80 peroperatively, postoperatively 10 L/min by non-rebreather facemask			0.30 peroperatively, postoperatively 4 L/min by nasal cannula			14 days
7	Bray 2018 [62]	Australia	Cardiac arrest	62	Pre-hospital transport (mean 50 min)	1.00			2-4 L/min via bag-valve mask		≥ 94% (≥ 90% in amended protocol)	In-hospital
8	Butler 1987A *Skin oxygen study* [63]	UK	Limb ischaemia /amputation	20	48	0.28			No supplemental oxygen			14 days
9	Butler 1987B *Healing study* [63]	UK	Limb ischaemia /amputation	39	48	0.28			No supplemental oxygen			1 year
10	Girardis 2016 [44]	Italy	Critical care	480	ICU stay (median 144)	≥ 0.40	≤ 20 kPa (150 mmHg)	97%-100%		9.3-13.3 kPa (70-100 mmHg)	94%-98%	60 days
11	Gomersall 2002 [45]	Hong Kong	AECOPD	36	Length of hospital stay (median 144)		> 9.0 kPa (67.5 mmHg)			> 6.6 kPa (50 mm Hg)		In-hospital
12	Heidari 2017 [64]	Iran	Acute coronary syndrome	79	6	4-6 L/min by nasal cannula			No supplemental oxygen			In-hospital
13	Hofmann 2017 [37,65]	Sweden	Myocardial infarction	6629	6-12 (IQR 11.64)	6 L/min by open facemask			No supplemental oxygen (unless SpO_2 < 90%)			1 year
14	Huynh Ky 2017 [66]	Canada	Acute coronary syndrome	39	Maximum 24 (mean 12)			97%			92%	Not specified
15	ICU-ROX investigators 2019 [17]	New Zealand	Critical care, mechanically ventilated	1000	ICU admission, maximum 672	Conventional oxygen administration (FiO_2 < 0.30 discouraged during mechanical ventilation)					91-96%	180 days

					(median 120)							
16	Ishii 2018 [67]	Japan	Critical care, invasively mechanically ventilated	51	Until first analysis of arterial blood sampling	1.00				100 mmHg (13.3 kPa)		3 days
17	Jakkula 2018 [39]	Finland	Cardiac arrest	123	36		20-25 kPa (150-187.5 mmHg)			10–15 kPa (75-112.5 mmHg)	95%-98%	6 months
18	Jun 2019 [46]	-	AECOPD and myocardial infarction, invasively mechanically ventilated	58	-	0.50-0.70 for the first 48 hours, hereafter 0.40-0.50			0.30-0.50			-
19	Khoshnood 2018 [68,69]	Sweden	Myocardial infarction	160	Pre-hospital transport and PCI (mean 1.4)	10 L/min by open facemask			No supplemental oxygen (unless SpO_2 < 94%)			6 months
20	Kuisma 2006 [70]	Finland	Cardiac arrest	32	1	1.00			0.30		≥ 95%	In-hospital
21	Lång 2018 [71]	Finland	Traumatic brain injury	70	Mechanical ventilation, maximum 336 (mean 136)	0.70			0.40			6 months
22	Mazdeh 2015 [72]	Iran	Stroke	52	12	0.50			No supplemental oxygen			6 months
23	Meyhoff 2009 [36]	Denmark	Acute abdominal surgery	385	2 (postop)	0.80 peroperatively, postoperatively 0.80 by non-			0.30 peroperatively, postoperatively 0.30 by non-			3 months

						rebreather facemask			rebreather facemask			
24	NCT02378545 [73]	UK	Sepsis	50	ED stay	15 L/min by non-re-breather facemask					94%	90 days
25	NCT02687217 [47]	India	Acute appendicitis	60	2	≥ 0.50 peroperatively, 0.31 postoperatively			0.21 peroperatively, 0.28 postoperatively			-
26	Padma 2010 [74]	India	Stroke		12	10 L/min by open facemask			No supplemental oxygen or up to 2 L/min by open facemask		≥ 95%	3 months
27	Panwar 2016 [41]	Australia, New Zealand, France	Critical care, invasively mechanically ventilated	104	Mechanical ventilation (median 114)			≥ 96%			88-92%	90 days
28	Perrin 2011 [42]	New Zealand	Acute exacerbation of asthma	106	1	8 L/min by open facemask					93-95%	1 h
29	Ranchord 2012 [75]	New Zealand, UK	Myocardial infarction	148	6	6 L/min by open facemask. Concentrations were delivered		≥ 92%			93-96%	30 days
30	Rawles 1976 [48]	UK	Myocardial infarction	200	24	6 L/min by open facemask			No supplemental oxygen (air 6 L/min by open facemask)			In-hospital
31	Rodrigo 2003 [76]	Uruguay	Acute exacerbations of asthma	77	0.33	1.00 oxygen by non-rebreather facemask			0.28 by open facemask			20 min
32	Rodrigues de Freitas Vianna 2017 [77]	Brazil	Critical care, invasively mechanically ventilated		Endotracheal suctioning procedure	1.00			0.20 above baseline FiO_2			30 min
33	Roffe 2010 [78]	UK	Stroke	63	12 (nocturnally)	2 L/min via nasal cannula			No supplemental oxygen			14 days

34	Roffe 2017A *Continuous oxygen* [38]	UK	Stroke	4002	72	3 L/min by nasal cannula if baseline $SpO_2 \leq 93\%$ and 2 L/min if baseline $SpO_2 > 93\%$			No supplemental oxygen			90 days
35	Roffe 2017B *Nocturnal oxygen* [38]	UK	Stroke	4001	10 x 3 (nocturnally)	3 L/min if baseline $SpO_2 \leq 93\%$ or less and 2 L/min if baseline > 93%			No supplemental oxygen			90 days
36	Sepehrvand 2019 [79]	Canada	Acute heart failure	50	72			≥ 96%			90-92%	30 days
37	Shi 2017 [80]	China	Stroke	18	4	10 L/min by open facemask			No supplemental oxygen			7 days
38	Sills 2003 [81]	UK	Stroke	25	8 (nocturnally)	2 L/min by nasal cannula			No supplemental oxygen			3 days
39	Singhal 2005 [82]	US	Stroke	16	8	45 L/min by open facemask			0-3 L/min by nasal cannula		≥ 96%	3 months
40	Singhal 2013 [83]	US	Stroke	85	8	30-45 L/min by open facemask			No supplemental oxygen (air 30-45 L/min by open facemask)			3 months
41	Stewart 2019 [84]	New Zealand	Acute coronary syndrome		-			≥ 95%			90-94%	1 year
42	Stub 2015 [85]	Australia	Myocardial infarction	638	Pre-hospital transport and PCI (mean 1.09)	8 L/min by open facemask					94%	6 months
43	Taher 2016 [86]	Iran	Traumatic brain injury		6	0.80			0.50			6 months
44	Thomas 2019 [87]	UK	Cardiac arrest	35	1	1.00					94-98%	90 days
45	Ukholkina 2005 [88]	Russia	Myocardial infarction		3.5	0.40-0.60			No supplemental oxygen			-
46	Wijesinghe 2012 [43]	New Zealand	Pneumonia	150	1	8 L/min by open facemask					93-95%	1 hour
47	Wilson 1997 [89]	UK	Myocardial infarction	50	24	4L/min by open facemask			No supplemental oxygen			-

48	Wu 2014 [90]	China	AECOPD	9	0.25	group B: 6–7 L/min by nebulisation mask, group C: 8–9 L/min by nebulisation mask			group A: 4–5 L/min by nebulisation mask			30 minutes
49	Young 2014 [91]	New Zealand	Cardiac arrest	18	72	1.00 prehospitally, conventional oxygen administration in ED and ICU		> 95% (suggested in ED and ICU)			90-94%	72 h
50	Zughaft 2013 [49]	Sweden	Stable angina or acute coronary syndrome	304	PCI	3 L/min by nasal cannula			No supplemental oxygen (air 3 L/min by nasal cannula)			1 year

*The specific FiO_2 is stated when delivered by mechanical ventilation, bag-valve mask (with flow ≥ 10 L/min), or venturi masks, unless otherwise specified

Table 2. Summary of findings

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Higher	Lower	Relative (95% CI)	Absolute (95% CI)		
All-cause mortality - overall low risk of bias except for blinding												
8	randomised trials	serious [a]	not serious	serious [b]	not serious [c]	none	798/9362 (8.5%)	562/6794 (8.3%)	**RR 0.98** (0.89 to 1.09)	**2 fewer per 1,000** (from 9 fewer to 7 more)	⨁⨁◯◯ LOW	CRITICAL
All-cause mortality - All trials												
34	randomised trials	serious [d]	not serious	serious [b]	not serious [e]	publication bias strongly suspected [f]	1102/11037 (10.0%)	812/8402 (9.7%)	**RR 1.04** (0.96 to 1.13)	**4 more per 1,000** (from 4 fewer to 13 more)	⨁◯◯◯ VERY LOW	CRITICAL
Serious adverse events - overall low risk of bias except for blinding												
3	randomised trials	serious [a]	not serious	serious [b]	not serious [g]	none	705/5313 (13.3%)	387/2743 (14.1%)	**RR 0.99** (0.89 to 1.12)	**1 fewer per 1,000** (from 16 fewer to 17 more)	⨁⨁◯◯ LOW	CRITICAL
Serious adverse events - All trials												
6	randomised trials	serious [h]	not serious	serious [b]	not serious [i]	none	924/5727 (16.1%)	578/3147 (18.4%)	**RR 1.03** (0.95 to 1.13)	**6 more per 1,000** (from 9 fewer to 24 more)	⨁⨁◯◯ LOW	CRITICAL
Quality of life - All trials												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Higher	Lower	Relative (95% CI)	Absolute (95% CI)		
6	randomised trials	serious [j]	serious [k]	serious [b]	not serious	none	4851	2594	-	MD **0.37 higher** (1.55 lower to 2.29 higher)	⨁◯◯◯ VERY LOW	IMPORTANT
Lung injury - All trials												
10	randomised trials	serious [l]	not serious	serious [b]	serious [m]	publication bias strongly suspected [f]	248/5934 (4.2%)	172/3293 (5.2%)	**RR 0.92** (0.76 to 1.11)	**4 fewer per 1,000** (from 13 fewer to 6 more)	⨁◯◯◯ VERY LOW	IMPORTANT
Sepsis - All trials												
4	randomised trials	serious [n]	not serious	serious [b]	serious [o]	none	33/649 (5.1%)	20/658 (3.0%)	**RR 1.64** (0.96 to 2.80)	**19 more per 1,000** (from 1 fewer to 55 more)	⨁◯◯◯ VERY LOW	IMPORTANT
Cardiovascular events - All trials												
16	randomised trials	serious [p]	not serious	serious [b]	serious [q]	publication bias strongly suspected [f]	277/9580 (2.9%)	225/7027 (3.2%)	**RR 1.06** (0.86 to 1.31)	**2 more per 1,000** (from 4 fewer to 10 more)	⨁◯◯◯ VERY LOW	IMPORTANT

CI: Confidence interval; **RR:** Risk ratio; **MD:** Mean difference

Explanations

[a] Participants and personnel and/or outcome assessors were not blinded.
[b] Differences in inspiratory oxygen fraction and target of arterial oxygenation in the experimental and control groups between trials.
[c] Required information size to reject a relative change of 20% was reached. Futility was reached for a relative change of 15%.
[d] 26/34 trials were overall at high risk of bias.

[e] Required information size to reject a relative change of 20% was reached. Futility was reached for a relative change of 15%.
[f] Funnel plot indicated asymmetry; however, Harbord test indicated no small-study effects.
[g] Required information size to reject a relative change of 20% was reached.
[h] 27/35 trials were at overall high risk of bias.
[i] Required information size to reject a relative change of 20% was reached.
[j] 2/6 trials were overall high risk of bias.
[k] I^2=57% (P=0.04), Signs of heterogeneity in forest plot.
[l] 5/10 trials were overall high risk of bias.
[m] Required information size to detect/reject a relative change of 20% was not reached.
[n] 3/4 trials were overall high risk of bias.
[o] Only 2.89% of the required information size was reached.
[p] 5/16 trials were overall high risk of bias.
[q] Only 25.72% of the required information size was reached.

Figure 1. Forest plot on mortality in trials with overall low risk of bias except for blinding versus trials with overall high risk of bias. Size of squares for risk ratio reflects weight of trial in pooled analysis. Horizontal bars represent 95% confidence interval.

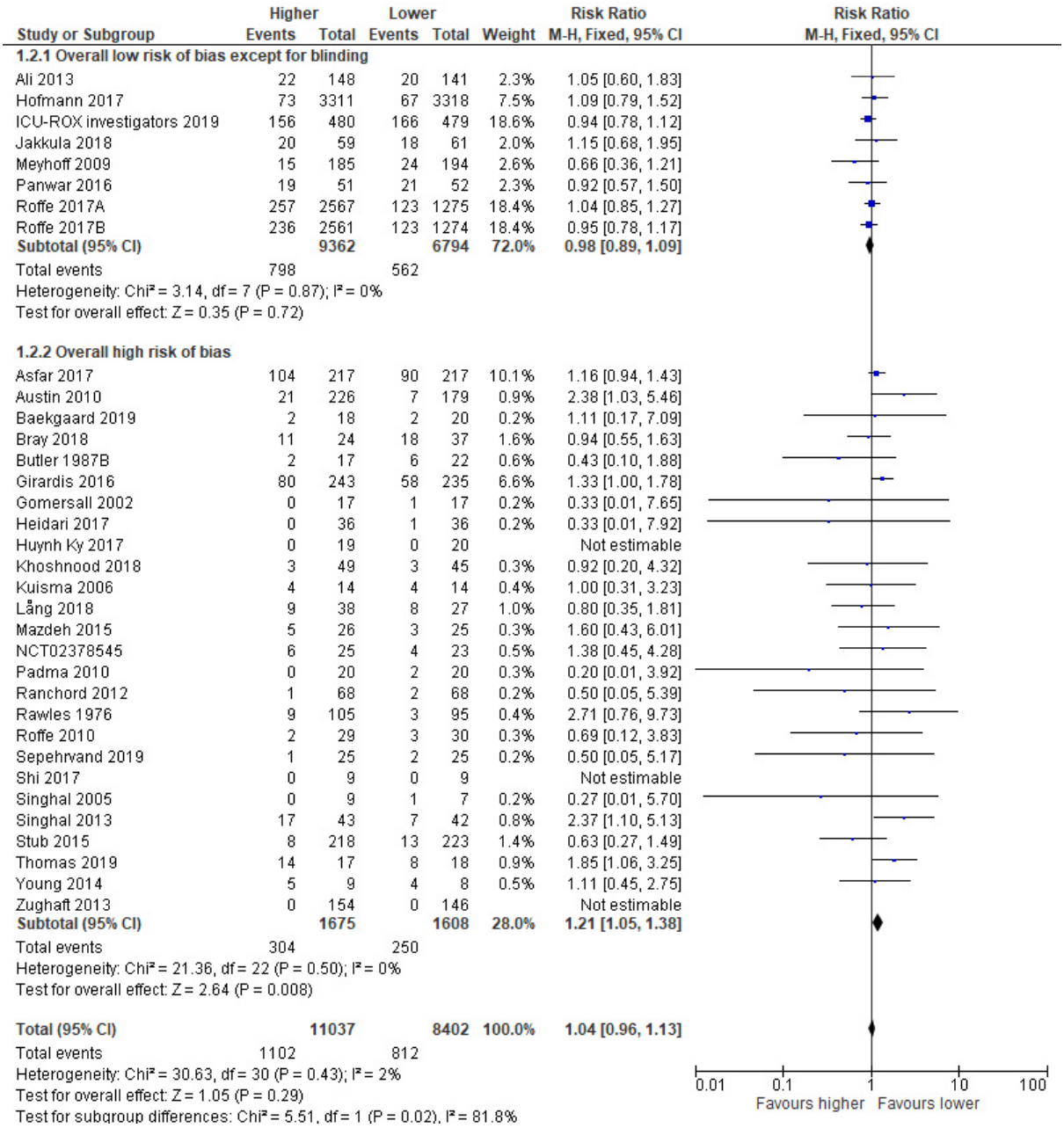


Figure 2. Trial sequential analysis of overall low risk of bias except for blinding trials of the effect of higher versus lower oxygen supplementation on mortality using an alpha of 3.3%, a power of 90%, control event proportion of 8.27% (from the included trials), a diversity (D2) of 0%, and a relative risk increase of 15%. The relative risk was 0.98 with a TSA-adjusted CI 0.86-1.12. Futility was reached, suggesting that a relative change of 15% can be excluded.

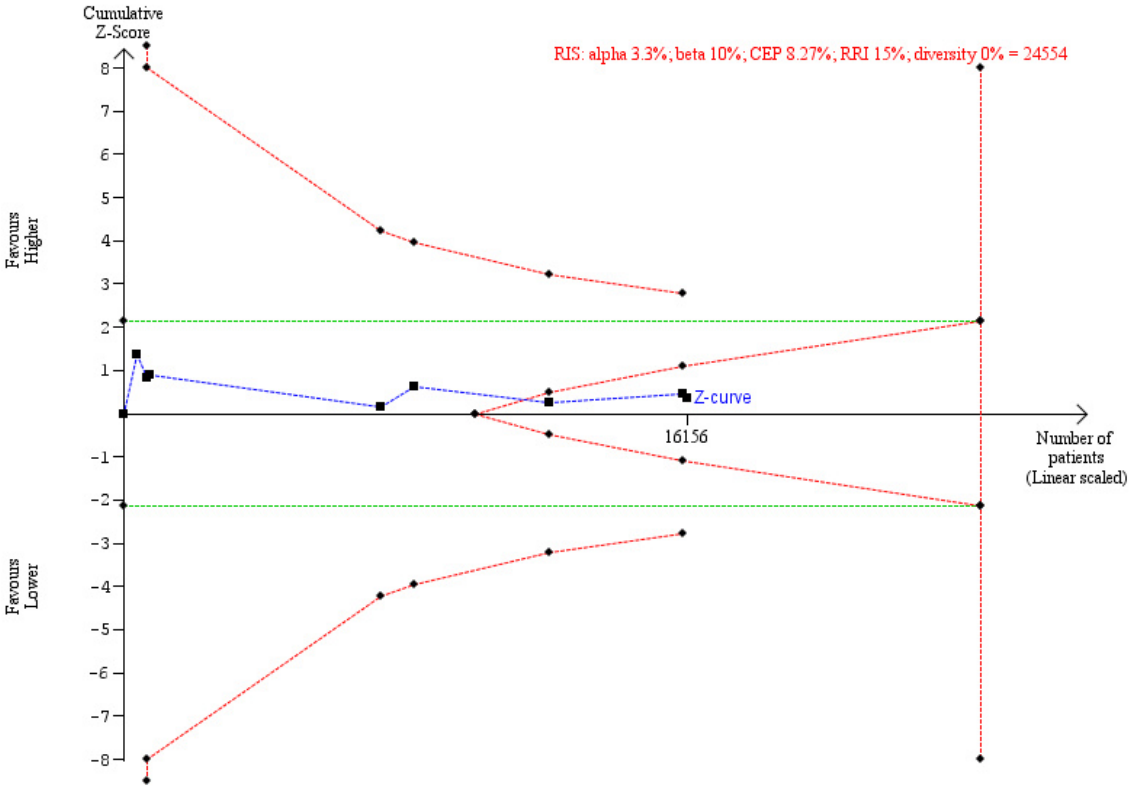


Figure 3. Forest plot on mortality stratified by population group. Size of squares for risk ratio reflects weight of trial in pooled analysis. Horizontal bars represent 95% confidence interval.

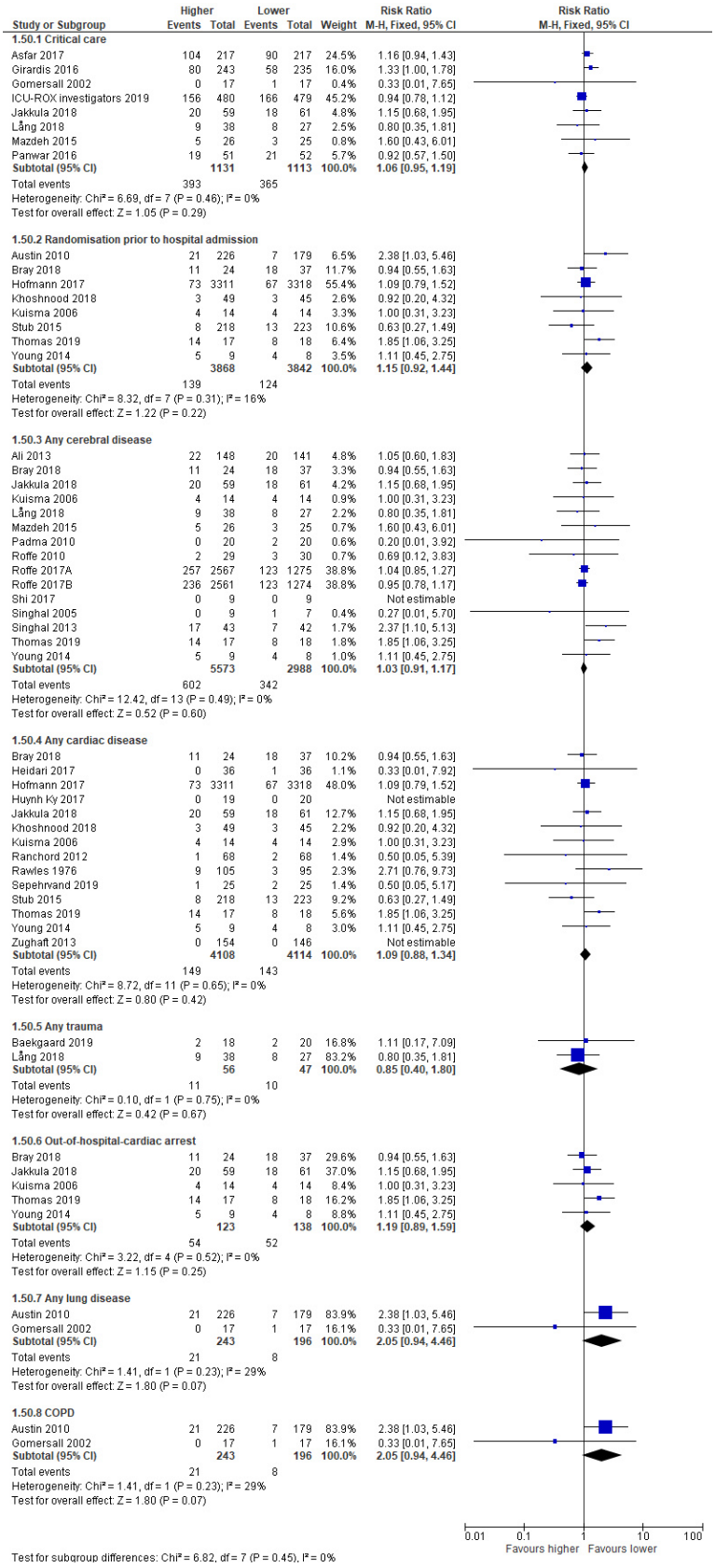


Figure 4. Forest plot on the proportion of participants with at least one serious adverse event, as reported by trialists, in trials with overall low risk of bias except for blinding versus trials with overall high risk of bias. Size of squares for risk ratio reflects weight of trial in pooled analysis. Horizontal bars represent 95% confidence interval.

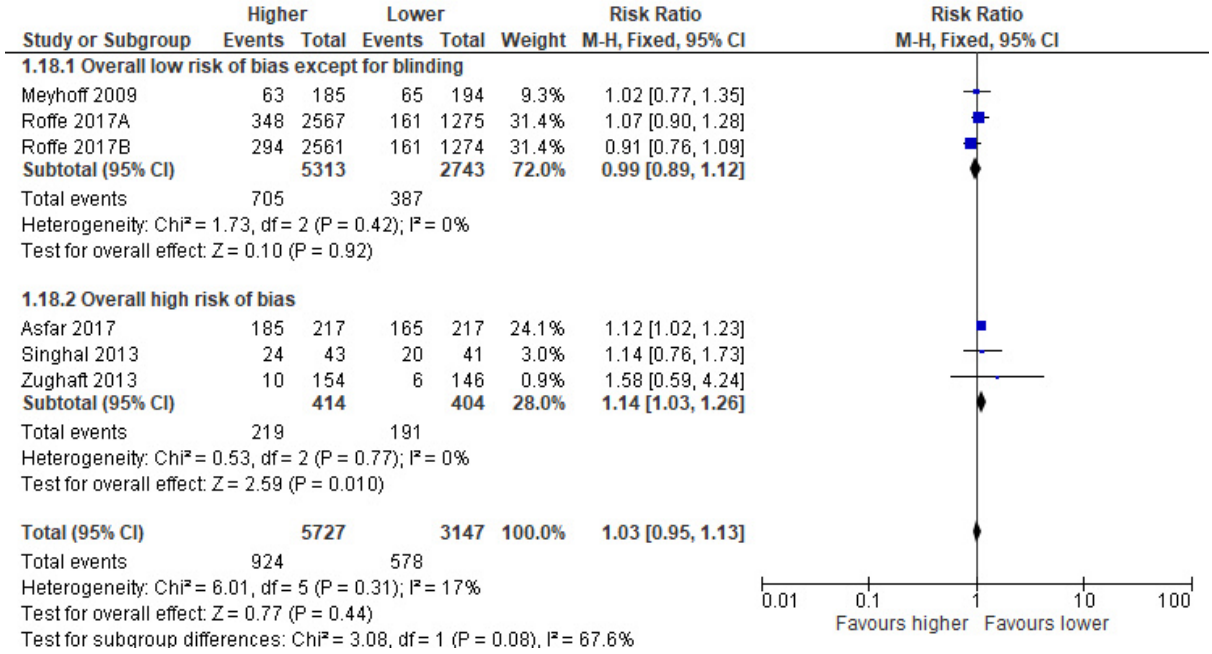
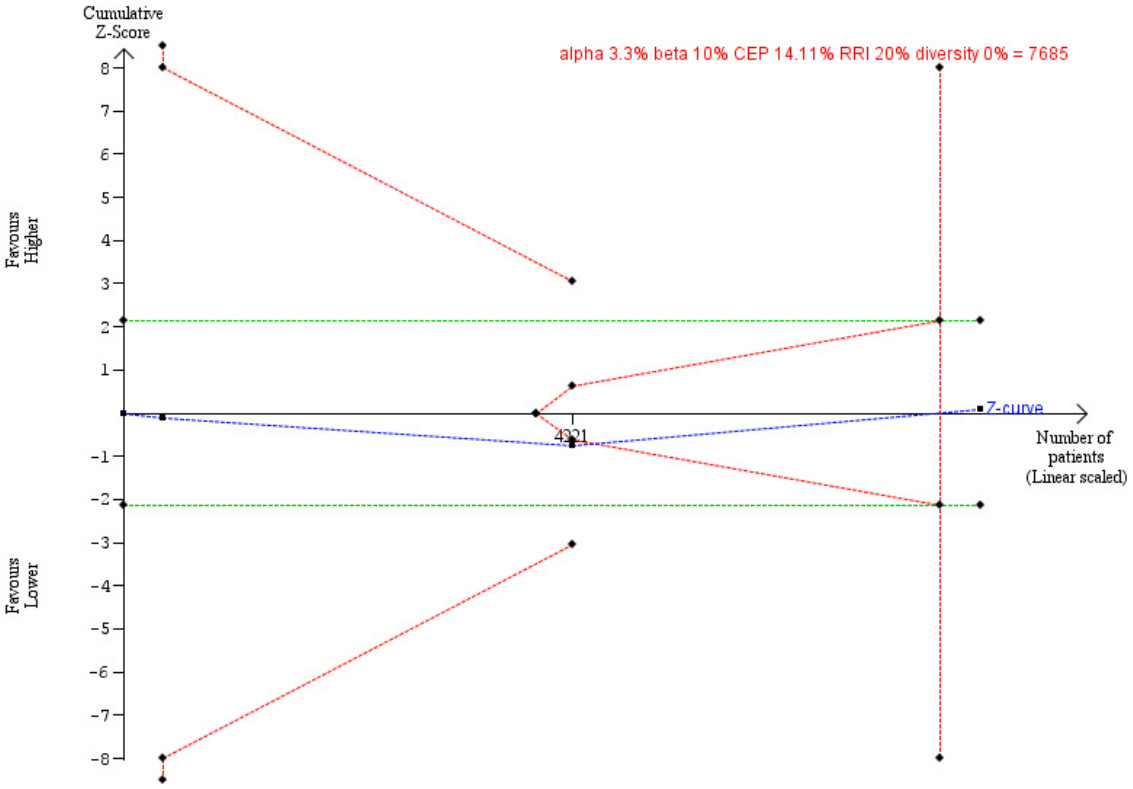


Figure 5. Trial sequential analysis of overall low risk of bias except for blinding trials of the effect of higher versus lower oxygen supplementation on the proportion of participants with at least one serious adverse event, as reported by trialists using an alpha of 3.3%, a power of 90%, control event proportion of 14.11% (from the included trials), a diversity (D2) of 0%, and a relative risk increase of 20%. The relative risk was 0.99 with a TSA-adjusted CI 0.83-1.19. Required information size was reached, suggesting that a relative change of 20% can be excluded.



DECLARATIONS

Funding

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors. M Barbateskovic' and OL Schjørrings' contributions to this work, were supported by Innovation fund Denmark [grant number 4108-00011B].

Author contributions

All authors contributed to the study protocol. Search strategy was built by MB who also performed the literature search. MB, OLS and SRK performed the literature screening, data extraction and risk of bias evaluation. MB conducted the analyses. The first draft of the manuscript was written by MB and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Conflicts of interest

MB, SRK, JCJ: None known. OLS is a member of the Management Committee of the HOT-ICU (Handling Oxygenation Targets in the Intensive Care Unit) trial, investigating higher versus lower oxygenation targets in patients admitted to the ICU. CSM reports direct and indirect departmental research funding from Ferring Pharmaceuticals, Radiometer, Merck, Sharp & Dohme Corp., and Boehringer Ingelheim as well as lecture fee from Radiometer outside the submitted work. CSM was the principal investigator of the PROXI trial (PeRioperative OXygen fraction – effect on surgical site Infection and pulmonary complications after abdominal surgery) investigating higher versus lower levels of perioperative inspiratory oxygen. Furthermore, he is a principle site investigator of the HOT-ICU trial and sponsor for the VIXIE trial investigating perioperative inspiratory oxygen (NCT03494387). BSR is the sponsor and primary investigator of the HOT-ICU trial. AP is a member of the Management Committee of the HOT-ICU trial. JW is a member of the task at Copenhagen Trial Unit (CTU) to develop theory and software for doing Trial Sequential Analysis (TSA) available as freeware including a comprehensive manual at www.ctu/tsa.

ADDITIONAL FILES

Additional file: Electronic supplementary material

REFERENCES

1. O'Driscoll BR, Howard LS, Earis J, Mak V, British Thoracic Society Emergency Oxygen Guideline Group, Bts Emergency Oxygen Guideline Development Group (2017) BTS guideline for oxygen use in adults in healthcare and emergency settings. Thorax 72 (Suppl 1):ii1-i90

2. Kallstrom TJ, American Association for Respiratory Care (2002) AARC clinical practice guideline. Oxygen therapy for adults in the acute care facility - 2002 revision & update. Respiratory Care 47 (6):717-720

3. Australian Resuscitation Council (2015) Guideline 11.6.1. Targeted oxygen therapy in adult advanced life support, 2014. resus.org.au/download/section_11/anzcor-guideline-11-6-1-targeted-oxygen-therapy-jan16.pdf.

4. Beasley R, Chien J, Douglas J, Eastlake L, Farah C, King G, Moore R, Pilcher J, Richards M, Smith S, Walters H (2015) Thoracic Society of Australia and New Zealand oxygen guidelines for acute oxygen use in adults: 'Swimming between the flags'. Respirology 20 (8):1182-1191. doi:10.1111/resp.12620

5. Rhodes A, Evans LE, Alhazzani W, Levy MM, Antonelli M, Ferrer R, Kumar A, Sevransky JE, Sprung CL, Nunnally ME, Rochwerg B, Rubenfeld GD, Angus DC, Annane D, Beale RJ, Bellinghan GJ, Bernard GR, Chiche JD, Coopersmith C, De Backer DP, French CJ, Fujishima S, Gerlach H, Hidalgo JL, Hollenberg SM, Jones AE, Karnad DR, Kleinpell RM, Koh Y, Lisboa TC, Machado FR, Marini JJ, Marshall JC, Mazuski JE, McIntyre LA, McLean AS, Mehta S, Moreno RP, Myburgh J, Navalesi P, Nishida O, Osborn TM, Perner A, Plunkett CM, Ranieri M, Schorr CA, Seckel MA, Seymour CW, Shieh L, Shukri KA, Simpson SQ, Singer M, Thompson BT, Townsend SR, Van der Poll T, Vincent JL, Wiersinga WJ, Zimmerman JL, Dellinger RP (2017) Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock: 2016. Intensive Care Med 43 (3):304-377. doi:10.1007/s00134-017-4683-6

6. de Graaff AE, Dongelmans DA, Binnekade JM, de Jonge E (2011) Clinicians' response to hyperoxia in ventilated patients in a Dutch ICU depends on the level of FiO2. Intensive Care Medicine 37 (1):46-51

7. Eastwood G, Bellomo R, Bailey M, Taori G, Pilcher D, Young P, et al (2012) Arterial oxygen tension and mortality in mechanically ventilated patients. Intensive Care Medicine 38 (1):91-98

8. Itagaki T, Nakano Y, Okuda N, Izawa M, Onodera M, Imanaka H, et al (2015) Hyperoxemia in mechanically ventilated, critically ill subjects: incidence and related factors. Respiratory Care 60 (3):335-340

9. Suzuki S, Eastwood GM, Peck L, Glassford NJ, Bellomo R (2013) Current oxygen management in mechanically ventilated patients: a prospective observational cohort study. Journal of Critical Care 28 (5):647-654

10. de Jonge E, Peelen L, Keijzers PJ, Joore H, de Lange D, van der Voort PH, et al (2008) Association between administered oxygen, arterial partial oxygen pressure and mortality in mechanically ventilated intensive care unit patients. Critical Care 12 (6):R156-R156

11. Panwar R, Capellier G, Schmutz N, Davies A, Cooper DJ, Bailey M, et al (2013) Current oxygenation practice in ventilated patients - an observational cohort study. Anaesthesia and Intensive Care 41 (4):505-514

12. Rachmale S, Li G, Wilson G, Malinchoc M, Gajic O (2012) Practice of excessive F(IO(2)) and effect on pulmonary outcomes in mechanically ventilated patients with acute lung injury. Respiratory Care 57 (11):1887-1893

13. Helmerhorst HJ, Roos-Blom MJ, van Westerloo DJ, de Jonge E (2015) Association between arterial hyperoxia and outcome in subsets of critical illness: a systematic review, meta-analysis, and meta-regression of cohort studies. Critical Care Medicine 43 (7):1508-1519
14. Damiani E, Adrario E, Girardis M, Romano R, Pelaia P, Singer M, et al (2014) Arterial hyperoxia and mortality in critically ill patients: a systematic review and meta-analysis. Critical Care 18 (6):711-711
15. Chu DK, Kim LH, Young PJ, Zamiri N, Almenawer SA, Jaeschke R, et al (2018) Mortality and morbidity in acutely ill adults treated with liberal versus conservative oxygen therapy (IOTA): a systematic review and meta-analysis. Lancet 391 (10131):1693-1705
16. Siemieniuk RAC, Chu DK, Kim LH, Güell-Rous MR, Alhazzani W, Soccal PM, et al (2018) Oxygen therapy for acutely ill medical patients: a clinical practice guideline. BMJ 363:k4169-k4169
17. Mackle D, Bellomo R, Bailey M, Beasley R, Deane A, Eastwood G, Finfer S, Freebairn R, King V, Linke N, Litton E, McArthur C, McGuinness S, Panwar R, Young P (2019) Conservative Oxygen Therapy during Mechanical Ventilation in the ICU. N Engl J Med. doi:10.1056/NEJMoa1903297
18. Barbateskovic M, Schjørring OL, Jakobsen JC, Meyhoff CS, Rasmussen BS, Perner A, et al (2018) Oxygen supplementation for critically ill patients - a protocol for a systematic review. Acta Anaesthesiologica Scandinavica 62 (7):1020-1030
19. Higgins JP, Green S, editor Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.
20. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. Bmj 339:b2700. doi:10.1136/bmj.b2700
21. Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C (2014) Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. BMC Medical Research Methodology 14:120-120
22. Harbord RM, Egger M, Sterne JA (2006) A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. Statistics in Medicine 25 (20):3443-3457
23. Review Manager 5 (RevMan 5) (2014). Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen
24. Wetterslev J, Thorlund K, Brok J, Gluud C (2009) Estimating required information size by quantifying diversity in random-effects model meta-analyses. BMC Medical Research Methodology 9:86-86
25. Wetterslev J, Jakobsen JC, Gluud C (2017) Trial Sequential Analysis in systematic reviews with meta-analysis. BMC Med Res Methodol 17 (1):39. doi:10.1186/s12874-017-0315-7
26. Brok J, Thorlund K, Gluud C, Wetterslev J (2008) Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses. Journal of Clinical Epidemiology 61 (8):763-769
27. Brok J, Thorlund K, Wetterslev J, Gluud C (2009) Apparently conclusive meta-analyses may be inconclusive - trial sequential analysis adjustment of random error risk due to repetitive testing of accumulating data in apparently conclusive neonatal meta-analyses. International Journal of Epidemiology 38 (1):287-298
28. Higgins JP, Whitehead A, Simmonds M (2011) Sequential methods for random-effects meta-analysis. Statistics in Medicine 30 (9):903-921

29. Imberger G, Gluud C, Boylan J, Wetterslev J (2015) Systematic reviews of anesthesiologic interventions reported as statistically significant: problems with power, precision, and type 1 error protection. Anesthesia and Analgesia 121 (6):1611-1622
30. Mascha EJ (2015) Alpha, beta, meta: guidelines for assessing power and type I error in meta-analyses. Anesthesia and Analgesia 121 (6):1430-1433
31. Pogue JM, Yusuf S (1997) Cumulating evidence from randomized trials: utilizing sequential monitoring boundaries for cumulative meta-analysis. Controlled Clinical Trials 18 (6):580-593
32. Terkawi AS, Mavridis D, Flood P, Wetterslev J, Terkawi RS, Bin Abdulhak AA, et al (2016) Does ondansetron modify sympathectomy due to subarachnoid anesthesia?: meta-analysis, meta-regression, and trial sequential analysis. Anesthesiology 124 (4):846-869
33. Thorlund K, Devereaux PJ, Wetterslev J, Guyatt G, Ioannidis JPA, Thabane L, et al (2009) Can trial sequential monitoring boundaries reduce spurious inferences from meta-analyses? International Journal of Epidemiology 38 (1):276-286
34. Thorlund K, Imberger G, Johnston BC, Walsh M, Awad T, Thabane L, Gluud C, Devereaux PJ, Wetterslev J (2012) Evolution of heterogeneity (I2) estimates and their 95% confidence intervals in large meta-analyses. PLoS One 7 (7):e39471. doi:10.1371/journal.pone.0039471
35. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al (2008) GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 336 (7650):924-926
36. Meyhoff CS, Wetterslev J, Jorgensen LN, Henneberg SW, Høgdall C, Lundvall L, et al (2009) Effect of high perioperative oxygen fraction on surgical site infection and pulmonary complications after abdominal surgery: the PROXI randomized clinical trial. JAMA 302 (14):1543-1550
37. Hofmann R, James SK, Jernberg T, Lindahl B, Erlinge D, Witt N, et al (2017) Oxygen therapy in suspected acute myocardial infarction. New England Journal of Medicine 377 (13):1240-1249
38. Roffe C, Nevatte T, Sim J, Bishop J, Ives N, Ferdinand P, et al (2017) Effect of routine low-dose oxygen supplementation on death and disability in adults with acute stroke: the stroke oxygen study randomized clinical trial. JAMA 318 (12):1125-1135
39. Jakkula P, Reinikainen M, Hästbacka J, Loisa P, Tiainen M, Pettilä V, et al (2018) Targeting two different levels of both arterial carbon dioxide and arterial oxygen after cardiac arrest and resuscitation: a randomised pilot trial. Intensive Care Medicine 44 (12):2112-2121
40. Ali K, Warusevitane A, Lally F, Sim J, Sills S, Pountain S, et al (2013) The stroke oxygen pilot study: a randomized controlled trial of the effects of routine oxygen supplementation early after acute stroke - effect on key outcomes at six months. PLOS ONE 8 (6):e59274-e59274
41. Panwar R, Hardie M, Bellomo R, Barrot L, Eastwood GM, Young PJ, et al (2016) Conservative versus liberal oxygenation targets for mechanically ventilated patients. A pilot multicenter randomized controlled trial. American Journal of Respiratory and Critical Care Medicine 193 (1):43-51
42. Perrin K, Wijesinghe M, Healy B, Wadsworth K, Bowditch R, Bibby S, et al (2011) Randomised controlled trial of high concentration versus titrated oxygen therapy in severe exacerbations of asthma. Thorax 66 (11):937-941
43. Wijesinghe M, Perrin K, Healy B, Weatherall M, Beasley R (2012) Randomized controlled trial of high concentration oxygen in suspected community-acquired pneumonia. J R Soc Med 105 (5):208-216. doi:10.1258/jrsm.2012.110084
44. Girardis M, Busani S, Damiani E, Donati A, Rinaldi L, Marudi A, et al (2016) Effect of conservative vs conventional oxygen therapy on mortality among patients in an intensive care unit: the oxygen-ICU randomized clinical trial. JAMA 316 (15):1583-1589

45. Gomersall CD, Joynt GM, Freebairn RC, Lai CK, Oh TE (2002) Oxygen therapy for hypercapnic patients with chronic obstructive pulmonary disease and acute respiratory failure: a randomized, controlled pilot study. Critical Care Medicine 30 (1):113-116
46. Jun J, Sun L, Wang Y, Liu F, Yang G, Han B (2019) Invasive mechanical ventilation with high concentration oxygen therapy for AECOPD patients with acute myocardial infarction. Chest 156 (4):A958
47. NCT02687217. Effect of Peri-operative Supplemental Oxygen in Wound Infection After Appendectomy 2016. https://clinicaltrials.gov/ct2/show/NCT02687217 Accessed 4 December 2019. .
48. Rawles JM, Kenmure AC (1976) Controlled trial of oxygen in uncomplicated myocardial infarction. BMJ 1 (6018):1121-1123
49. Zughaft D, Bhiladvala P, Van Dijkman A, Harnek J, Madsen Hardig B, Bjork J (2013) The analgesic effect of oxygen during percutaneous coronary intervention (the OXYPAIN Trial). Acute Cardiac Care 15 (3):63-68
50. Hrobjartsson A, Emanuelsson F, Skou Thomsen AS, Hilden J, Brorson S (2014) Bias due to lack of patient blinding in clinical trials. A systematic review of trials randomizing patients to blind and nonblind sub-studies. International Journal of Epidemiology 43 (4):1272-1283
51. Savovic J, Turner RM, Mawdsley D, Jones HE, Beynon R, Higgins JPT, et al (2018) Association between risk-of-bias assessments and results of randomized trials in Cochrane Reviews: the ROBES meta-epidemiologic study. American Journal of Epidemiology 187 (5):1113-1122
52. Barbateskovic M, Schjorring OL, Russo Krauss S, Jakobsen JC, Meyhoff CS, Dahl RM, Rasmussen BS, Perner A, Wetterslev J (2019) Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the intensive care unit. Cochrane Database Syst Rev 2019 (11). doi:10.1002/14651858.CD012631.pub2
53. Cabello JB, Burls A, Emparanza JI, Bayliss S, Quinn T (2016) Oxygen therapy for acute myocardial infarction. Cochrane Database of Systematic Reviews (12)
54. Sepehrvand N, James SK, Stub D, Khoshnood A, Ezekowitz JA, Hofmann R (2018) Effects of supplemental oxygen therapy in patients with suspected acute myocardial infarction: a meta-analysis of randomised clinical trials. Heart 104 (20):1694-1698
55. Dellinger RP, Levy MM, Rhodes A, Annane D, Gerlach H, Opal SM, et al (2013) Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock: 2012. Critical Care Medicine 41 (2):580-637
56. Rasmussen BS, Perner A, Wetterslev J, Meyhoff CS, Schjørring OL (2018) Oxygenation targets in acutely ill patients: still a matter of debate. Lancet 392 (10163):2436-2437
57. Asfar P, Schortgen F, Boisramé-Helms J, Charpentier J, Guérot E, Megarbane B, et al (2017) Hyperoxia and hypertonic saline in patients with septic shock (HYPERS2S): a two-by-two factorial, multicentre, randomised, clinical trial. Lancet Respiratory Medicine 5 (3):180-190
58. Austin MA, Wills KE, Blizzard L, Walters EH, Wood-Baker R (2010) Effect of high flow oxygen on mortality in chronic obstructive pulmonary disease patients in prehospital setting: randomised controlled trial. BMJ 341:c5462-c5462
59. Baekgaard JS, Isbye D, Ottosen CI, Larsen MH, Andersen JH, Rasmussen LS, Steinmetz J (2019) Restrictive vs liberal oxygen for trauma patients-the TRAUMOX1 pilot randomised clinical trial. Acta Anaesthesiol Scand 63 (7):947-955. doi:10.1111/aas.13362
60. Bardsley G, Pilcher J, McKinstry S, Shirtcliffe P, Berry J, Fingleton J, Weatherall M, Beasley R (2018) Oxygen versus air-driven nebulisers for exacerbations of chronic obstructive pulmonary disease: a randomised controlled trial. BMC Pulm Med 18 (1):157. doi:10.1186/s12890-018-0720-7

61. Bickel A, Gurevits M, Vamos R, Ivry S, Eitan A (2011) Perioperative hyperoxygenation and wound site infection following surgery for acute appendicitis: a randomized, prospective, controlled trial. Archives of Surgery 146 (4):464-470
62. Bray JE, Hein C, Smith K, Stephenson M, Grantham H, Finn J (2018) Oxygen titration after resuscitation from out-of-hospital cardiac arrest: a multi-centre, randomised controlled pilot study (the EXACT pilot trial). Resuscitation 128:211-215
63. Butler CM, Ham RO, Lafferty K, Cotton LT, Roberts VC (1987) The effect of adjuvant oxygen therapy on transcutaneous pO2 and healing in the below-knee amputee. Prosthet Orthot Int 11 (1):10-16. doi:10.3109/03093648709079373
64. Heidari F, Rahzani K, Iranpoor D, Rezaeed K (2017) The effect of oxygen on the outcomes of non-ST-segment elevation acute coronary syndromes. IJC Metabolic & Endocrine 14:67-71
65. Jernberg T, Lindahl B, Alfredsson J, Berglund E, Bergstrom O, Engstrom A, Erlinge D, Herlitz J, Jumatate R, Kellerth T, Lauermann J, Lindmark K, Lingman M, Ljung L, Nilsson C, Omerovic E, Pernow J, Ravn-Fischer A, Sparv D, Yndigegn T, Ostlund O, James SK, Hofmann R (2018) Long-Term Effects of Oxygen Therapy on Death or Hospitalization for Heart Failure in Patients With Suspected Acute Myocardial Infarction. Circulation 138 (24):2754-2762. doi:10.1161/circulationaha.118.036220
66. Huynh Ky M, Bouchard PA, Morin J, L'Her E, Sarrazin JF, Lellouche F (2017) Closed-loop adjustment of oxygen flowrate with FreeO2 in patients with acute coronary syndrome: comparison of automated titration with FreeO2 (set at two SpO2 target) and of manual titration. A randomized controlled study. Annals of Intensive Care 7 (Suppl 1):O59-O59
67. Ishii K, Morimatsu H, Hyodo T, Ono K, Hidaka H, Koyama Y, et al (2018) Relationship between inspired oxygen concentration and atelectasis formation after extubation. Critical Care Medicine 46 (1 Suppl 1):533-533
68. Khoshnood A, Carlsson M, Akbarzadeh M, Bhiladvala P, Roijer A, Nordlund D, et al (2018) Effect of oxygen therapy on myocardial salvage in ST elevation myocardial infarction: the randomized SOCCER trial. European Journal of Emergency Medicine 25 (2):78-84
69. Khoshnood A, Akbarzadeh M, Roijer A, Meurling C, Carlsson M, Bhiladvala P, et al (2017) Effects of oxygen therapy on wall-motion score index in patients with ST elevation myocardial infarction - the randomized SOCCER trial. Echocardiography 34 (8):1130-1137
70. Kuisma M, Boyd J, Voipio V, Alaspää A, Roine RO, Rosenberg P (2006) Comparison of 30 and the 100% inspired oxygen concentrations during early post-resuscitation period: a randomised controlled pilot study. Resuscitation 69 (2):199-206
71. Lång M, Skrifvars MB, Siironen J, Tanskanen P, Ala-Peijari M, Koivisto T, et al (2018) A pilot study of hyperoxemia on neurological injury, inflammation and oxidative stress. Acta Anaesthesiologica Scandinavica 62 (6):801-810
72. Mazdeh M, Taher A, Torabian S, Seifirad S (2015) Effects of normobaric hyperoxia in severe acute stroke: a randomized controlled clinical trial study. Acta Medica Iranica 53 (11):676-680
73. NCT02378545. Trial of Hyperoxic O2 Therapy vs. Normoxic O2 Therapy in Sepsis (HO2T or NO2T) 2015. https://clinicaltrials.gov/ct2/show/NCT02378545 Accessed 4 Devember 2019.
74. Padma MV, Bhasin A, Bhatia R, Garg A, Singh MB, Tripathi M, et al (2010) Normobaric oxygen therapy in acute ischemic stroke: a pilot study in Indian patients. Annals of Indian Academy of Neurology 13 (4):284-288
75. Ranchord AM, Argyle R, Beynon R, Perrin K, Sharma V, Weatherall M, et al (2012) High-concentration versus titrated oxygen therapy in ST elevation myocardial infarction: a pilot randomized controlled trial. American Heart Journal 163 (2):168-175

76. Rodrigo GJ, Rodriquez Verde M, Peregalli V, Rodrigo C (2003) Effects of short-term 28% and 100% oxygen on PaCO2 and peak expiratory flow rate in acute asthma: a randomized trial. Chest 124 (4):1312-1317
77. Vianna JR, Pires Di Lorenzo VA, Simoes MM, Jamami M (2017) Comparing the Effects of Two Different Levels of Hyperoxygenation on Gas Exchange During Open Endotracheal Suctioning: A Randomized Crossover Study. Respir Care 62 (1):92-101. doi:10.4187/respcare.04665
78. Roffe C, Sills S, Pountain SJ, Allen M (2010) A randomized controlled trial of the effect of fixed-dose routine nocturnal oxygen supplementation on oxygen saturation in patients with acute stroke. Journal of Stroke and Cerebrovascular Diseases 19 (1):29-35
79. Sepehrvand N, Alemayehu W, Rowe BH, McAlister FA, van Diepen S, Stickland M, Ezekowitz JA (2019) High vs. low oxygen therapy in patients with acute heart failure: HiLo-HF pilot trial. ESC Heart Fail 6 (4):667-677. doi:10.1002/ehf2.12448
80. Shi S, Qi Z, Ma Q, Pan R, Timmins GS, Zhao Y, Shi W, Zhang Y, Ji X, Liu KJ (2017) Normobaric Hyperoxia Reduces Blood Occludin Fragments in Rats and Patients With Acute Ischemic Stroke. Stroke 48 (10):2848-2854. doi:10.1161/strokeaha.117.017713
81. Sills S, Halim M, Roffe C (2003) A pilot study of routine nocturnal oxygen supplementation in patients with acute stroke. Age and Ageing 32 (Suppl 2):ii41-ii41
82. Singhal AB, Benner T, Roccatagliata L, Koroshetz WJ, Schaefer PW, Lo EH, et al (2005) A pilot study of normobaric oxygen therapy in acute ischemic stroke. Stroke 36 (4):797-802
83. Singhal A (2013) A phase IIb clinical trial of normobaric oxygen therapy (NBO) in acute ischemic stroke (AIS). Neurology 80 (Suppl 7):S02.001
84. Stewart R (2019) Design and conduct of the New Zealand oxygen therapy in acute coronary syndromes trial. Heart Lung and Circulation 28:S16
85. Stub D, Smith K, Bernard S, Nehme Z, Stephenson M, Bray JE, et al (2015) Air versus oxygen in S-segment-elevation myocardial infarction. Circulation 131 (24):2143-2150
86. Taher A, Pilehvari Z, Poorolajal J, Aghajanloo M (2016) Effects of normobaric hyperoxia in traumatic brain injury: a randomized controlled clinical trial. Trauma Monthly 21 (1):e26772-e26772
87. Thomas M, Voss S, Benger J, Kirby K, Nolan JP (2019) Cluster randomised comparison of the effectiveness of 100% oxygen versus titrated oxygen in patients with a sustained return of spontaneous circulation following out of hospital cardiac arrest: a feasibility study. PROXY: post ROSC OXYgenation study. BMC Emerg Med 19 (1):16. doi:10.1186/s12873-018-0214-1
88. Ukholkina GB, Kostianov IIu, Kuchkina NV, Grendo EP, Gofman IaB (2005) Effect of oxygenotherapy used in combination with reperfusion in patients with acute myocardial infarction. Kardiologiia 45 (5):59-59
89. Wilson AT, Channer KS (1997) Hypoxaemia and supplemental oxygen therapy in the first 24 hours after myocardial infarction: the role of pulse oximetry. J R Coll Physicians Lond 31 (6):657-661
90. Wu WW, Hong HH, Shao XP, Rui L, Jing M, Lu GD (2014) Effect of oxygen-driven nebulization at different oxygen flows in acute exacerbation of chronic obstructive pulmonary disease patients. Am J Med Sci 347 (5):343-346. doi:10.1097/MAJ.0b013e3182937766
91. Young P, Bailey M, Bellomo R, Bernard S, Dicker B, Freebairn R, et al (2014) HyperOxic Therapy OR NormOxic Therapy after out-of-hospital cardiac arrest (HOT OR NOT): a randomised controlled feasibility trial. Resuscitation 85 (12):1686-1691

PAPER VI

CHIMS: Clinical Heterogeneity In Meta-Analysis Score

- a new tool to assess and quantify clinical heterogeneity in meta-analyses of interventions

Barbateskovic M[1,2*], Koster TM[3], Eck RJ[4], Maagaard M[1], Afshari A[5], Blokzijl F[6], Cronhjort M[7], Dieperink W[3], Fabritius ML[8], Feinberg J[9], French C[10], Gareb B[11], Geisler A[12], Granholm A[13], Hiemstra B[14], Hu R[15], Imberger G[16], Jensen BT[17], Jonsson AB[13], Karam O[18], Kong DZ[19,20], Korang SK[1], Koster G[4], Lai B[21], Liang N[22], Lundstrøm LH[23], Marker S[2,13], Meyhoff T[2,13], Nielsen EE[9], Nørskov AK[1], Petersen MW[13], Risom EC[24], Rygård SL[13], Safi S[1], Sethi N[1], Sjövall F[25], Lauridsen SV[26,27], van Bakelen N[11], Volbeda M[3], van der Horst ICC[28], Gluud C[1], Perner A[2,13], Møller MH[2,13], Keus F[13], Wetterslev J[1,2]

Institutional addresses

1. The Copenhagen Trial Unit (CTU), Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark
2. Centre for Research in Intensive Care (CRIC), Copenhagen, Denmark
3. Department of Critical Care, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
4. Department of internal medicine, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
5. Department of Pediatric and Obstetric Anesthesia, University of Copenhagen, Rigshospitalet, Denmark
6. Department of Cardiothoracic Surgery, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
7. Karolinska Institutet, Department of Clinical Science and Education, Södersjukhuset, Section of anaesthesia and intensive care, Stockholm, Sweden
8. Department of Anaesthesiology, Centre of Head and Orthopaedics, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark
9. Department of Internal Medicine, Cardiology Section, Holbaek Hospital, Holbaek, Denmark
10. Centre for Integrated Critical Care, The University of Melbourne, Victoria Australia
11. Department of Oral and Maxillofacial Surgery, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
12. Department of Anaesthesiology, Zealand University Hospital, Koege, Denmark
13. Department of Intensive Care, Copenhagen University Hospital – Rigshospitalet, Copenhagen, Denmark
14. Department of Anesthesiology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands
15. Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing, China
16. Department of Anaesthesia and Pain Medicine, Western Health, Centre for Integrated Critical Care, University of Melbourne, Australia
17. Department of Urology, Aarhus University Hospital, Denmark
18. Division of Pediatric Critical Care Medicine, Children's Hospital of Richmond at VCU, Richmond, Virginia, USA

19. Liaoning University of Traditional Chinese Medicine, Shenyang, China
20. Department of Science and Technology Management, The Affiliated Hospital of Liaoning University of Traditional Chinese Medicine, Shenyang, China
21. The Third Affiliated Hospital of Beijing, University of Chinese Medicine, Beijing 100029, China
22. Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing, 100700, China
23. Department of Anaesthesiology and Intensive Care, Nordsjællands Hospital, Hillerød, Denmark
24. Department of Cardiothoracic Anesthesiology, Copenhagen University Hospital - Rigshospitalet, Copenhagen, Denmark
25. Department for Intensive- and perioperative care, Skane University Hospital, Malmö, Sweden, Department for Clinical Sciences, Mitochondrial Medicine, Lund University, Lund, Sweden
26. WHO-CC, the Parker Institute, Bispebjerg and Frederiksberg Hospital, Copenhagen University Hospital, Copenhagen Denmark
27. Department of Urology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark
28. Department of Intensive Care, Maastricht University Medical Center +, University Maastricht, Maastricht, The Netherlands

Corresponding author (*)
Marija Barbateskovic, Ph.D. student
The Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, DK-2100 Copenhagen, Denmark
Tel: +45 3545 7113
E-mail: marija.barbateskovic@ctu.dk

Keywords: Meta-analysis; Systematic review; Evidence; Quality; Heterogeneity; Tool

Competing interests

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: MB and JW report grants from Innovation Fund Denmark during the conduct of the study. TSM reports grants from the Novo Nordisk Foundation and grants from Sofus Friis Foundation outside the submitted work. AP reports two grants from the Novo Nordisk Foundation related to intensive care research. The authors report no other relationships or activities that could appear to have influenced the submitted work.

ABSTRACT

Objective: To develop and validate Clinical Heterogeneity In Meta-analysis Score (CHIMS), a new tool for assessing and quantifying clinical heterogeneity in meta-analyses of interventions.

Study design and setting: The development of CHIMS was based on consensus work informed by empirical literature and expertise. We drafted the CHIMS tool, refined it and validated CHIMS for interrater scale reliability and agreement in three groups.

Results: CHIMS measures clinical heterogeneity on a scale that includes four domains with 11 items overall: setting (time of conduct/country development status/units type), population (age; sex; patient inclusion criteria/baseline disease severity, co-morbidities), intervention (intervention intensity/strength/duration of intervention; timing; control intervention; co-interventions), outcome (definition of outcome; timing of outcome assessment). The CHIMS tool is completed in two steps: first two authors independently assess clinical heterogeneity in the four domains. Second, after agreeing upon scores of individual items a consensus score is achieved.

Interrater scale reliability and agreement ranged from moderate to almost perfect depending on the type of raters.

Conclusion: CHIMS is the first tool developed for assessing and quantifying clinical heterogeneity in meta-analyses of interventions. We found CHIMS to be a reliable tool for assessing and quantifying clinical heterogeneity in meta-analyses.

1.0 INTRODUCTION

A meta-analysis of high-quality randomised clinical trials is considered the best available evidence in health care management and often forms the basis of clinical practice guidelines and for protocols of randomised clinical trials [1]. Still, undetected clinical, methodological and/or statistical heterogeneity may lead to inappropriate conclusions or recommendations.

Several potential sources of heterogeneity exist among trials included in systematic reviews and meta-analyses. Clinical heterogeneity can be characterised by variability in settings, participants, characteristics of interventions and comparators, use of co-interventions, and the types and timing of outcome assessments. Methodological heterogeneity, or difference in risk of bias, is characterised by variability in trial design and quality in distinct domains. Statistical heterogeneity is characterised by variability in treatment effects between trials [2]. The presence and magnitude of statistical heterogeneity is associated with risk of bias and may be associated with clinical sources of heterogeneity [3, 4], arise from other unknown or unrecorded trial characteristics, or from random errors ('play of chance') due to sparse data and repetitive testing [3-7]. In the context of systematic reviews, clinical heterogeneity can be defined as differences in the clinical characteristics of trials, which may lead to variations in the pooled treatment effect estimates across trials that are not covered by the bias assessment of the included trials [3, 4, 8].

In contrast to methodological and statistical heterogeneity [9], clinical heterogeneity in systematic reviews has only been given sporadic attention [5, 7]. Although subgroup analyses and meta-regression analyses may detect differences in treatment effect size associated with trial characteristics, the overall clinical heterogeneity is usually not assessed. We are not aware of any tool designed to assess and quantify clinical heterogeneity in meta-analyses. To improve the interpretation of systematic reviews and possibly their external validity it is crucial to increase our understanding of clinical heterogeneity. It is also essential to investigate whether clinical heterogeneity is associated with statistical heterogeneity. If this is so, a tool to detect and quantify clinical heterogeneity could be redundant if it was covered by the assessment of the degree of statistical heterogeneity. Furthermore, as methodological heterogeneity, according to bias of the included trials, does not include clinical differences between trials of the included interventions, such as dosage or length of follow-up, it may be essential to quantify overall clinical heterogeneity.

Accordingly, we aimed to develop a tool for assessing and quantifying clinical heterogeneity in meta-analyses of interventions, and to test the reliability of the tool. In a supplementary exploratory analysis, we estimated the association, if any, between clinical and statistical heterogeneity.

2.0 METHODS

The development and interrater scale reliability and agreement assessments of the Clinical Heterogeneity In Meta-analyses Score (CHIMS) tool was conducted following our pre-published protocol and reported following the Guidelines for Reporting Reliability and Agreement Studies (GRRAS) [10, 11].

2.1 Development of CHIMS

We constructed CHIMS during a pilot phase based on consensus work informed by empirical literature and expertise by Gagnier and colleagues [12, 13] (Figure 1a): a methodologic review of guidance of the literature on clinical heterogeneity in systematic reviews and their consensus-based recommendations for investigating clinical heterogeneity in systematic reviews (based on the method using a modified Delphi technique with three phases: 1. pre-meeting item generation; 2. face-to-face consensus meeting; and 3. post-meeting feedback).

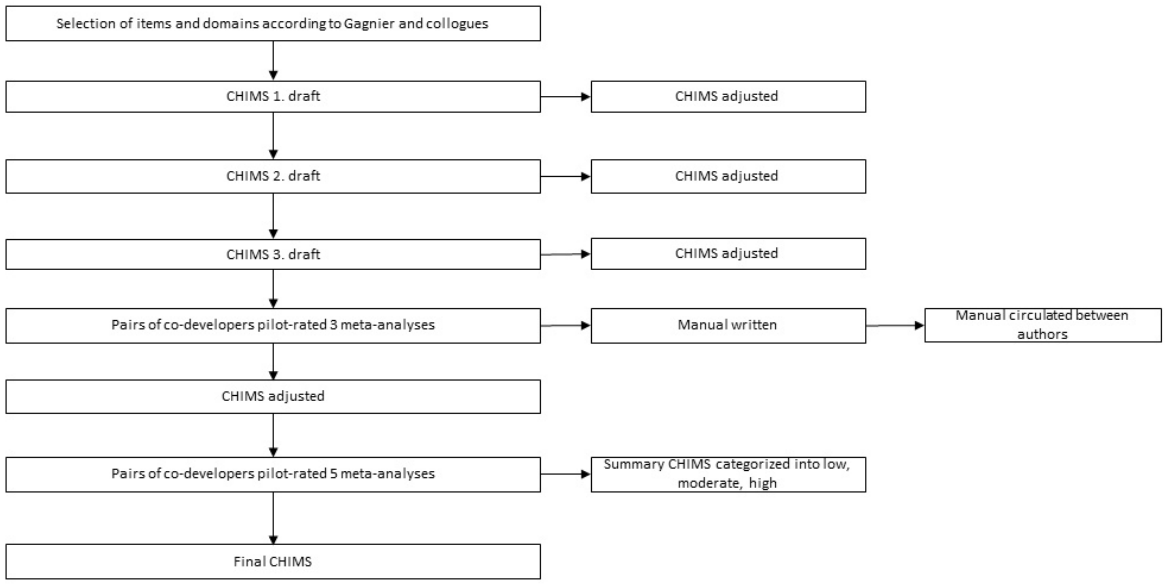


Figure 1a. Process of the development of CHIMS

One author drafted the CHIMS tool which was reviewed by the author/project group and revised according to comments and circulated three times. Initially, a complete list of Cochrane reviews within the field of intensive care medicine was created [14, 15]. Two authors scored the first three meta-analyses with subsequent adjustment of the CHIMS tool and wrote a draft manual providing guidance on the use of CHIMS. The manual was circulated between the authors and revised. Hereafter, two authors scored the next five meta-analyses from the same list and the summary CHIMS score was categorised into low, moderate, or high clinical heterogeneity. A final version of the CHIMS tool was produced to be evaluated for reliability.

3.0 THE CHIMS TOOL

CHIMS was developed to detect and quantify clinical heterogeneity in meta-analyses of interventions. It is intended to help researchers who conduct meta-analyses within all medical fields quantifying clinical heterogeneity and for researchers and guideline panels critically appraising meta-analyses.

CHIMS measures clinical heterogeneity on an ordinal scale that includes four domains with 11 items overall, covering essential domains describing clinical heterogeneity [12, 13] (Table 1).

Table 1. The clinical heterogeneity in a meta-analysis score (CHIMS) tool

Domains	**Items**	**Score**	**Explanation of score for extreme differences between trials in a meta-analysis**
Setting heterogeneity	**1. Years reported (A), performed in developed vs developing country (B), unit type (C)**	0	No differences: A) years reported differ < 15, B) No developed vs developing countries, AND C) slight variations in the unit or facility type and there is low risk of affecting other fields of heterogeneity
		1	Slight variation (at least one of A-C involved): A) years reported differ ≥ 15, OR B) developed vs developing countries, OR treating units not similar, OR C) if there are slight variations in the unit or facility type but there is risk of affecting other fields of heterogeneity
		2	Considerable variation (all of A-C involved): A) years reported differ ≥ 15, AND B) developed vs developing countries, AND C) treating units not similar (all of A-C involved), OR if the trials in the opinion of the assessor differs markedly in setting heterogeneity
Population heterogeneity	**2. Age**	0	Mean/median age ≤ 10 years difference
		1	11 to 20 years difference in mean/median age
		2	Mean/median age > 20 years difference
	3. Sex	0	% women ≤ 20 % absolute difference between trials
		1	21% to 30% absolute difference of % women between trials
		2	More than > 30% absolute difference between trials

	4. Participant inclusion criteria and baseline disease severity	0 1 2	Different trials include patients that are equally ill or the difference in risk or score for disease severity of patients ≤ 20% Condition/patient population differs slightly with 50% or more overlap of types of participants and/or the difference in risk or score for disease severity of patients is 21% to 30% Condition/patient population differs considerable and/or the difference in risk or score for disease severity of patients > 30% Use relative difference when inclusion criteria are assessed (disease severity scores).
	5. Co-morbidities	0 1 2	Difference in frequency of important comorbidities ≤ 20% or no co-morbidities are reported in the included trials and differences in co-morbidities are assumed absent Slight differences in important co-morbidities, between 21% and 30%, or no co-morbidities are reported in the trials, but differences in co-morbidities are assumed Differences in frequency of important comorbidities > 30% or highly likely variations in co-morbidities Use absolute difference when comparing important comorbidities.
Intervention heterogeneity	**6. Intensity, strengths, or duration of intervention**	0 1 2	Little variation: differences in dose, strengths, devices, cut-offs, or duration of interventions ≤20% Slight variation: 21% to 30% differences in dose, strengths, devices, cut-offs, or duration intervention, or if dose, strength, cut-offs or duration of intervention cannot be assessed from the information in the included trials Considerable variation: if different types of interventions are used, or different doses, strengths, devices, cut-offs, or duration of intervention > 30% Use relative differences when assessing intensity, strengths, duration.

	7. Timing	0 1 2	Criteria for starting the intervention are similar, or relative differences of timing of intervention differs ≤ 20% Criteria for starting the intervention differ slightly, or the relative timing difference is 21% to 30% Criteria for starting the intervention differ, or relative timing difference exceeds > 30%
	8. Control intervention	0 1 2	All control interventions are the same Control interventions include placebo AND no intervention, assess as item 6 if an active intervention is used Including trials with different active control interventions OR trials with active and placebo/no intervention
	9. Co-interventions	0 1 2	No apparent differences in co-interventions, OR standard care is not described or assumed to be the same, OR equally applied in groups, OR different co-interventions are used but the effects of the co-interventions are assumed to be small Slight variation in co-interventions or the same cointerventions are used with slight variation (< 30% difference in e.g. doses or numbers of participants using the co-intervention) Considerable differences if it is assumed that the cointervention is not usual care, or differences in use or e.g. doses of cointerventions > 30% Use absolute difference when assessing co-interventions.
Outcome heterogeneity	**10. Definition of the outcome in the meta-analysis**	0 1 2	Same definition of outcome Slight variations in definition of outcome Considerable variations in definition of outcome
	11. Timing of outcome measurement	0 1 2	Less than one month between follow-up of outcome More than one but less than or equal to 3 months between follow-ups More than 3 months between follow-up of outcome

- The first domain aims to detect setting heterogeneity by assessing differences between trials in: time of conduct; type of country development status; localisation within the health care system.
- The second domain aims to detect population heterogeneity by assessing differences between trials in: age; sex; patient inclusion criteria and baseline disease severity; comorbidities.
- The third domain aims to detect intervention heterogeneity by assessing differences between trials in: intervention intensity/strength/duration (dose, frequency, duration, device, cut-off values); timing of intervention(s); heterogeneity of control-interventions; use of co-interventions.
- The fourth domain aims to detect outcome heterogeneity between trials by assessing differences between trial in: outcomes definitions and timing of outcome assessment(s).

The tool is used in two steps. 1. Assess clinical heterogeneity in each of the four domains (11 items) (Table 1). The 11 items are each scored as to low clinical heterogeneity (0 points), moderate (or unknown/undescribed) clinical heterogeneity (1 point), or high clinical heterogeneity (2 points), with a total range of 0 to 22 with equal weight assigned to each item. Guidance on how to score each item is provided in the CHIMS manual (Supplementary appendix A). 2. Sum the item scores into an overall CHIMS score.

3.1 Assessment of scale reliability and agreement

A sample of 60 meta-analyses was deemed sufficient to evaluate CHIMS as 10-20 evaluations per category is considered sufficient to accurately estimate the coefficients of a regression model [16] and two times the squared amount of categories ($2 \cdot categories^2$) to approximate a normal distribution to be used for the analysis of quadratic weighted kappa [17].

We applied CHIMS to the 60 meta-analyses with a dichotomous primary outcome with at least three randomised clinical trials included (Figure 1b). We selected in a consecutive order 20 titles (which had not already been used in the development of the CHIMS) from the list of Cochrane reviews within the intensive care setting. Another 20 Cochrane reviews of interventions focusing on clinical scenarios outside the intensive care setting were selected to cover a wide range of non-intensive care interventions. These were picked by browsing The Cochrane Database of Systematic Reviews by topic. Finally, a convenience sample of 20 mainly non-Cochrane reviews with meta-analyses, of which around half were within the field of intensive care, were selected.

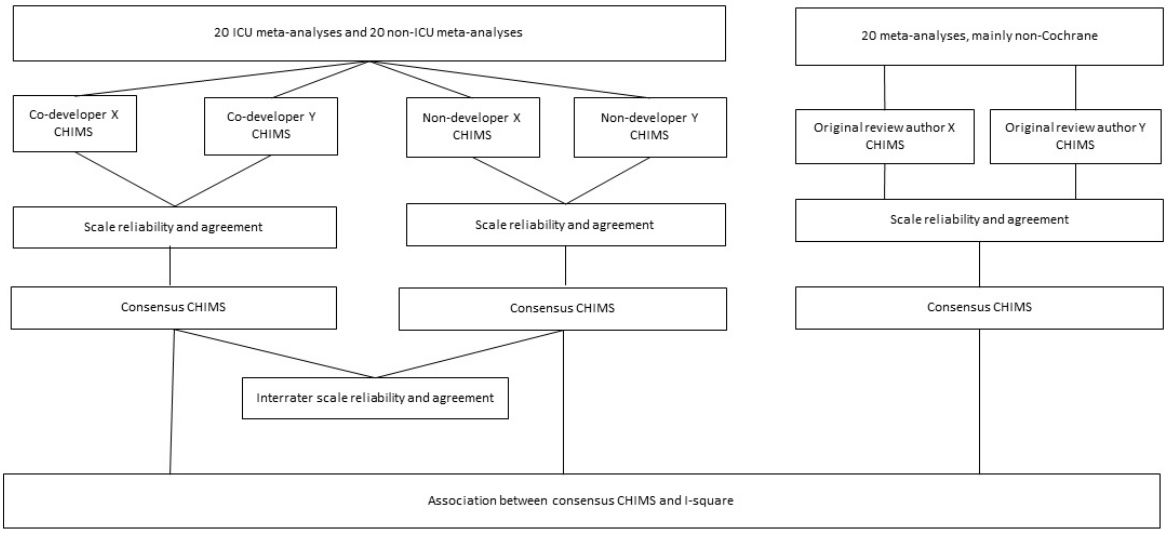


Figure 1b. Interrater scale reliability and agreement testing of CHIMS

We evaluated CHIMS for interrater scale reliability by CHIMS scoring of the 60 meta-analyses [11]. Two independent evaluators involved in the development of CHIMS (co-developers) and two independent evaluators not involved in the development of CHIMS and neither in the meta-analyses (non-developers) scored the same 40 meta-analyses published in the Cochrane Database of Systematic Reviews. Finally, the sample of 20 mainly non-Cochrane reviews with meta-analyses were CHIMS scored by two of the review's original authors.

The two non-developers of CHIMS and the 20 pairs of original review authors were instructed only by reading the guidance document – no additional guidance was given.

After individual and independent scoring of CHIMS, the evaluators pairwise agreed upon each item score, thereby achieving a total consensus CHIMS.

According to our protocol [10], interrater scale reliability of the summarised total CHIMS was investigated in four scenarios: between co-developers of CHIMS; non-developers of CHIMS; pairs of original review authors; and between consensus scores of co-developers and non-developers of CHIMS. The consensus CHIMS was achieved by reaching consensus for each of the 11 items of the CHIMS and recalculating a summarised, total consensus CHIMS.

We stratified the analyses of interrater scale reliability between co-developers of CHIMS and non-developers of CHIMS according to meta-analyses of intensive care unit (ICU) interventions or non-ICU interventions. We analysed the possible difference between the distributions of consensus CHIMS in ICU and non-ICU meta-analyses using the Mann-Whitney test, presenting box and whiskers plots with medians, interquartile ranges and full ranges.

The interrater reliabilities of the summarised total CHIMS were analysed with intraclass correlation coefficients (ICC) using one-way random reliability analysis of exact agreement on average CHIMS and for single measures (single meta-analysis) for co-developers and non-developers of CHIMS. A two-way random reliability analysis of exact agreement was used for pairs of original review authors. For pairs of original review authors, we also analysed interrater reliability within the domains of CHIMS.

Quadratic weighted kappa values for the agreement between the protocolised categorical classification of CHIMS (low: 0-11; moderate 12-18; high 19-22), defined after a pilot scoring, were calculated. Moreover, quadratic weighted kappa values for the agreement between the protocolised categorical classification of CHIMS and the categorical classification of I^2 in the meta-analyses (low $I^2 \leq 30\%$; moderate $I^2 > 30\%$ to $\leq 60\%$; high $I^2 > 60\%$) modified from Higgins et al. were calculated [18]. Imputed relative distances between ordinal categories in the calculation of the quadratic weighted kappa were set to one.

Additionally, linear regression analyses were performed for any associations between the raters' summarised total CHIMS. Finally, we analysed the possible association between the consensus CHIMS and I^2 in 60 meta-analyses using linear regression. Pearson's correlation coefficients, R^2, and P-values for the linear regression coefficients being equal to zero were calculated. We plotted regression lines and regression standardised residuals including P-P plots to investigate whether residuals were normally distributed as required for a linear regression models to be adequate.

Agreement was classified as suggested by Landis and Koch: values less than 0 indicated poor, 0-0.20 slight, 0.21-0.40 fair, 0.41-0.60 moderate, 0.61-0.80 substantial, and 0.81-1.00 almost perfect agreement [19]. All ICC and kappa values are presented with 95% confidence interval (CI). SPSS version 17 (SPSS Statistics for Windows, Chicago: SPSS Inc.) was used for the analysis of scale reliability. http://vassarstats.net/kappa.html was used to calculate kappa values.

3.2 Interrater scale reliability and agreement of CHIMS

Four raters independently applied CHIMS to 20 meta-analyses of ICU-interventions and 20 meta-analyses of non-ICU interventions, for a total of 160 evaluations. Twenty pairs (of 35 different raters) of original review authors applied CHIMS to 20 meta-analyses, for a total of 40 evaluations (Supplementary appendix B). In total, 721 trials were included in the 60 meta-analyses assessed with a median of 8 (interquartile range 5-15) trials per meta-analysis.

Main characteristics of the meta-analyses evaluated, their reference, and supplemental figures are presented in the Supplementary appendix C.

CHIMS varied between 0 and 21 points in the 60 meta-analyses. Average CHIMS for all raters varied between (mean ± SD) 11.5 **±** 5.4 and 14.2 ± 3.9 and the difference between average CHIMS for pairs of raters ranged between 0.3 and 2.4 (Table 2).

Table 2. Interrater agreements stratified for types of raters as developers of CHIMS, original review authors, and non-developers**.**

Coefficients → / **Datasets analysed** ↓	**Scale reliability: intraclass correlation coefficient on average measures (95% CI)**	**Intraclass correlation coefficient on single measures (95% CI)**	**Pearson's correlation coefficient (95% CI)**	**R^2, P-value for test of linear regression coefficient equal to 0, and model fit**	**Constant (95% CI) in linear regression equation, raters mean ± SD**	**Quadratic weighted kappa (95% CI) for agreement between low, moderate, or high CHIMS[#]**
Two co-developers of CHIMS*	0.85 (0.72 to 0.92)	0.74 (0.56 to 0.85)	0.76 (0.53 to 0.98)	0.54 P<0.0001 Residual plots suggest goodness of model fit	3.0 (-0.21 to 6.2) 1.Rater: 13.6 **±** 3.6 2.Rater: 13.3 **±** 3.7	0.61 (0.18 to 1.00)
Pairs of original review authors**	0.94 (0.85 to 0.98)	0.89 (0.75 to 0.96)	0.90 (0.69 to 1.12)	0.82 P<0.0001 Residual plots suggest goodness of fit	0.13 (-2.9 to 3.2) 1.Rater: 13.2 ± 6.2 2.Rater: 12.1 ± 6.2	0.72 (0.42 to 1.00)
Two non-developers of CHIMS*	0.74 (0.51 to 0.86)	0.59 (0.34 to 0.76)	0.72 (0.56 to 0.88)	0.52 P<0.0001 Residual plots suggest goodness of model fit	7.9 (5.8 to 10.0) 1.Rater: 13.9 **±** 3.9 2.Rater: 11.5 **±** 5.4	0.41 (0.14 to 0.69)
Consensus scores from co-developers and non-developers of CHIMS*	0.91 (0.83 to 0.95)	0.84 (0.72 to 0.91)	0.85 (0.81 to 1.22)	0.73 P<0.0001 Residual plots suggest goodness of model fit	0.75 (-3.7 to 2.2) 1.Rater: 14.2 **±** 3.9 2.Rater: 13.7 **± 4**.5	0.68 (0.38 to 0.98)

3.2.1 Co-developers of CHIMS

Interrater scale reliability for two co-developers of CHIMS was almost perfect with an ICC of 0.85 (95% confidence interval 0.72 to 0.92) for average measures and substantial with an ICC of 0.74 (0.56 to 0.85) for single measures. Pearson's correlation coefficient was 0.76 (0.53 to 0.98). Quadratic weighted kappa values for the agreement between categorical CHIMS for two co-developers was substantial with a kappa of 0.61 (0.18 to 1.00). Consensus CHIMS score between developers of CHIMS stratified for ICU and non-ICU meta-analyses were median 18 (range 9-20) and median 12 (range 7-18), respectively (P=0.001, Mann-Whitney test for different distributions of CHIMS; Supplementary appendix C). The interrater scale reliability between two developers of CHIMS in ICU meta-analyses and non-ICU meta-analyses were almost perfect as well (Table 3).

Table 3. Interrater agreements stratified for ICU and non-ICU meta-analyses

Coefficients → **Datasets analysed ↓**	**Scale reliability: intraclass correlation coefficients on average measures (95% CI)**	**Intraclass correlation coefficients on single measures (95% CI)**	**Pearson's correlation coefficient (95% CI)**	**Raters' means ± SD**
ICU meta-analyses Interrater agreement* between two co-developers of CHIMS*	0.71 (0.29 to 0.89)	0.55 (0.17 to 0.80)	0.54 (0.12 to o.89)	1.Rater: 15.3 ± 3.2 2.Rater: 15.7 ± 3.0
Non-ICU meta-analyses Interrater agreement* between two co-developers of CHIMS	0.82 (0.56 to 0.93)	0.70 (0.39 to 0.87)	0.75 (0.35 to 0.91)	1.Rater: 12.0 ± 3.2 2.Rater: 11.0 ± 2.7
ICU meta-analyses Interrater agreement* between two non-developers of CHIMS	0.78 (0.45 to 0.91)	0.64 (0.29 to 0.84)	0.69 (0.27 to 0.86)	1.Rater: 15.6 ± 3.9 2.Rater: 14.1 ± 4.7
Non-ICU meta-analyses Interrater agreement* between two non-developers of CHIMS	0.55 (-0.13 to 0.82)	0.38 (-0.06 to 0.69)	0.63 (0.17 to 0.70)	1.Rater: 12.4 ± 3.4 2.Rater: 9.0 ± 4.9

* One-way random reliability analysis of exact agreement in 20 meta-analyses rated with CHIMS. CI is confidence interval. SPSS ver. 17 was used.

3.2.2 Non-developers of CHIMS

Interrater scale reliability for two non-developers of CHIMS was substantial with an ICC of 0.74 (0.51 to 0.86) for average measures and moderate for single measures with an ICC of 0.59 (0.34 to 0.76). Pearson's correlation coefficient was 0.72 (0.56 to 0.88). Quadratic weighted kappa values for the agreement between categorical CHIMS for two non-developers was moderate with a kappa of 0.41 (0.14 to 0.69). Consensus CHIMS score between non-developers of CHIMS stratified for ICU and non-ICU meta-analyses were median 17 (range 7-21) and median 12 (range 5-19), respectively (P=0.016, Mann-Whitney test for different distributions of CHIMS; Supplementary appendix C). The interrater scale reliability between two non-developers of CHIMS on average measures in ICU meta-analyses and non-ICU meta-analyses were substantial and moderate, respectively (Table 3), and moderate and fair, respectively for single measures (Table 3).

3.2.3 Pairs of original review authors

Interrater scale reliability of CHIMS for two original review authors was almost perfect with an ICC of 0.94 (0.85 to 0.98) for average measures and 0.89 (0.75 to 0.96) for single measures. Pearson's correlation coefficient was 0.90 (0.69 to 1.12) (Figure 3). Quadratic weighted kappa values for the agreement between two original review authors was substantial with a kappa of 0.72 (0.42 to 1.00).

Interrater scale reliability of CHIMS for two original review authors on the four CHIMS domains were consistent with the summary scale reliability ranging across domains from 0.68 to 0.93 on average measures and from 0.51 to 0.87 for single meta-analyses (Supplementary appendix 3).

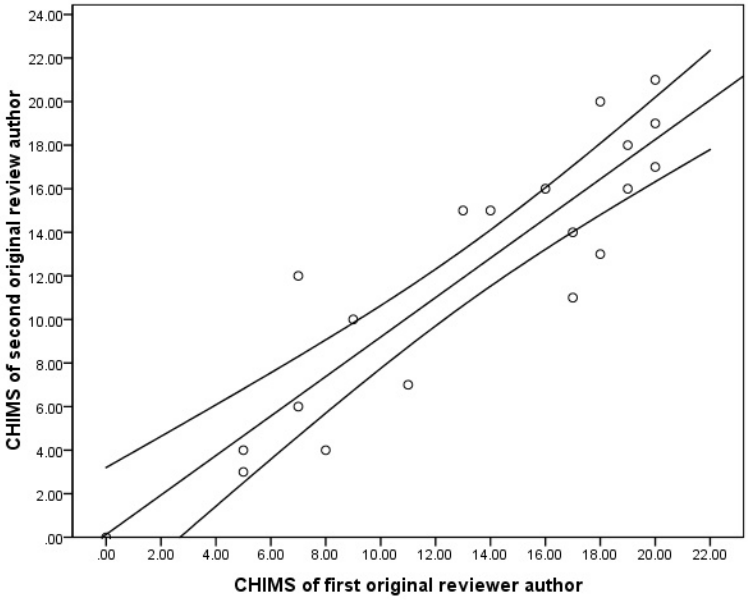


Figure 3. Fitted regression line (Y = 0.90 · X + 0.13) of CHIMS from second original review author on CHIMS from first original review author in 20 meta-analyses from mainly non-Cochrane reviews. Hyperbolic lines around fitted line represents 95% CI for the regression line. R^2=0.82.

3.2.4 Consensus scores between developers and non-developers of CHIMS

Interrater scale reliability of consensus CHIMS between developers and non-developers of CHIMS was almost perfect with an ICC of 0.91 (0.83 to 0.95) for average measures 0.84 (0.72 to 0.91) for single measures. Pearson's correlation coefficient was 0.85 (0.81 to 1.22) (Supplemental appendix C). Quadratic weighted kappa values for the agreement between the categorical consensus CHIMS was substantial with a kappa of 0.68 (0.38 to 0.98).

Linear regression showed that a linear model explained from 52% to 82% of the co-variation in CHIMS between raters regardless of the meta-analyses being ICU or non-ICU meta-analyses (Table 2). Model of fit analyses justified a linear regression model as standardised residuals were normally distributed.

3.2.5 Association between clinical and statistical heterogeneity (consensus CHIMS and I^2)

Consensus CHIMS from both developers and non-developers of CHIMS supplemented with consensus CHIMS for pairs of original review authors indicated an absence of association with regression coefficients close to zero with narrow CIs: -0.02 (-1.6 to 1.4) and -0.13 (-2.0 to 0.7), respectively (Table 4 and Supplementary appendix C). In fact, a linear model seems unjustified, as analyses of standardised residuals indicated absence of a normal distribution. Quadratic weighted kappa values for the agreement between categorical consensus CHIMS and categorical statistical heterogeneity was not calculable because the observed concordance was smaller than mean chance concordance (Table 4).

Table 4. Regression of consensus CHIMS on statistical heterogeneity (I^2) and Kappa between categorised CHIMS and categorised I^2

Coefficients → / **Dataset analysed** ↓	**Regression coefficient (95% CI)**	**R^2, P-value for test of regression coefficient equal to 0, and model fit**	**Constant (95% CI) in linear regression equation**	**Quadratic weighted kappa (95% CI) for agreement between low, moderate, or high CHIMS# and low, moderate and high statistical heterogeneity**
Consensus CHIMS (from two co-developers)* versus I^2**	-0.02 (-1.6 to 1.4)	0.000 P=0.88 Residual plots suggest lack of model fit	21.5 (-0.07 to 43.1)	Kappa is not calculated for this data set because observed concordance is smaller than mean-chance concordance
Consensus CHIMS (from two non-developers)* versus I^2**	-0.13 (-2.0 to 0.7)	0.016 P=0.34 Residual plots suggest lack of model fit	28.7 (9.2 to 48.2)	Kappa is not calculated for this data set because observed concordance is smaller than mean-chance concordance

Regression analysis of consensus CHIMS and I^2 in 60 meta-analyses rated with CHIMS.
Low CHIMS 0 to 10; moderate CHIMS 11 to 18; high CHIMS 19 to 22.
* Supplemented with consensus scores from original review authors.
** Low $I^2 \leq 30\%$; moderate $I^2 > 30\%$ to $\leq 60\%$; high $I^2 > 60\%$. NC is not calculated. SPSS ver. 17 was used.

4.0 DISCUSSION

We aimed at developing a tool to assess and quantify clinical heterogeneity in meta-analyses of interventions, and to assess its reliability and agreement across different raters and topics. We constructed the CHIMS tool and evaluated it in three groups. The highest interrater scale reliability and agreement on both average and single summarised measures of CHIMS and categorical classification of CHIMS (low, moderate, high) were achieved in groups of original review authors. Co-developers achieved lower interrater scale reliability and agreement compared to original review authors. Non-developers of CHIMS who were not involved in the rated meta-analyses achieved the lowest interrater reliability and agreement. Although interrater scale reliability and agreement between non-developers of CHIMS were only moderate to substantial for average measures, single measures, and categorical classifications of CHIMS, respectively, the reliability and agreement increased to substantial and almost perfect, respectively, when either scores from two co-developers of CHIMS or two original review authors were compared. The external reliability (or generalisability) tested by assessing consensus scores from the group of co-developers and non-developers of CHIMS was almost perfect when analysing interrater scale reliability and substantial when analysing CHIMS categories stressing the fact that consensus is important to achieve when assessing clinical heterogeneity with CHIMS. Consensus scores of co-developers and non-developers showed significant higher CHIMS scores within intensive care meta-analyses compared to non-ICU meta-analyses.

Moreover, we observed absence of a linear association between clinical heterogeneity measured with CHIMS and statistical heterogeneity quantified by I^2 as regression coefficients were close to zero with narrow confidence intervals. To summarise our exploratory analyses, it appears that clinical and statistical heterogeneity are two different aspects of heterogeneity in meta-analyses.

4.1 Strengths and limitations

Our approach used in the development of CHIMS has several strengths. We relied strongly on the consensus reports and expert panel from which the items and domains covered in CHIMS originate [12, 13]. The CHIMS tool was developed over several steps and a final version of the CHIMS tool was extensively evaluated in a relatively large sample of meta-analyses of different settings, populations, interventions and outcomes by three groups of raters. It includes a domain and item-based approach supported by signalling questions in a manual similar to other tools used in the systematic review process [2, 20, 21] and it appears that the raters found the CHIMS tool operational determined by only two clarifying questions among non-developers and original review authors.

Knowledge about the medical field and interventions assessed in the meta-analyses seems preferable when assessing clinical heterogeneity with CHIMS; other expertise such as knowledge of trial methodology or statistics is not required. Application of the tool requires some time investment as full trial reports from all trials included in a meta-analysis have to be explored carefully, especially when many trials with low clinical heterogeneity in one or more domains are included in the assessed meta-analysis. Conversely, the scoring of CHIMS can be completed rather quickly in the presence of high clinical heterogeneity for an item when just two trials differ substantially (see manual, supplementary appendix A). Nonetheless, we recommend looking for the specific information needed to assess all items in all trials to get a full overview of the clinical heterogeneity in the meta-analysis.

In some circumstances some items may overlap. This is the case when a meta-analysis is conducted in a 'lumping' review that includes all participants regardless of e.g. age, and thus may lead to high clinical heterogeneity

between the included trials for the items of age, but also for items such as participant inclusion criteria, baseline disease severity, and co-morbidities, consequently leading to double counts.

In our sample, clinical heterogeneity in meta-analyses of interventions in the field of intensive care appear to be high as compared to the group of meta-analyses in other medical fields. This difference indicates higher clinical heterogeneity in meta-analyses in the field of intensive care, but it may also be a chance finding. Nevertheless, the domains and items included in the CHIMS tool have been selected to be key categories/topics especially with the purpose of investigating clinical heterogeneity in meta-analyses regardless of the medical field [13].

A reason for the imperfect agreement between the categories low, moderate and high CHIMS may be attributable to the somewhat arbitrary cut off between these categories, which may be reflected in the analyses of the quadratic weighted kappa values.

4.2 Implications

The CHIMS tool is designed to be applicable in all medical fields and intended to be used by multiple users conducting or assessing meta-analyses. Our analyses illustrate that CHIMS is a reliable tool for assessing and quantifying clinical heterogeneity in meta-analyses. We consider to use CHIMS in the systematic review process to quantify overall clinical heterogeneity, to highlight clinical heterogeneity within specific domains and it may be practical when assessing indirectness and inconsistency in GRADE [5]. Other implications include the possibility of comparing CHIMS across meta-analyses and with statistical heterogeneity such as I^2 or D^2 [22]. However, our finding of lack of association between clinical and statistical heterogeneity should be considered hypothesis-generating due to the limited number of investigated meta-analyses and scenarios. In any case, we recommend these results to be explored further. We encourage investigators to provide feedback and report experiences to the corresponding author.

In conclusion, CHIMS is the first tool developed to assess and quantify clinical heterogeneity in meta-analyses. Interrater scale reliability for overall CHIMS in various scenarios varied from moderate to almost perfect. Reliability was almost perfect between original review authors and between consensus scores of non-developers and co-developers of the CHIMS. We consider CHIMS a reliable tool and recommend using CHIMS for the assessment of the overall clinical heterogeneity in meta-analyses.

Author contributions

MB and JW conceived the project. MB organised, collected data, and oversaw the project. JW drafted the first version of CHIMS, and MB, TMK, FK, CG, MHM, ICCH and AP contributed to the development. MB, TMK, and JW drafted the guidance manual. MB and TMK pilot-tested CHIMS and comprised the pair of co-developers; RE and MM comprised the pair of non-developers. 41/45 co-authors each CHIMS scored at least one meta-analysis. JW performed the statistical analyses. MB and JW wrote the first draft of the manuscript. All authors reviewed, commented on the draft, and finally approved the manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. MB's and JW's contribution to this work, was supported by Innovation Fund Denmark [grant number 4108-00011B].

Supplementary appendix A: CHIMS tool and guidance document

Supplementary appendix B: Data from assessment of reliability and agreement of CHIMS

Supplementary appendix C: Supplementary material for reliability and agreement analyses

REFERENCES

1. Garattini S, Jakobsen JC, Wetterslev J, Bertele V, Banzi R, Rath A, Neugebauer EA, Laville M, Masson Y, Hivert V, Eikermann M, Aydin B, Ngwabyt S, Martinho C, Gerardi C, Szmigielski CA, Demotes-Mainard J, Gluud C. Evidence-based clinical practice: Overview of threats to the validity of evidence and how to minimise them. Eur J Intern Med 2016;32:13-21.
2. Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.handbook.cochrane.org.
3. Savovic J, Turner RM, Mawdsley D, Jones HE, Beynon R, Higgins JPT, Sterne JAC. Association between risk-of-bias assessments and results of randomized trials in Cochrane reviews: The ROBES meta-epidemiologic study. Am J Epidemiol 2018;187(5):1113-1122.
4. Rhodes KM, Turner RM, Savovic J, Jones HE, Mawdsley D, Higgins JPT. Between-trial heterogeneity in meta-analyses may be partially explained by reported design characteristics. J Clin Epidemiol 2018;95:45-54.
5. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schunemann HJ. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336(7650):924-6.
6. Keus F, Wetterslev J, Gluud C, van Laarhoven CJ. Evidence at a glance: error matrix approach for overviewing available evidence. BMC Med Res Methodol 2010;10:90.
7. Reade MC, Delaney A, Bailey MJ, Angus DC. Bench-to-bedside review: avoiding pitfalls in critical care meta-analysis--funnel plots, risk estimates, types of heterogeneity, baseline risk and the ecologic fallacy. Crit Care 2008;12(4):220.
8. Imberger G. Clinical guidelines and the question of uncertainty. Br J Anaesth 2013;111(5):700-2.
9. Thompson SG. Why sources of heterogeneity in meta-analysis should be investigated. BMJ 1994;309(6965):1351-5.
10. Barbateskovic M, Koster TM, Keus F, Gluud C, Møller MH, van der Horst ICC, Perner A, Wetterslev J. A protocol for constructing a tool to assess clinical heterogeneity in meta-analyses, assessment of interrater variability, and a pilot study of the association between clinical and statistical heterogeneity, Copenhagen Trial Unit 2019, http://ctu.dk/media/13724/2019-protocol-chims-protocol-manual-ver-11.0_11-03-2019.pdf
11. Kottner J, Audige L, Brorson S, Donner A, Gajewski BJ, Hrobjartsson A, Roberts C, Shoukri M, Streiner DL. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. J Clin Epidemiol 2011;64(1):96-106.
12. Gagnier JJ, Moher D, Boon H, Beyene J, Bombardier C. Investigating clinical heterogeneity in systematic reviews: a methodologic review of guidance in the literature. BMC Med Res Methodol 2012;12:111.
13. Gagnier JJ, Morgenstern H, Altman DG, Berlin J, Chang S, McCulloch P, Sun X, Moher D. Consensus-based recommendations for investigating clinical heterogeneity in systematic reviews. BMC Med Res Methodol 2013;13:106.
14. Koster TM, Wetterslev J, Gluud C, Keus F, van der Horst ICC. Systematic overview and critical appraisal of meta-analyses of interventions in intensive care medicine. Acta Anaesthesiol Scand 2018;62(1041-9).
15. Koster TM, Wetterslev J, Gluud C, Jakobsen JC, Kaufmann T, Eck RJ, Koster G, Hiemstra B, van der Horst ICC, Keus E. Apparently conclusive meta-analyses on interventions in critical care may be inconclusive-a meta-epidemiological study. J Clin Epidemiol 2019;114:1-10.
16. Gerke O, Moller S, Debrabant B, Halekoh U. Experience applying the guidelines for reporting reliability and agreement studies (GRRAS) indicated five questions should be addressed in the planning phase from a statistical point of view. Diagnostics 2018;8(4):E69.
17. Gisev N, Bell JS, Chen TF. Interrater agreement and interrater reliability: key concepts, approaches, and applications. Res Social Adm Pharm 2013;9(3):330-8.
18. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327(7414):557-60.
19. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159-74.
20. Whiting P, Savovic J, Higgins JP, Caldwell DM, Reeves BC, Shea B, Davies P, Kleijnen J, Churchill R. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol 2016;69:225-34.

21. Sterne JA, Hernan MA, Reeves BC, Savovic J, Berkman ND, Viswanathan M, Henry D, Altman DG, Ansari MT, Boutron I, Carpenter JR, Chan AW, Churchill R, Deeks JJ, Hrobjartsson A, Kirkham J, Juni P, Loke YK, Pigott TD, Ramsay CR, Regidor D, Rothstein HR, Sandhu L, Santaguida PL, Schunemann HJ, Shea B, Shrier I, Tugwell P, Turner L, Valentine JC, Waddington H, Waters E, Wells GA, Whiting PF, Higgins JP. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016;355:i4919.
22. Wetterslev J, Thorlund K, Brok J, Gluud C. Estimating required information size by quantifying diversity in random-effects model meta-analyses. BMC Med Res Methodol 2009;9:86.

Electronic Supplementary Material (ESM)

This supplementary has been provided by the authors to give readers additional information about their work.

Supplement to: Barbateskovic M, Krauss RK, Collet MO, Larsen LK, Jakobsen JC, Perner A, Wetterslev J. Pharmacological interventions for prevention and management of delirium in intensive care patients: a systematic overview of reviews and meta-analyses

Supplementary material for Pharmacological interventions for prevention and management of delirium in intensive care patients: a systematic overview of reviews and meta-analyses

Table 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Checklist

Section/topic	#	Checklist item	Line number(s)
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1-3
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	26-53
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	64-101
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	102-104
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	108-109
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	114-145
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	147-152, Electronic supplementary
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Electronic supplementary
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	155-164
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	165-167

Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	169-180
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	182-184, 197-199
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	N/A, results presented narratively, summary measures stated in protocol[1]
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	186-195
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	N/A, specified in protocol
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	N/A, specified in protocol
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	207, PRISMA flowchart
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	214-148, Table 2, ESM Table 3
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	Table 1
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	Table 2, ESM Table 3
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	N/A, results presented narratively
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	N/A
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	N/A, results presented narratively
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	357-364

Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	378-406
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	473-489
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	517-520

Search strategies

Cochrane Library from inception to 21.07.17 – 336 hits

#1 MeSH descriptor: [Antipsychotic Agents] explode all trees
#2 MeSH descriptor: [Hypnotics and Sedatives] explode all trees
#3 MeSH descriptor: [Benzodiazepines] explode all trees
#4 MeSH descriptor: [Analgesics, Opioid] explode all trees
#5 MeSH descriptor: [Melatonin] explode all trees
#6 (medication or drug* or agent* or pharmacologic* or antipsychotic* or sedative* or opioid* or benzodiazepin*or aripiprazole or clozapine or haloperidol or olanzapine or quetiapine or risperidone or ziprasidone or dexmedetomidine or clonidine or cholinesterase inhibitor or rivastigmine or donepezil or melatonin* or ketamine):ti,ab,kw
#7 (#1 or #2 or #3 or #4 or #5 or #6)
#8 MeSH descriptor: [Delirium] explode all trees
#9 (deliri*):ti,ab,kw
#10 ((acute organic) near/3 (psychosyndrome* or brain syndrome*)):ti,ab,kw
#11 ((acute brain) near/3 (dysfunction* or failure* or syndrome*)):ti,ab,kw
#12 ((postoperati* or post-operati* or postsurg* or post-surg*) near/1 (cognitive dysfunction or brain dysfunction or psychosis)):ti,ab,kw
#13 (acute) near (psycho-organic syndrome* or psychoorganic syndrome* or organic psychosyndrome* or organic psycho-syndrome*):ti,ab,kw
#14 (#8 or #9 or #10 or #11 or #12 or #13)
#15 MeSH descriptor: [Critical Care] explode all trees
#16 MeSH descriptor: [Intensive Care Units] explode all trees
#17 MeSH descriptor: [Heart Arrest] explode all trees
#18 MeSH descriptor: [Myocardial Infarction] explode all trees
#19 MeSH descriptor: [Shock] explode all trees
#20 MeSH descriptor: [Craniocerebral Trauma] explode all trees
#21 MeSH descriptor: [Stroke] explode all trees
#22 MeSH descriptor: [Sepsis] explode all trees
#23 MeSH descriptor: [Shock, Septic] explode all trees
#24 MeSH descriptor: [Thoracic Surgery] explode all trees
#25 MeSH descriptor: [Thorax] explode all trees
#26 (((intensive or critical*) near/3 (care or unit or department* or ill*)) or ICU):ti,ab,kw
#27 (cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct* or shock or (traumatic brain injury or TBI or head trauma) or (stroke or intracranial bleeding or intracranial hemorrhage) or (sepsis or septic shock)) or ((cardiothoracic or thorax or thoracic or chest or heart or cardiac) near/2 (surgical or surgery or operat*)) or ((acute*) near/2 (surgery or operat*)):ti,ab,kw
#28 (#15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27)
#29 (#7 and #14 and #28)

MEDLINE (Ovid) 1950 to 21.07.17 – 1362 hits

1. exp Antipsychotic Agents/
2. exp "Hypnotics and Sedatives"/
3. exp Benzodiazepines/
4. exp Analgesics, Opioid/
5. exp Melatonin/
6. (medication or drug* or agent* or pharmacologic* or antipsychotic* or sedative* or opioid* or benzodiazepin* or aripiprazole or clozapine or haloperidol or olanzapine or quetiapine or risperidone or ziprasidone or dexmedetomidine or clonidine or cholinesterase inhibitor or rivastigmine or donepezil or melatonin* or ketamine).tw.
7. 1 or 2 or 3 or 4 or 5 or 6
8. exp Delirium/
9. deliri*.tw.
10. (acute organic adj1 (psychosyndrome* or brain syndrome*)).tw.
11. (acute brain adj1 (dysfunction* or failure* or syndrome*)).tw.
12. ((postoperati* or post-operati* or postsurg* or post-surg*) adj1 (cognitive dysfunction or brain dysfunction or psychosis)).tw.
13. ((psycho-organic syndrome* or psychoorganic syndrome* or organic psychosyndrome* or organic psycho-syndrome*) adj3 acute).tw. 14. 8 or 9 or 10 or 11 or 12 or 13
15. exp Critical Care/
16. exp Intensive-Care-Units/
17. exp Heart Arrest/
18. exp Myocardial Infarction/
19. exp Shock/
20. exp Craniocerebral Trauma/
21. exp Stroke/
22. exp Sepsis/
23. exp Shock, Septic/
24. (((intensive or critical*) adj3 (care or unit or department* or ill*)) or ICU).tw.
25. (cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct* or shock or (traumatic brain injury or TBI or head trauma) or (stroke or intracranial bleeding or intracranial hemorrhage) or (sepsis or septic shock)).tw.
26. exp Thoracic Surgery/
27. exp Thorax/su [Surgery]
28. ((cardiothoracic or thorax or thoracic or chest or heart or cardiac) adj2 (surgical or surgery or operat*)).tw.
29. (acute* adj2 (surgery or operat*)).tw.
30. 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29
31. 7 and 14 and 30

Embase (Ovid) 1974 to 21.07.17 – 2165 hits

1. *neuroleptic agent/
2. *sedative agent/

3. *opiate/
4. *benzodiazepine derivative/
5. *cholinesterase inhibitor/
6. *melatonin/
7. (medication or drug* or agent* or pharmacologic* or antipsychotic* or sedative* or opioid* or benzodiazepin* or aripiprazole or clozapine or haloperidol or olanzapine or quetiapine or risperidone or ziprasidone or dexmedetomidine or clonidine or cholinesterase inhibitor or rivastigmine or donepezil or melatonin* or ketamine).tw.
8. 1 or 2 or 3 or 4 or 5 or 6 or 7
9. *delirium/
10. deliri*.tw.
11. (acute organic adj1 (psychosyndrome* or brain syndrome*)).tw.
12. (acute brain adj1 (dysfunction* or failure* or syndrome*)).tw.
13. ((postoperati* or post-operati* or postsurg* or post-surg*) adj1 (cognitive dysfunction or brain dysfunction or psychosis)).tw.
14. ((psycho-organic syndrome* or psychoorganic syndrome* or organic psychosyndrome* or organic psycho-syndrome*) adj3 acute).tw.
15. 9 or 10 or 11 or 12 or 13 or 14
16. *critical illness/
17. *intensive care/
18. *intensive care unit/
19. *heart arrest/
20. *heart infarction/
21. *shock/
22. *traumatic brain injury/
23. *cerebrovascular accident/
24. *sepsis/
25. *septic shock/
26. (((intensive or critical*) adj3 (care or unit or department* or ill*)) or ICU).tw.
27. (cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct* or shock or (traumatic brain injury or TBI or head trauma) or (stroke or intracranial bleeding or intracranial hemorrhage) or (sepsis or septic shock)).tw.
28. exp heart surgery/
29. exp thorax/su [Surgery]
30. ((cardiothoracic or thorax or thoracic or chest or heart or cardiac) adj2 (surgical or surgery or operat*)).tw.
31. (acute* adj2 (surgery or operat*)).tw.
32. 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31
33. 8 and 15 and 32

Science Citation Index-Expanded 1900 to 21.07.17 – 546 hits

#6 (#5 AND #2 AND #1)
#5: (#4 OR #3)

#4: TS=(cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct* or shock or (traumatic brain injury or TBI or head trauma) or (stroke or intracranial bleeding or intracranial hemorrhage) or (sepsis or septic shock) or (cardiac surgery or heart surgery or thoracic surgery or thorax surgery) or (cardiac operation or heart operation or thoracic operation or thorax operation) or (acute* surgery or acute* operat*))
#3: TS=((((intensive or critical*) near3 (care or unit or department* or ill*)) or ICU))
#2: TI=((deliri*) or (acute organic psychosyndrome* or acute organic brain syndrome*) or (acute brain dysfunction* or acute brain failure* or acute brain syndrome*) or (postoperati* cognitive dysfunction or post-operati* cognitive dysfunction or postsurg* cognitive dysfunction or post-surg* cognitive dysfunction or postoperati* brain dysfunction or post-operati* brain dysfunction or postsurg* brain dysfunction or post-surg* brain dysfunction or postoperati* psychosis or post-operati* psychosis or postsurg* psychosis or post-surg* psychosis) or (acute psycho-organic syndrome* or acute psychoorganic syndrome* or acute organic psychosyndrome* or acute organic psycho-syndrome*))
#1: TOPIC: ((medication or drug* or agent* or pharmacologic* or antipsychotic* or sedative* or opioid* or benzodiazepin*or aripiprazole or clozapine or haloperidol or olanzapine or quetiapine or risperidone or ziprasidone or dexmedetomidine or clonidine or cholinesterase inhibitor or rivastigmine or donepezil or melatonin* or ketamine))

BIOSIS Previews 1969 to 21.07.17 – 263 hits

#6 (#5 AND #2 AND #1)
#5 (#4 OR #3)
#4 TOPIC: ((cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct* or shock or (traumatic brain injury or TBI or head trauma) or (stroke or intracranial bleeding or intracranial hemorrhage) or (sepsis or septic shock) or (cardiac surgery or heart surgery or thoracic surgery or thorax surgery) or (cardiac operation or heart operation or thoracic operation or thorax operation) or (acute* surgery or acute* operat*)))
#3 TOPIC: (((((intensive or critical*) near3 (care or unit or department* or ill*)) or ICU)))
#2 TITLE: (((deliri*) or (acute organic psychosyndrome* or acute organic brain syndrome*) or (acute brain dysfunction* or acute brain failure* or acute brain syndrome*) or (postoperati* cognitive dysfunction or post-operati* cognitive dysfunction or postsurg* cognitive dysfunction or post-surg* cognitive dysfunction or postoperati* brain dysfunction or post-operati* brain dysfunction or postsurg* brain dysfunction or post-surg* brain dysfunction or postoperati* psychosis or post-operati* psychosis or postsurg* psychosis or post-surg* psychosis) or (acute psycho-organic syndrome* or acute psychoorganic syndrome* or acute organic psychosyndrome* or acute organic psycho-syndrome*)))
#1 TOPIC: (((medication or drug* or agent* or pharmacologic* or antipsychotic* or sedative* or opioid* or benzodiazepin*or aripiprazole or clozapine or haloperidol or olanzapine or quetiapine or risperidone or ziprasidone or dexmedetomidine or clonidine or cholinesterase inhibitor or rivastigmine or donepezil or melatonin* or ketamine)))

Cinahl 1981 to 21.07.17 – 342 hits

S23 (S7 AND S10 AND S22)
S22 (S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21)
S21 AB (cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct* or shock or (traumatic brain injury or TBI or head trauma) or (stroke or intracranial bleeding or intracranial hemorrhage) or (sepsis or septic shock)or (cardiac surgery or heart surgery or thoracic surgery or thorax surgery) or (cardiac operation or heart operation or thoracic operation or thorax operation) or (acute* surgery or acute* operat*))
S20 AB (((intensive or critical*) and (care or unit or department* or ill*)) or ICU)
S19 MJ septic shock
S18 MJ sepsis
S17 MJ stroke
S16 MJ brain injuries
S15 MJ Shock
S14 MJ Myocardial Infarction
S13 MJ Heart Arrest
S12 MJ Intensive Care Units
S11 MJ critical care
S10 (S8 OR S9)
S9 AB (((deliri*) or (acute organic psychosyndrome* or acute organic brain syndrome*) or (acute brain dysfunction* or acute brain failure* or acute brain syndrome*) or (postoperati* cognitive dysfunction or post-operati* cognitive dysfunction or postsurg* cognitive dysfunction or post-surg* cognitive dysfunction or postoperati* brain dysfunction or post-operati* brain dysfunction or postsurg* brain dysfunction or post-surg* brain dysfunction or postoperati* psychosis or post-operati* psychosis or postsurg* psychosis or post-surg* psychosis) or (acute psycho-organic syndrome* or acute psychoorganic syndrome* or acute organic psychosyndrome* or acute organic psycho-syndrome*)))
S8 MJ Delirium
S7 (S1 OR S2 OR S3 OR S4 OR S5 OR S6)
S6 AB (medication or drug* or agent* or pharmacologic* or antipsychotic* or sedative* or opioid* or benzodiazepin*or aripiprazole or clozapine or haloperidol or olanzapine or quetiapine or risperidone or ziprasidone or dexmedetomidine or clonidine or cholinesterase inhibitor or rivastigmine or donepezil or melatonin* or ketamine)
S5 MJ Melatonin
S4 MJ Analgesics, Opioid
S3 MM Antianxiety Agents, Benzodiazepine
S2 MJ hypnotics and sedatives
S1 MJ Antipsychotic Agents

Latin American Caribbean Health Sciences Literature (LILACS) 1982 to 21.07.17 – 22 hits
http://bases.bireme.br/cgi-bin/wxislind.exe/iah/online/?IsisScript=iah/iah.xis&base=LILACS&lang=i

Medication or drug or drugs or agent or agents or pharmacologic or pharmacological or antipsychotic or antipsychotics or sedative or sedatives or opioid or opioids or benzodiazepine or benzodiazepines or aripiprazole or clozapine or haloperidol or olanzapine or quetiapine or risperidone or ziprasidone or dexmedetomidine or clonidine or cholinesterase inhibitor or rivastigmine or donepezil or melatonin or melatonine or ketamine

AND

Delirium or acute organic psychosyndrome or acute organic brain syndrome or acute brain dysfunction or acute brain failure or acute brain syndrome or postoperative cognitive dysfunction or post-operation cognitive dysfunction or postsurgical cognitive dysfunction or post-surgical cognitive dysfunction or postoperative brain dysfunction or post-operative brain dysfunction or postsurgical brain dysfunction or post-surgical brain dysfunction or postoperative psychosis or post-operative psychosis or postsurgical psychosis or post-surgical psychosis or acute psycho-organic syndrome or acute psychoorganic syndrome or acute organic psychosyndrome or acute organic psycho-syndrome

AND

Intensive care or critical care or critically ill or critical illness or ICU or cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct or shock or traumatic brain injury or TBI or head trauma or stroke or intracranial bleeding or intracranial hemorrhage or sepsis or septic shock or acute surgery or emergency surgery or urgent surgery or trauma surgery or acute operation or emergency operation or urgent operation or trauma operation or acute resection or emergency resection or urgent resection or trauma resection or acute section or emergency section or urgent section or trauma section or cardiac surgery or heart surgery or thoracic surgery or thorax surgery or cardiac operation or heart operation or thoracic operation or thorax operation or acute surgery or acute operation

Data extraction form

REVIEW IDENTIFICATION

Authors	
Year	
Title	

REVIEW ELIGIBILITY

Review			**Relevant participants**			**Relevant intervention**			**Relevant outcomes**		
Yes	No	Unclear	Yes	No	Unclear	Yes	No	Unclear	Yes	No	Unclear
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

DO NOT PROCEED IF ANY OF THE ABOVE ANSWERS IS 'NO'

Include ☐	**Exclude** ☐ Record reason for exclusion

Recommendation on the use of haloperidol for the management of delirium	Yes	No	Not stated
	☐	☐	☐

PRISMA CHECKLIST

Section/Topic	**#**	**Checklist item**	**Reported on page #**
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both	
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known	
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS)	
METHODS			

Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale	
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched	
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated	
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis)	
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators	
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made	
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis	
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means)	
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I2) for each meta-analysis	
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies)	
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified	
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram	
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations	
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12)	
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot	
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency	
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15)	
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16])	
DISCUSSION			

Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers)	
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias)	
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research	
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review	

REVIEW ASSESSED TO BE SYSTEMATIC ACCORDING TO PRISMA

YES	NO
☐	☐

RISK OF BIAS ASSESSMENT OF SYSTEMATIC REVIEWS USING ROBIS

Identifying concerns with the review

DOMAIN 1: STUDY ELIGIBILITY CRITERIA		
Describe the study eligibility criteria, any restrictions on eligibility and whether there was evidence that objectives and eligibility criteria were pre-specified:		
1.1	Did the review adhere to pre-defined objectives and eligibility criteria?	Y/PY/PN/N/NI
1.2	Were the eligibility criteria appropriate for the review question?	Y/PY/PN/N/NI
1.3	Were eligibility criteria unambiguous?	Y/PY/PN/N/NI
1.4	Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Y/PY/PN/N/NI
1.5	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	Y/PY/PN/N/NI
Concerns regarding specification of study eligibility criteria Rationale for concern:		LOW/HIGH/UNCLEAR

DOMAIN 2: IDENTIFICATION AND SELECTION OF STUDIES		
Describe methods of study identification and selection (e.g. number of reviewers involved):		
2.1	Did the search include an appropriate range of databases/electronic sources for published and unpublished reports?	Y/PY/PN/N/NI
2.2	Were methods additional to database searching used to identify relevant reports?	Y/PY/PN/N/NI
2.3	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	Y/PY/PN/N/NI

2.4	Were restrictions based on date, publication format, or language appropriate?	Y/PY/PN/N/NI
2.5	Were efforts made to minimise error in selection of studies?	Y/PY/PN/N/NI
Concerns regarding methods used to identify and/or select studies Rationale for concern:		LOW/HIGH/UNCLEAR

DOMAIN 3: DATA COLLECTION AND STUDY APPRAISAL		
Describe methods of data collection, what data were extracted from studies or collected through other means, how risk of bias was assessed (e.g. number of reviewers involved) and the tool used to assess risk of bias:		
3.1	Were efforts made to minimise error in data collection?	Y/PY/PN/N/NI
3.2	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Y/PY/PN/N/NI
3.3	Were all relevant study results collected for use in the synthesis?	Y/PY/PN/N/NI
3.4	Was risk of bias (or methodological quality) formally assessed using appropriate criteria?	Y/PY/PN/N/NI
3.5	Were efforts made to minimise error in risk of bias assessment?	Y/PY/PN/N/NI
Concerns regarding methods used to collect data and appraise studies Rationale for concern:		LOW/HIGH/UNCLEAR

DOMAIN 4: SYNTHESIS AND FINDINGS		
Describe synthesis methods:		
4.1	Did the synthesis include all studies that it should?	Y/PY/PN/N/NI
4.2	Were all pre-defined analyses reported or departures explained?	Y/PY/PN/N/NI
4.3	Was the synthesis appropriate given the nature and similarity in the research questions, study designs and outcomes across included studies?	Y/PY/PN/N/NI
4.4	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Y/PY/PN/N/NI
4.5	Were the findings robust, e.g. as demonstrated through funnel plot or sensitivity analyses?	Y/PY/PN/N/NI
4.6	Were biases in primary studies minimal or addressed in the synthesis?	Y/PY/PN/N/NI
Concerns regarding the synthesis and findings Rationale for concern:		LOW/HIGH/UNCLEAR

Y=YES, PY=PROBABLY YES, PN=PROBABLY NO, N=NO, NI=NO INFORMATION

Judging risk of bias

Summarize the concerns identified during 'Identifying concerns with the review' assessment:

Domain	Concern	Rationale for concern
1. Concerns regarding specification of study eligibility criteria		
2. Concerns regarding methods used to identify and/or select studies		
3. Concerns regarding used to collect data and appraise studies		
4. Concerns regarding the synthesis and findings		

RISK OF BIAS IN THE SYSTEMATIC REVIEW

Describe whether conclusions were supported by the evidence:		
A	Did the interpretation of findings address all of the concerns identified in Domains 1 to 4?	Y/PY/PN/N/NI
B	Was the relevance of identified studies to the review's research question appropriately considered?	Y/PY/PN/N/NI
C	Did the reviewers avoid emphasizing results on the basis of their statistical significance?	Y/PY/PN/N/NI
Risk of bias in the systematic review		RISK: LOW/HIGH/UNCLEAR

Y=YES, PY=PROBABLY YES, PN=PROBABLY NO, N=NO, NI=NO INFORMATION

CHARACTERISTICS OF SYSTEMATIC REVIEW

Number of trials included	
Number of participants included	
ICU population (e.g. medical)	
Diagnostic criteria of delirium	
Type of pharmacological agent(s) included	
Primary and secondary outcomes	
Results on primary and secondary outcomes	
Type of meta-analytic and sequential analysis used	
Authors' conclusion	

Table 2. Reviews checked against the PRISMA criteria[a]

	Review identification	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15	#16	#17	#18	#19	#20	#21	#22	#23	#24	#25	#26	#27
1	Adams [2]	1	0	1	1	0	0	1	0	1	0	0	0	0	0	0	0	1	1	1	0	0	0	0	1	1	1	1
2	Al Qadheeb [3]	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1
3	Bathula [4]	1	1	1	1	0	1	1	0	1	1	1	1	0	0	0	0	0	1	1	0	0	0	0	1	1	1	1
4	Cao [5]	1	0	1	0	0	1	1	1	1	1	0	1	1	1	0	1	1	1	0	1	1	0	1	1	1	1	0
5	Chen [6]	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	Constantin [7]	1	1	1	1	0	1	1	0	1	1	1	0	1	1	1	0	1	1	0	0	1	0	0	1	1	1	1
7	Devlin [8]	0	0	1	1	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0
8	Devlin [9]	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
9	Fraser [10]	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1
10	Gerlach [11]	0	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	1	0
11	Hawkins [12]	0	1	1	1	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	1	0
12	Hoy [13]	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
13	Lawrance [14]	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
14	Li [15]	1	0	1	1	0	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	Liao [16]	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0	1	0	1	1	0
16	Lin [17]	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
17	Lonergan [18]	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	1	1	0	0	0	1	0	1	1
18	Mo [19]	0	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	1	1	1	0

19	Nelson [20]	0	0	1	1	0	1	1	0	1	1	1	0	0	0	0	0	1	1	0	0	0	0	0	1	1	1	1
20	Pasin [21]	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
21	Porhomayon [22]	1	1	1	1	0	1	1	0	1	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	1	1	0
22	Rea [23]	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0
23	Schrijver [24]	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	0	1	1	1	0	0	1	0	1	1	1	0
24	Serafim [25]	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	0	1	1	1	1	1	0	0	1	1	1	1
25	Szumita [26]	0	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0
26	Tan [27]	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
27	Teitelbaum [28]	0	0	1	0	0	0	1	1	1	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	1	0
28	Teslyar [29]	1	1	1	1	0	1	1	0	1	1	0	1	1	1	1	0	1	1	1	1	1	0	0	1	1	1	0
29	Wang [30]	1	0	1	0	0	1	1	1	1	0	0	1	1	1	0	0	1	1	0	1	1	0	1	1	1	1	0
30	Xia [31]	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
31	Zaal [32]	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1
32	Restrepo Bernal [33]	0	1	1	1	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
33	Bledowski [34]	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	Celis Rodriguez [35]	0	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35	Elefritz [36]	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
36	Flannery [37]	1	1	1	1	0	1	1	0	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	1	1	1	0
37	Geng [38]	1	1	1	1	0	1	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	1
38	Girardis [39]	0	1	1	1	0	0	1	0	0	1	0	1	0	0	1	0	0	1	0	0	0	1	0	1	1	0	0
39	Khan [40]	1	1	1	1	0	1	1	0	1	1	1	1	0	0	0	0	1	1	1	0	0	0	0	1	1	1	1

40	Nguyen [41]	0	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0
41	Pelland [42]	1	1	1	1	0	1	1	0	1	1	1	0	0	0	0	0	1	1	0	0	0	0	0	1	0	1	0
42	Rosenzweig [43]	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	1	1
43	Santos [44]	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	1	1	1	0
44	Tran [45]	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1
45	Tremblay [46]	0	1	1	0	0	0	1	0	1	0	0	0	0	0	0	0	1	1	0	1	0	0	0	1	0	1	0
46	Zhang [47]	1	0	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	0	0	0	0	1	1	0	0
47	Ford [48]	0	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
48	Gosch [49]	0	0	1	1	0	1	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0
49	Hirota [50]	1	1	1	1	0	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0
50	Liu [51]	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
51	Meagher [52]	0	0	1	0	0	1	1	0	0	0	0	1	0	0	0	0	1	1	1	0	0	0	0	1	0	1	0
52	Orena [53]	1	1	1	1	0	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	1	1	1	1
53	Schrader [54]	0	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
54	Sockalingam [55]	0	1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
55	Tabet [56]	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0
56	Tse [57]	0	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	1
57	Zhang [58]	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

[a] #1 Title, #2 Abstract, #3 Rationale, #4 Objectives, #5 Protocol, #6 Eligibility criteria, #7 Information sources, #8 Search, #9 Study selection, #10 Data collection process, #11 Data items, #12 RoB in individual studies, #13 Summary measures, #14 Synthesis of results, #15 RoB across studies, #16 Additional analyses, #17 Study selection, #18 Study characteristics, #19 RoB within studies, #20 Results of individual studies, #21 Synthesis of results, #22 RoB across studies, #23 Additional analysis, #24 Summary of evidence, #25 Limitations, #26 Conclusions, #27 Funding

Table 3. Comparison of dexmedetomidine vs other sedatives on additional outcomes reported by the systematic and semi systematic reviews

	Chen [6]	Tan [27]	Lin [17]	Fraser [10]	Xia [31]	Tran [45]	Liu [51]
Hypotension	RR 1.22, 0.86 to 1.74; 6 RCTs including 1587 patients	RR 1.43, 0.78–2.6; 12 RCTs including 1545 patients	RR 1.06, 0.72 to 1.56; 4 RCTs including 739 patients		RR 1.12, 0.86 to 1.47; 6 RCTs including 1015 patients		RR 1.12, 0.90 to 1.39; 5 RCTs including 564 patients
Hypertension	3 trials included, meta-analysis not performed				RR 1.56, 1.11 to 2.20, 3 RCTs including 846 ICU patients		
Bradycardia	RR 2.11, 1.39 to 3.20; 6 RCTs including 1587 mechanically ventilated patients	RR 1.82, 0.66 to 5.03; 10 RCTs including 1164 patients	RR 2.08, 1.16 to 3.74; 3 RCTs including 650 patients		RR 1.36, 0.85 to 2.18; 2 RCTs including 788 ICU patients		RR 3.17, 1.41 to 7.10, 4 RCTs including 475 patients after cardiac surgery
Atrial fibrillation		RR 0.95, 0.68 to 1.33; number of RCTs and patients included in analysis was not reported	RR 0.90, 0.62 to 1.29; 3 RCTs including 683 patients				RR 1.04, 0.84 to 1.30; 6 RCTs including 854 patients
Tachycardia	4 RCTs included, meta-analysis not performed		RR 0.27, 0.08 to 0.97, 3 RCTs including 683 patients				
First-degree atrioventricular block	2 RCTs included, meta-analysis not performed						
Hyperglycaemia	3 RCTs included, meta-analysis not performed	RR 1.05, 0.64 to 1.71; number of RCTs and patients included in analysis was not reported	RR 0.78, 0.61 to 0.99, 3 RCTs including 622 patients				
Hypoglycaemia	2 RCTs included, meta-analysis not performed						
Length of ICU stay	Geometric mean by -0.15 (-0.15 to -0.01), corresponding to a reduction of 14% in the geometric mean (0.01% to 24%), 5 RCTs, 1223 patients	WMD -0.48 d, -0.18 to -0.78 d, 12 RCTs including 1264 patients	MD -3.44, -11.40 to 4.52; 4 RCTs including 534 patients	Beneficial effect of non-benzodiazepine, (4/6 RCTs included dexmedetomidine as comparator) use compared	MD -0.81 d, CI -1.48 to -0.15; 5 RCTs including 655 patients, respectively)	2 trials included, meta-analysis not performed	MD -9.72 h, −29.22 to 9.78; 5 RCTs including 448 patients

				with benzodiazepine sedation, MD –1.64, -2.57 to -0.70; 6 RCTs including 1225 patients			
Length of hospital stay		found no evidence of a difference when comparing dexmedetomidine with traditional sedatives (neither summary measures nor number of RCTs and patients included in analysis were reported).	MD -0.38, -0.95 to 0.19; 3 RCTs including 445 patients				
Duration of mechanical ventilation	Geometric mean duration of mechanical ventilation reduced by 0.25 (0.10 to 0.40), corresponding to a reduction of 22% in the geometric mean (10% to 33%), 4 RCTs including 1120 patients	WMD -0.51, -1.75 to 0.73; 12 RCTs including 1901 patients	MD -0.87, -1.67 to -0.07, 6 RCTs including 857 patients	Found a beneficial effect on non-benzodiazepine use compared with benzodiazepine sedation (-1.87, -2.51 to -1.22; 4 RCTs including 1101 patients	MD 0.53 h, -2.66 to 3.72; 5 RCTs including 895 patients	1 RCT included	MD -0.95 h, −1.26 to −0.64; 7 RCTs including 807 patients
Proportion of sedation time spent at target sedation level	Reported participants treated with dexmedetomidine overall spent a higher proportion of time at the target sedation level. Meta-analysis not performed						
Duration of weaning	1 trial included						
Reintubation			RR 1.21, 0.33 to 4.41, 2 RCTs including 355 patients				
Coma	1 RCT included						
Self-extubation		RR 1.36 0.31 to 5.90; number of RCTs and patients included in analysis was not reported					
Nausea and vomiting		RR 1.03, 0.66 to 1.59; number of RCTs and patients	RR 1.02, 0.72 to 1.46, 3 RCTs				

	included in analysis was not reported	including 622 patients	
Myocardial infarction	RR 0.62, 0.07 to 5.63; number of RCTs and patients included in analysis was not reported		
Morphine		MD 1.25, -0.98 to 3.49; 3 RCTs including 205 patients	
Any postoperative infection		RR 0.89, 0.38 to 2.12, 3 RCTs including 683 patients	
Intracranial pressure			2 RCTs included, meta-analysis not performed
Cerebral perfusion pressure			1 RCT included
Arterial pressure			3 RCTs included, meta-analysis not performed

Reference List

1. Barbateskovic M, Larsen LK, Oxenboll-Collet M, et al. Pharmacological interventions for delirium in intensive care patients: a protocol for an overview of reviews. *Syst Rev* 2016; 5(1): 211.
2. Adams R, Brown GT, Davidson M, et al. Efficacy of dexmedetomidine compared with midazolam for sedation in adult intensive care patients: a systematic review. *Br J Anaesth* 2013; 111(5): 703-10.
3. Al-Qadheeb NS, Balk EM, Fraser GL, et al. Randomized ICU trials do not demonstrate an association between interventions that reduce delirium duration and short-term mortality: a systematic review and meta-analysis. *Crit Care Med* 2014; 42(6): 1442-54.
4. Bathula M and Gonzales JP. The pharmacologic treatment of intensive care unit delirium: a systematic review. *Ann Pharmacother* 2013; 47(9): 1168-74.
5. Cao F, Zhang H, and Feng X. Role of dexmedetomidine in the perioperative period of patients undergoing coronary artery bypass graft surgery: A meta-analysis. *Med J Chin PLA,* 2014; 39(12): 981-86.
6. Chen K, Lu Z, Xin YC, et al. Alpha-2 agonists for long-term sedation during mechanical ventilation in critically ill patients. *Cochrane Database Syst Rev* 2015; 1: Cd010269.
7. Constantin JM, Momon A, Mantz J, et al. Efficacy and safety of sedation with dexmedetomidine in critical care patients: a meta-analysis of randomized controlled trials. *Anaesth Crit Care Pain Med* 2016; 35(1): 7-15.
8. Devlin JW and Skrobik Y. Antipsychotics for the prevention and treatment of delirium in the intensive care unit: what is their role? *Harv Rev Psychiatry* 2011; 19(2): 59-67.
9. Devlin JW, Al-Qadhee NS, and Skrobik Y. Pharmacologic prevention and treatment of delirium in critically ill and non-critically ill hospitalised patients: a review of data from prospective, randomised studies. *Best Pract Res Clin Anaesthesiol* 2012; 26(3): 289-309.
10. Fraser GL, Devlin JW, Worby CP, et al. Benzodiazepine versus nonbenzodiazepine-based sedation for mechanically ventilated, critically ill adults: a systematic review and meta-analysis of randomized trials. *Crit Care Med* 2013; 41(9 Suppl 1): S30-8.
11. Gerlach AT, Murphy CV, and Dasta JF. An updated focused review of dexmedetomidine in adults. *Ann Pharmacother* 2009; 43(12): 2064-74.
12. Hawkins SB, Bucklin M, and Muzyk AJ. Quetiapine for the treatment of delirium. *J Hosp Med* 2013; 8(4): 215-20.
13. Hoy SM and Keating GM. Dexmedetomidine: a review of its use for sedation in mechanically ventilated patients in an intensive care setting and for procedural sedation. *Drugs* 2011; 71(11): 1481-501.
14. Lawrence KR and Nasraway SA. Conduction disturbances associated with administration of butyrophenone antipsychotics in the critically ill: a review of the literature. *Pharmacotherapy* 1997; 17(3): 531-7.
15. Li B, Wang H, Wu H, and Gao C. Neurocognitive dysfunction risk alleviation with the use of dexmedetomidine in perioperative conditions or as ICU sedation: a meta-analysis. *Medicine (Baltimore)* 2015; 94(14): e597.
16. T LYYJXDYLDCGLG. Efficacy and Safety of Antipsychotics in the Treatment of Delirium: A Systematic Review. *Chinese Journal of Evidence-Based Medicine* 2015; (12): 1401-6.
17. Lin YY, He B, Chen J, and Wang ZN. Can dexmedetomidine be a safe and efficacious sedative agent in post-cardiac surgery patients? a meta-analysis. *Crit Care* 2012; 16(5): R169.
18. Lonergan E, Luxenberg J, and Areosa Sastre A. Benzodiazepines for delirium. *Cochrane Database Syst Rev* 2009; (4): Cd006379.
19. Mo Y and Zimmermann AE. Role of dexmedetomidine for the prevention and treatment of delirium in intensive care unit patients. *Ann Pharmacother* 2013; 47(6): 869-76.

20. Nelson S, Muzyk AJ, Bucklin MH, Brudney S, and Gagliardi JP. Defining the Role of Dexmedetomidine in the Prevention of Delirium in the Intensive Care Unit. *Biomed Res Int* 2015; 2015: 635737.
21. Pasin L, Landoni G, Nardelli P, et al. Dexmedetomidine reduces the risk of delirium, agitation and confusion in critically Ill patients: a meta-analysis of randomized controlled trials. *J Cardiothorac Vasc Anesth* 2014; 28(6): 1459-66.
22. Porhomayon J, Joude P, Adlparvar G, El-Solh AA, and Nader ND. The Impact of High Versus Low Sedation Dosing Strategy on Cognitive Dysfunction in Survivors of Intensive Care Units: A Systematic Review and Meta-Analysis. *J Cardiovasc Thorac Res* 2015; 7(2): 43-8.
23. Rea RS, Battistone S, Fong JJ, and Devlin JW. Atypical antipsychotics versus haloperidol for treatment of delirium in acutely ill patients. *Pharmacotherapy* 2007; 27(4): 588-94.
24. Schrijver EJ, de Graaf K, de Vries OJ, Maier AB, and Nanayakkara PW. Efficacy and safety of haloperidol for in-hospital delirium prevention and treatment: A systematic review of current evidence. *Eur J Intern Med* 2016; 27: 14-23.
25. Serafim RB, Bozza FA, Soares M, et al. Pharmacologic prevention and treatment of delirium in intensive care patients: A systematic review. *J Crit Care* 2015; 30(4): 799-807.
26. Szumita PM, Baroletti SA, Anger KE, and Wechsler ME. Sedation and analgesia in the intensive care unit: evaluating the role of dexmedetomidine. *Am J Health Syst Pharm* 2007; 64(1): 37-44.
27. Tan JA and Ho KM. Use of dexmedetomidine as a sedative and analgesic agent in critically ill adult patients: a meta-analysis. *Intensive Care Med* 2010; 36(6): 926-39.
28. Teitelbaum JS, Ayoub O, and Skrobik Y. A critical appraisal of sedation, analgesia and delirium in neurocritical care. *Can J Neurol Sci* 2011; 38(6): 815-25.
29. Teslyar P, Stock VM, Wilk CM, et al. Prophylaxis with antipsychotic medication reduces the risk of post-operative delirium in elderly patients: a meta-analysis. *Psychosomatics* 2013; 54(2): 124-31.
30. H WLCXRXL. Effectiveness and Safety of Dexmedetomidine for Postoperative Sedation in Cardiac Patients: A Meta-Analysis. *Chin J Evid-based Med* 2013; 13(1): 93-9.
31. Xia ZQ, Chen SQ, Yao X, et al. Clinical benefits of dexmedetomidine versus propofol in adult intensive care unit patients: a meta-analysis of randomized clinical trials. *J Surg Res* 2013; 185(2): 833-43.
32. Zaal IJ and Slooter AJ. Delirium in critically ill patients: epidemiology, pathophysiology, diagnosis and management. *Drugs* 2012; 72(11): 1457-71.
33. Restrepo Bernal D, Nino Garcia JA, and Ortiz Estevez DE. [Delirium Prevention]. *Rev Colomb Psiquiatr* 2016; 45(1): 37-45.
34. Bledowski J and Trutia A. A review of pharmacologic management and prevention strategies for delirium in the intensive care unit. *Psychosomatics* 2012; 53(3): 203-11.
35. Celis-Rodriguez E, Birchenall C, de la Cal MA, et al. Clinical practice guidelines for evidence-based management of sedoanalgesia in critically ill adult patients. *Med Intensiva* 2013; 37(8): 519-74.
36. Elefritz JL, Murphy CV, Papadimos TJ, and Lyaker MR. Methadone analgesia in the critically ill. *J Crit Care* 2016; 34: 84-8.
37. Flannery AH, Oyler DR, and Weinhouse GL. The Impact of Interventions to Improve Sleep on Delirium in the ICU: A Systematic Review and Research Framework. *Crit Care Med* 2016; 44(12): 2231-2240.
38. Geng J, Qian J, Cheng H, Ji F, and Liu H. The Influence of Perioperative Dexmedetomidine on Patients Undergoing Cardiac Surgery: A Meta-Analysis. *PLoS One* 2016; 11(4): e0152829.
39. Girardis M, Cantaroni C, Savoia G, Melotti R, and Conti G. A critical appraisal of the quality of analgosedation guidelines in critically ill patients. *Minerva Anestesiol* 2016; 82(2): 230-5.
40. Khan BA, Zawahiri M, Campbell NL, et al. Delirium in hospitalized patients: implications of current evidence on clinical practice and future avenues for research--a systematic evidence review. *J Hosp Med* 2012; 7(7): 580-9.

41. Nguyen HM and Pon D. Off-Label Use of Dexmedetomidine for the Treatment of Delirium in the Intensive Care Unit. *P t* 2016; 41(10): 642-643.
42. Pelland C and Trudel JF. [Atypical antipsychotic efficacy and safety in managing delirium: a systematic review and critical analysis]. *Psychol Neuropsychiatr Vieil* 2009; 7(2): 109-19.
43. Rosenzweig AB and Sittambalam CD. A new approach to the prevention and treatment of delirium in elderly patients in the intensive care unit. *J Community Hosp Intern Med Perspect* 2015; 5(4): 27950.
44. Santos E, Cardoso D, Neves H, et al. Effectiveness of haloperidol prophylaxis in critically ill patients with a high risk of delirium: a systematic review. *JBI Database System Rev Implement Rep* 2017; 15(5): 1440-1472.
45. Tran A, Blinder H, Hutton B, and English SW. A Systematic Review of Alpha-2 Agonists for Sedation in Mechanically Ventilated Neurocritical Care Patients. *Neurocrit Care* 2017.
46. Tremblay P and Gold S. Prevention of Post-operative Delirium in the Elderly Using Pharmacological Agents. *Can Geriatr J* 2016; 19(3): 113-126.
47. Zhang Z, Chen K, Ni H, Zhang X, and Fan H. Sedation of mechanically ventilated adults in intensive care unit: a network meta-analysis. *Sci Rep* 2017; 7: 44979.
48. Ford AH and Almeida OP. Pharmacological interventions for preventing delirium in the elderly. *Maturitas* 2015; 81(2): 287-92.
49. Gosch M and Nicholas JA. Pharmacologic prevention of postoperative delirium. *Z Gerontol Geriatr* 2014; 47(2): 105-9.
50. Hirota T and Kishi T. Prophylactic antipsychotic use for postoperative delirium: a systematic review and meta-analysis. *J Clin Psychiatry* 2013; 74(12): e1136-44.
51. Liu X, Xie G, Zhang K, et al. Dexmedetomidine vs propofol sedation reduces delirium in patients after cardiac surgery: A meta-analysis with trial sequential analysis of randomized controlled trials. *J Crit Care* 2017; 38: 190-196.
52. Meagher DJ, McLoughlin L, Leonard M, et al. What do we really know about the treatment of delirium with antipsychotics? Ten key issues for delirium pharmacotherapy. *Am J Geriatr Psychiatry* 2013; 21(12): 1223-38.
53. Orena EF, King AB, and Hughes CG. The role of anesthesia in the prevention of postoperative delirium: a systematic review. *Minerva Anestesiol* 2016; 82(6): 669-83.
54. Schrader SL, Wellik KE, Demaerschalk BM, et al. Adjunctive haloperidol prophylaxis reduces postoperative delirium severity and duration in at-risk elderly patients. *Neurologist* 2008; 14(2): 134-7.
55. Sockalingam S, Parekh N, Bogoch, II, et al. Delirium in the postoperative cardiac patient: a review. *J Card Surg* 2005; 20(6): 560-7.
56. Tabet N and Howard R. Pharmacological treatment for the prevention of delirium: review of current evidence. *Int J Geriatr Psychiatry* 2009; 24(10): 1037-44.
57. Tse L, Schwarz SK, Bowering JB, et al. Pharmacological risk factors for delirium after cardiac surgery: a review. *Curr Neuropharmacol* 2012; 10(3): 181-96.
58. Zhang H, Lu Y, Liu M, et al. Strategies for prevention of postoperative delirium: a systematic review and meta-analysis of randomized trials. *Crit Care* 2013; 17(2): R47.

Electronic Supplementary Material

This supplementary has been provided by the authors to give readers additional information about their work.

Supplement to: Barbateskovic M, Krauss SR, Collet MO, Andersen-Ranberg NC, Mathiesen O, Jakobsen JC, Perner A, Wetterslev J. Haloperidol for the treatment of delirium in critically ill patients: a systematic review with meta-analysis and Trial Sequential Analysis

Correspondence: Barbateskovic M, Copenhagen Trial Unit, Centre for Clinical Intervention Research, DK2100 Copenhagen, Denmark, Tel: +45 3545 7113, E-mail: marija.barbateskovic@ctu.dk

Supplementary for

Haloperidol for the treatment of delirium in critically ill patients: a systematic review with meta-analysis and Trial Sequential Analysis

PRISMA CHECKLIST

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	3-4
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	5
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	6
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	ESM
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	7
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	7

Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	7
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7-8
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	8
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	7
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	8-10
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	10, ESM
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Table 1, ESM
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	11-12, Figure 2
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	Figure 3+4, ESM
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	12-15, Figure 3+4, ESM
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	12-15
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	12-15, ESM
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	16, Table 2
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	16-17
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	18
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	19

SEARCH STRATEGIES

Cochrane Central Register of Controlled Trials (CENTRAL)

(The Cochrane Library, from inception to 5 March 2019)

#1 MeSH descriptor: [Haloperidol]
#2 ((alased) or (aloperidin) or (aloperidine) or (apo-haloperidol) or (avant) or (binison) or (brotopon) or (celenase) or (cereen) or (cerenace) or (cizoren) or (depidol) or (dores) or (dozic) or (duraperidol) or (einalon s) or (fortunan) or (govotil) or (Haldol) or (haldol solutab) or (halidol) or (halo-p) or (halojust) or (halomed) or (haloneural) or (haloper) or (haloperidol) or (haloperidol hydrochoride) or (haloperidol intensol) or (haloperidol lactate) or (haloperil) or (haloperin) or (haloperitol) or (halopidol) or (halopol) or (halosten) or (haricon) or (haridol-d) or (keselan) or (linton) or (lodomer-2) or (mcn jr 1625) or (mcn jr1625) or (mixidol) or (novoperidol) or (nsc 170973) or (nsc170973) or (peluces) or (perida) or (peridol) or (peridor) or (r 1625) or (r1625) or (selezyme) or (seranace) or (serenace) or (serenase) or (serenelfi) or (siegoperidol) or (sigaperidol) or (trancodol-10) or (trancodol-5))
#3 (#1 or #2)

#4 MeSH descriptor: [Delirium] explode all trees
#5 deliri*
#6 (acute organic near (psychosyndrome* or brain syndrome*))
#7 (acute brain near (dysfunction* or failure* or syndrome*))
#8 ((postoperati* or post-operati* or postsurg* or post-surg*) near/1 (cognitive dysfunction or brain dysfunction or psychosis))
#9 ((psycho-organic syndrome* or psychoorganic syndrome* or organic psychosyndrome* or organic psycho-syndrome*) near/3 acute)
#10 (metabolic encephalopathy or exogenous psychosis)
#11 (acute confusion* or acute psycho-organic syndrome or obnubilat*)
#12 (clouded state or clouding of consciousness*)
#13 (cloud* near/3 consciousness*)
#14 (toxic near/1 (psychosis or confusion))
#15 (#4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14)

#16 (#3 and #15)

MEDLINE (Ovid)

From 1950 to 5 March 2019

1. exp Haloperidol/
2. ((alased) or (aloperidin) or (aloperidine) or (apo-haloperidol) or (avant) or (binison) or (brotopon) or (celenase) or (cereen) or (cerenace) or (cizoren) or (depidol) or (dores) or (dozic) or (duraperidol) or (einalon s) or (fortunan) or (govotil) or (Haldol) or (haldol solutab) or (halidol) or (halo-p) or (halojust) or (halomed) or (haloneural) or (haloper) or (haloperidol) or (haloperidol hydrochoride) or (haloperidol intensol) or (haloperidol lactate) or (haloperil) or (haloperin) or (haloperitol) or (halopidol) or (halopol) or (halosten) or (haricon) or (haridol-d) or (keselan) or (linton) or (lodomer-2) or (mcn jr 1625) or (mcn jr1625) or (mixidol) or (novoperidol) or (nsc 170973) or (nsc170973) or (peluces) or (perida) or (peridol) or (peridor) or (r 1625) or (r1625) or (selezyme) or (seranace) or (serenace) or (serenase) or (serenelfi) or (siegoperidol) or (sigaperidol) or (trancodol-10) or (trancodol-5)).mp.
3. 1 or 2

4. exp Delirium/
5. deliri*.mp.
6. (acute organic adj (psychosyndrome* or brain syndrome*)).mp.
7. (acute brain adj (dysfunction* or failure* or syndrome*)).mp.
8. ((postoperati* or post-operati* or postsurg* or post-surg*) adj1 (cognitive dysfunction or brain dysfunction or psychosis)).mp.
9. ((psycho-organic syndrome* or psychoorganic syndrome* or organic psychosyndrome* or organic psycho-syndrome*) adj3 acute).mp.
10. (metabolic encephalopathy or exogenous psychosis).mp.
11. (acute confusion* or acute psycho-organic syndrome or obnubilat*).mp.

12. (clouded state or clouding of consciousness*).mp.
13. (cloud* adj3 consciousness*).mp.
14. (toxic adj1 (psychosis or confusion)).mp.
15. 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14

16. 3 and 15

Embase (OvidSP)

From 1974 to 5 March 2019

1. *haloperidol/
2. *haloperidol decanoate/
3. *haloperidol-induced catalepsy/
4. *reduced haloperidol/
5. (alased or aloperidin or aloperidine or apo-haloperidol or avant or binison or brotopon or celenase or cereen or cerenace or cizoren or depidol or dores or dozic or duraperidol or einalon s or fortunan or govotil or Haldol or haldol solutab or halidol or halo-p or halojust or halomed or haloneural or haloper or haloperidol or haloperidol hydrochoride or haloperidol intensol or haloperidol lactate or haloperil or haloperin or haloperitol or halopidol or halopol or halosten or haricon or haridol-d or keselan or linton or lodomer-2 or mcn jr 1625 or mcn jr1625 or mixidol or novoperidol or nsc 170973 or nsc170973 or peluces or perida or peridol or peridor or r 1625 or r1625 or selezyme or seranace or serenace or serenase or serenelfi or siegoperidol or sigaperidol or trancodol-10 or trancodol-5).ti,ab,kw,tw.
6. 1 or 2 or 3 or 4 or 5

7. *delirium/
8. *postoperative delirium/
9. deliri*.ti,ab,kw.
10. (acute organic adj (psychosyndrome* or brain syndrome*)).ti,ab,kw.
11. (acute brain adj (dysfunction* or failure* or syndrome*)).ti,ab,kw.
12. ((postoperati* or post-operati* or postsurg* or post-surg*) adj1 (cognitive dysfunction or brain dysfunction or psychosis)).ti,ab,kw.
13. ((psycho-organic syndrome* or psychoorganic syndrome* or organic psychosyndrome* or organic psycho-syndrome*) adj3 acute).ti,ab,kw.
14. (metabolic encephalopathy or exogenous psychosis).ti,ab,kw.
15. (acute confusion* or acute psycho-organic syndrome or obnubilat*).ti,ab,kw.
16. (clouded state or clouding of consciousness*).ti,ab,kw.
17. (cloud* adj3 consciousness*).ti,ab,kw.
18. (toxic adj1 (psychosis or confusion)).ti,ab,kw.
19. 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18

20. 6 and 19

Science Citation Index (web of science)

From 1900 to 5 March 2019

#1 TOPIC: (((alased) or (aloperidin) or (aloperidine) or (apo-haloperidol) or (avant) or (binison) or (brotopon) or (celenase) or (cereen) or (cerenace) or (cizoren) or (depidol) or (dores) or (dozic) or (duraperidol) or (einalon s) or (fortunan) or (govotil) or (Haldol) or (haldol solutab) or (halidol) or (halo-p) or (halojust) or (halomed) or (haloneural) or (haloper) or (haloperidol) or (haloperidol hydrochoride) or (haloperidol intensol) or (haloperidol lactate) or (haloperil) or (haloperin) or (haloperitol) or (halopidol) or (halopol) or (halosten) or (haricon) or (haridol-d) or (keselan) or (linton) or (lodomer-2) or (mcn jr 1625) or (mcn jr1625) or (mixidol) or (novoperidol) or (nsc 170973) or (nsc170973) or (peluces) or (perida) or (peridol) or (peridor) or (r 1625) or (r1625) or (selezyme) or (seranace) or (serenace) or (serenase) or (serenelfi) or (siegoperidol) or (sigaperidol) or (trancodol-10) or (trancodol-5)))

#2 TOPIC: (deliri*)
#3 TI=(acute organic and psychosyndrome* or brain syndrome*)
#4 TI=(acute brain and dysfunction* or failure* or syndrome*)
#5 TI=(postoperati* or post-operati* or postsurg* or post-surg* and cognitive dysfunction or brain dysfunction or psychosis)

#6 TS=(psycho-organic syndrome* or psychoorganic syndrome* or organic psychosyndrome* or organic psycho-syndrome* near/3 acute)
#7 TS=(metabolic encephalopathy or exogenous psychosis)
#8 TS=(acute confusion* or acute psycho-organic syndrome or obnubilat*)
#9 TI=(clouded state or clouding of consciousness*)
#10 TS=(cloud* near/3 consciousness*)
#11 TI=(toxic and psychosis or confusion)

#12 (#11 OR #10 OR #9 OR #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2)

#13 (#12 AND #1)

Biosis Previews (web of science)

From 1969 to 5 March 2019

#1 TOPIC: (((alased) or (aloperidin) or (aloperidine) or (apo-haloperidol) or (avant) or (binison) or (brotopon) or (celenase) or (cereen) or (cerenace) or (cizoren) or (depidol) or (dores) or (dozic) or (duraperidol) or (einalon s) or (fortunan) or (govotil) or (Haldol) or (haldol solutab) or (halidol) or (halo-p) or (halojust) or (halomed) or (haloneural) or (haloper) or (haloperidol) or (haloperidol hydrochoride) or (haloperidol intensol) or (haloperidol lactate) or (haloperil) or (haloperin) or (haloperitol) or (halopidol) or (halopol) or (halosten) or (haricon) or (haridol-d) or (keselan) or (linton) or (lodomer-2) or (mcn jr 1625) or (mcn jr1625) or (mixidol) or (novoperidol) or (nsc 170973) or (nsc170973) or (peluces) or (perida) or (peridol) or (peridor) or (r 1625) or (r1625) or (selezyme) or (seranace) or (serenace) or (serenase) or (serenelfi) or (siegoperidol) or (sigaperidol) or (trancodol-10) or (trancodol-5)))

#2 TOPIC: (deliri*)
#3 TI=(acute organic and psychosyndrome* or brain syndrome*)
#4 TI=(acute brain and dysfunction* or failure* or syndrome*)
#5 TI=(postoperati* or post-operati* or postsurg* or post-surg* and cognitive dysfunction or brain dysfunction or psychosis)
#6 TS=(psycho-organic syndrome* or psychoorganic syndrome* or organic psychosyndrome* or organic psycho-syndrome* near/3 acute)
#7 TS=(metabolic encephalopathy or exogenous psychosis)
#8 TS=(acute confusion* or acute psycho-organic syndrome or obnubilat*)
#9 TI=(clouded state or clouding of consciousness*)
#10 TS=(cloud* near/3 consciousness*)
#11 TI=(toxic and psychosis or confusion)
#12 (#11 OR #10 OR #9 OR #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2)

#13 (#12 AND #1)

Cumulative Index to Nursing & Allied Health Literature (CINAHL)

From inception to 5 March 2019

S1 MW haloperidol
S2 TX ((alased) or (aloperidin) or (aloperidine) or (apo-haloperidol) or (avant) or (binison) or (brotopon) or (celenase) or (cereen) or (cerenace) or (cizoren) or (depidol) or (dores) or (dozic) or (duraperidol) or (einalon s) or (fortunan) or (govotil) or (Haldol) or (haldol solutab) or (halidol) or (halo-p) or (halojust) or (halomed) or (haloneural) or (haloper) or (haloperidol) or (haloperidol hydrochoride) or (haloperidol intensol) or (haloperidol lactate) or (haloperil) or (haloperin) or (haloperitol) or (halopidol) or (halopol) or (halosten) or (haricon) or (haridol-d) or (keselan) or (linton) or (lodomer-2) or (mcn jr 1625) or (mcn jr1625) or (mixidol) or (novoperidol) or (nsc 170973) or (nsc170973) or (peluces) or (perida) or (peridol) or (peridor) or (r 1625) or (r1625) or (selezyme) or (seranace) or (serenace) or (serenase) or (serenelfi) or (siegoperidol) or (sigaperidol) or (trancodol-10) or (trancodol-5))
S3 (S1 OR S2)

S4 MW delirium
S5 TX deliri*
S6 AB (acute organic and (psychosyndrome* or brain syndrome*))
S7 AB (acute brain and (dysfunction* or failure* or syndrome*))

S8 AB ((postoperati* or post-operati* or postsurg* or post-surg*) and (cognitive dysfunction or brain dysfunction or psychosis))
S9 AB ((psycho-organic syndrome* or psychoorganic syndrome* or organic psychosyndrome* or organic psycho-syndrome*) and acute)
S10 AB metabolic encephalopathy or exogenous psychosis
S11 AB acute confusion* or acute psycho-organic syndrome or obnubilat*
S12 AB clouded state or clouding of consciousness*
S13 AB cloud* and consciousness*
S14 AB (toxic and (psychosis or confusion))
S15 (S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14)

S16 (S3 AND S15)

Latin American Caribbean Health Sciences Literature (LILACS)

From inception date 5 March 2019

(alased) or (aloperidin) or (aloperidine) or (apo-haloperidol) or (avant) or (binison) or (brotopon) or (celenase) or (cereen) or (cerenace) or (cizoren) or (depidol) or (dores) or (dozic) or (duraperidol) or (einalon s) or (fortunan) or (govotil) or (Haldol) or (haldol solutab) or (halidol) or (halo-p) or (halojust) or (halomed) or (haloneural) or (haloper) or (haloperidol) or (haloperidol hydrochoride) or (haloperidol intensol) or (haloperidol lactate) or (haloperil) or (haloperin) or (haloperitol) or (halopidol) or (halopol) or (halosten) or (haricon) or (haridol-d) or (keselan) or (linton) or (lodomer-2) or (mcn jr 1625) or (mcn jr1625) or (mixidol) or (novoperidol) or (nsc 170973) or (nsc170973) or (peluces) or (perida) or (peridol) or (peridor) or (r 1625) or (r1625) or (selezyme) or (seranace) or (serenace) or (serenase) or (serenelfi) or (siegoperidol) or (sigaperidol) or (trancodol-10) or (trancodol-5) [Words] and (delirium) or (delirious) OR (acute organic psychosyndrome) or (acute organic brain syndrome) OR (acute brain dysfunction) or (acute brain failure) or (acute brain syndrome) OR (postoperative cognitive dysfunction) or (post-operative cognitive dysfunction) or (postsurgical cognitive dysfunction) or (post-surgical cognitive dysfunction) OR (postoperative brain dysfunction) or (post-operative brain dysfunction) or (postsurgical brain dysfunction) or (post-surgical brain dysfunction) OR (postoperative psychosis) or (post-operative psychosis) or (postsurgical psychosis) or (post-surgical psychosis) OR (acute psycho-organic syndrome) or (acute psychoorganic syndrome) or (acute organic psychosyndrome) or (acute organic psycho-syndrome) OR (metabolic encephalopathy) or (exogenous psychosis) OR (acute confusion) or (acute psycho-organic syndrome) or (obnubilate) OR (clouded state) or (clouding of consciousness) OR (toxic psychosis) or (toxic confusion) [Words]

DATA COLLECTION FORM

	General					Interventions						Patient information - Sub. Gr.		
	Trial id	Year	Publ. Type	Protocol	Sub.Gr.1 OVERALL RoB	Exp (Haloperidol) IV vs oral	Exp (Haloperidol) Dose	Control (any) IV vs oral	Control (any) Dose	Intervention period	Max follow-up	Sub.Gr.2 Population	Sub.Gr.3 Used control intervention	Sub.Gr.4 Delirium diagnosis
1	Atalan	2013												
2	Bakri (Dex)	2015												
	Bakri (Ondan)	2015												
3	Breitbart (chlor)	1996												
	Breitbart (lorazepam)	1996												
4	Girard (placebo)	2018												
	Girard (ziprazidone)	2018												
5	Han	2004												
6	ORIC-I	2017												
7	Skrobik	2004												
8	Tagarakis	2012												

Randomisation and follow-up (best worst/worst best)									All-cause mortality									
E: No randomised	C: No randomised	Total randomised	*E: Lost to follow-up*	*C: Lost to follow-up*	E: No analysed	C: No analysed	Total analysed	**NOTES**	E: No of events	E: No analysed	C: No of events	C: No analysed	Blind pt+pers	Blind outcome assessor	Incomepl outcome data	Selective outcome reporting	Time point used	NOTES

Days alive without delirium within 28 days												Quality of life												
E: Number of days	E: IQR	E: Total analysed	C: Number of days	C: IQR	C: Total analysed	Blind pt+pers	Blind outcome assessor	Incomepl outcome data	Selective outcome reporting	Time point	Notes	E: Mean	E: SD	E: Total analysed	C: Mean	C: SD	C: Total analysed	Blind pt+pers	Blind outcome assessor	Incomepl outcome data	Selective outcome reporting	Time point	Scale used	Mean or change score	NOTES

Cognitive function														Delirium severity													
E: Mean	E: SD	E: Total analysed	C: Mean	C: SD	C: Total analysed	Blind pt+pers	Blind outcome assessor	Incomepl outcome data	Selective outcome reporting	Time point	Scale used	End or change score	NOTES	E: Mean	E: SD	E: Total analysed	C: Mean	C: SD	C: Total analysed	Blind pt+pers	Blind outcome assessor	Incomepl outcome data	Selective outcome reporting	Time point	Scale used	Mean or change score	NOTES

Prop. of pts with one or more SAR - composite outcome										Prolongation of QTc interval (exploratory)									
E: No of events	E: No analysed	C: No of events	C: No analysed	Blind pt+pers	Blind outcome assessor	Incomepl outcome data	Selective outcome reporting	Time point	NOTES	E: No of events	E: No analysed	C: No of events	C: No analysed	Blind pt+pers	Blind outcome assessor	Incomepl outcome data	Selective outcome reporting	Time point	NOTES

Extrapyramidal symptoms (post-hoc)										Delirium resolution (post hoc)									
E: No of events	E: No analysed	C: No of events	C: No analysed	Blind pt+pers	Blind outcome assessor	Incomepl outcome data	Selective outcome reporting	Time point	NOTES	E: No of patients	E: No analysed	C: No of patients	C: No analysed	Blind pt+pers	Blind outcome assessor	Incomepl outcome data	Selective outcome reporting	Time point	NOTES

DETAILS OF INCLUDED TRIALS AND RISK OF BIAS ASSESSMENT

ATALAN 2013 [1]	
Methods	Randomised clinical trial The trial was conducted in a community hospital in Turkey
Participants	**Sample size:** 53 (haloperidol 26; morphine 27) **Age (mean):** 66 years **Sex:** 74% males **Baseline disease severity:** APACHE II score (preoperative) 6,01 **Co-morbidities:** preoperative characteristics of the study population: • cigarette use: haloperidol 46,2%; morphine 33,3% • alcohol use: haloperidol 3,8%; morphine 18,5% • COPD: haloperidol 7,7%; morphine 7,4% • hypertension: haloperidol 61,5%; morphine 51,9% • noninsulin-dependent diabetes mellitus: haloperidol 34,6%; morphine 25,9% • insulin-dependent diabetes mellitus: haloperidol 15,4%; morphine 14,8% • previous stroke: haloperidol 7,7%; morphine 3,7% • psychotropic drugs: haloperidol 11,5%; morphine 3,7% • BMI: haloperidol 28,3; morphine 27,6 **Setting:** Patients with hyperactive delirium after cardiac surgery admitted to ICU. **Inclusion criteria:** patients with delirium who had cardiac surgery with or without cardiopulmonary bypass. **Exclusion criteria:** Patients who had a history of dementia and/or abnormal level of consciousness, Parkinson's disease, and recent seizures prior to surgery. **Delirium:** Delirium was defined according to the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders IV criteria. Postoperatively, patients were screened daily with CAM-ICU. For delirium to be diagnosed, both the first and second criteria (1. acute onset and fluctuating course; 2. inattention; 3. disorganized thinking; 4. altered level of consciousness) had to be present, with either criterion 3 or 4. Patients who were diagnosed with delirium were evaluated further by the Richmond Agitation and Sedation Scale (RASS). Based on this scale, patients who had a RASS score of < 0 were diagnosed with hypoactive delirium and who had a RASS score of > + 2 were diagnosed with hyperactive delirium. Patients with hypoactive delirium were excluded from the study; whereas patients with hyperactive delirium were randomised intro two groups. All delirious patients were re-evaluated every 12 hours by CAM-ICU and RASS until discharged from the hospital or for a maximum of 10 days following surgery. Patients were considered delirium free when they were free of symptoms for more than 24 hours.
Interventions	**Experimental intervention:** 5 mg haloperidol IM every hour until the adequate sedation and target RASS scores (between -1 and +1) were achieved. **Control intervention:** 5 mg morphine sulphate IM every hour until the adequate sedation and target RASS scores (between -1 and +1) were achieved. **Timing:** study medications were started after diagnosis of delirium **Duration:** Maximum 10 days.

	Co-intervention: In patients who were still agitated despite the administration of 20 mg/d of morphine or 20 mg/d of haloperidol, 2.5 mg of lorazepam perorally, twice a day was added to the treatment regimen. During admission to the ICU, every patient was ventilated in assist-control mode to maintain pH between 7.35 and 7.45, PaCO2 between 35 and 45 mmHg, and PaO2 > 95%. Ventilation was weaned as per ICU protocol. Postoperative analgesia was achieved by providing 1 g of paracetamol intravenously every 8 hours and 50 mg of dexketoprofen intravenously twice a day.
Outcomes	**Outcomes:** • Duration time of delirious behaviour • Daily total medication doses • Need for additional sedative drug • The percentage of patients who maintained a RASS score within the target scores • Reintubation • Redo-surgery • Length of ICU and hospital stay • Readmission to the ICU • Hospital mortality rate **Timing of outcome measurement:** until patients were discharged from the hospital or for a maximum of 10 days following surgery.
Notes	28 May 2019: E-mail sent to Dr Nazan Atan asking for additional information about the trial. Reminder sent 4 June 2019. Reply was not received.

Risk of bias assessment

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	Unclear risk	Not described
Blinding of outcome assessment (detection bias)	Low risk	The clinical evaluation was made by an intensivist together with a consultant psychiatrist, who was blinded to the study groups
Incomplete outcome data (attrition bias)	Low risk	No patient was lost to follow-up
Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	Unclear risk	It was unclear how the trial was funded

BAKRI 2015 [2]	
Methods	Randomised clinical trial The trial was conducted in Saudi Arabia
Participants	**Sample size:** 96 (haloperidol 32; dexmedetomidine 32; ondansetron 32) **Age (mean):** 31 years **Sex:** 91% males **Baseline disease severity:** mean injury severity score (ISS) 24,5 **Co-morbidities:** characteristics of the study population: • patients on mechanical ventilation on ICU admission: haloperidol 31,3%; dexmedetomidine 28,1%; ondansetron 21,8% • weight (kg): haloperidol 71; dexmedetomidine 74; ondansetron 72

	Setting: postoperative trauma patients admitted to the ICU **Inclusion criteria:** adult postoperative trauma patients admitted to the ICU who screened delirium-positive, by using Intensive Care Delirium Screening Checklist (ICDSC). **Exclusion criteria:** Patients were excluded if they had underlying neurological diseases, significant hearing loss, intracranial injury, ischemic/hemorrhagic strokes, or language barrier that would confound the evaluation of delirium. Similarly, severely injured, deeply comatose, or moribund patients were excluded. **Delirium:** The delirium was assessed by Intensive Care Delirium Screening Checklist (ICDSC). The ICDSC includes eight items, based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) criteria and features of delirium, which includes altered level of consciousness, inattention, disorientation, hallucination-delusion-psychosis, psychomotor agitation or retardation, inappropriate speech or mood, sleep–wake cycle disturbances, and symptom fluctuation according to a total score system from 0 to 8 points. Delirium-positive was defined if the patient had a score of 4 points or more on the ICDSC scale. Delirium-positive patients were assessed twice a day for 3 days after inclusion in the study. The ICDSC was assessed 1 h after study medications were given. Averages of the two scores were recorded every day.
Interventions	**Experimental intervention:** 5 mg haloperidol IV twice daily by infusion **Control intervention:** 1 µg/kg dexmedetomidine twice daily (infusion) or 4 mg ondansetron twice daily (infusion) **Timing:** study medications were started after diagnosis of delirium **Duration:** 3 days **Co-intervention:** The treating physicians were free to prescribe additional haloperidol as rescue when clinically needed in all the three groups. Rescue haloperidol was used in 9.4% of the patients in the haloperidol group, 15.6% of the patients in the dexmedetomidine groups and in 34.4% of the patients in the ondansetron group.
Outcomes	**Outcomes:** • number of patients with delirium at day 3 • number of patients requireing rescue haloperidol • ICDSC scores • mean arterial blood pressure • mean Visual Analog Scale (VAS) of pain at the time of delirium assessment • serious adverse events • prolongation of QTc interval **Timing of outcome measurement:** after three days of intervention
Notes	The study was a three-arm study. We have split the study into two studies: haloperidol versus dexmedetomidine and haloperidol versus ondansetron, thus, we have divided patients and events from the haloperidol group in two. 28 May 2019: E-mail sent to Dr Bakri asking for additional information on the trial. Reminder sent 4 June 2019. Reply was not received. We have calculated SDs for delirium severity from the reported numbers reported (1.1; 1.2 and 1.3), which we assume are SEMs.

Risk of bias assessment

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated random numbers
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	Low risk	The study medications were calculated and prepared by physicians who were not a part of the research team. Data were collected by researchers who were blinded to the

		study drugs. Patients were managed by the ICU staffs who were not included in the study
Blinding of outcome assessment (detection bias)	Low risk	Data were collected by researchers who were blinded to the study drugs
Incomplete outcome data (attrition bias)	Low risk	No patient was lost to follow-up
Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	Low risk	The study was carried out without external funding and the study appeared to be free of other components that could put it at risk of bias
BREITBART 1996 [3]		
Methods	Randomised clinical trial The trial was conducted in the US	
Participants	**Age (mean):** 39,2 years **Sex:** 77% males **Baseline disease severity:** Karnofsky Performance Status for all patients was 52,3. Medical status Profile showed that patients had multiple medical complications (mean 12,57). Most common were hematologic and metabolic disorders (e.g. anemia, leukopenia, thrombocytopenia, hypoalbuminemia) and infectious diseases (e.g. septicemia, systemic fungal infections, pneumocystis carinii pneumonia, tuberculosis, and disseminated viral infections). Severity of medical complications was moderate to severe. **Co-morbidities:** not reported **Setting:** AIDS patients admitted to a high dependency AIDS unit **Inclusion criteria:** medically hospitalised adult patients with delirium who met the case definition for AIDS and were undergoing treatment for AIDS-related medical problems **Exclusion criteria:** known hypersensitivity to neuroleptics or benzodiazepines; presence of neuroleptic malignant syndrome; concurrent treatment with neuroleptic drugs; seizure disorder; current systematic chemotherapy for Kaposi's sarcoma; withdrawal syndrome or anticholinergic delirium for which a more specific treatment was indicated; current or past diagnosis of schizophrenia, schizoaffective disorder, or bipolar disorder; and the study compomising medical treatment for the underlying aetiology. Efforts were made to also exclude patients in whom delirium appeared to be part of a terminal event (i.e., the patient was expected to die within 24 hours). **Delirium:** enrolled patients were treated to the study protocol if they met DSM-III-R criteria for delirium and scored above the threshold score diagnostic for delirium (score of 13 or greater) on the delirium rating scale. The delirium rating scale is a 10 item scale specifically integrating DSM-III criteria. The maximum possible score is 32.	
Interventions	**Experimental intervention:** a treatment protocol for study drug administration was followed. Each delirious patient was evaluated hourly with the Delirium Rating Scale - if the patient's score was 13 or greater, the next level dose of study drug was administered. After stabilisation (when the patient was asleep, calm, and not hallucinating or had scored 12 or below on the Delirium Rating Scale) a maintenance dose was started on day 2 and continued for up to 6 days of treatment protocol. Mean haloperidol dose the first 24 hours was 2,8 mg. Average maintenance dose was 1,4 mg. **Control intervention:** mean chlorpromazine dose the first 24 hours was 50 mg. Average maintenance dose was 36 mg. Mean lorazepam dose the first 24 hours was 3 mg. Average maintenance dose was 4,6 mg. **Timing:** study medications were started after diagnosis of delirium (treatment was initiated during the first 24 to 48 hours of delirium onset) **Duration:** treatment protocol up to 6 days **Co-intervention:** not reported	
Outcomes	**Outcomes:** • mortality	

	• scores on Delirium Rating Scale • scores on Mini-Mental State (cognitive function) **Timing of outcome measurement:** day 2 and end of treatment (delirium and cognitive scores) and 8 days from initiation of protocol (mortality)
Notes	The study was a three-arm study. We have split the study into two studies: haloperidol versus chlorpromazine and haloperidol versus lorazepam, thus, we have divided patients and events from the haloperidol group in two. Midway through the study, lorazepam was removed from the study due to treatment limiting adverse side effects in this group. 28 May 2019: E-mail sent to Dr Breitbart asking for additional information on risk of bias and outcomes. Reply was received. Additional data on delirium resolution was not received.

Risk of bias assessment

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	A biostatistician generated a randomization table for groups of 9 patients each assuring that an equal number of batches of 9 patients would receive equal numbers of the 3 study drugs - labelled Drugs A, B, or C (confirmed by Dr Breitbart via email correspondence)
Allocation concealment (selection bias)	Low risk	The hospital pharmacy kept all study drugs and when a patient on the study protocol was ordered study drug the pharmacist would consult the randomization table and dispense study drug A, B, or C, whichever was indicated as the next drug to dispense (confirmed by Dr Breitbart via email correspondence)
Blinding of participants and personnel (performance bias)	Low risk	Drugs A, B, and C were all dispensed in identical capsules. All 3 drugs were available and possible to administer in intravenous form if needed and done so with no identifying information, both clinicians and participants were blinded (confirmed by Dr Breitbart via email correspondence)
Blinding of outcome assessment (detection bias)	Low risk	Outcome assessors were blinded to the drug. They only knew if patient received drug A B or C (confirmed by Dr Breitbart via email correspondence)
Incomplete outcome data (attrition bias)	Low risk	No patient was lost to follow-up
Selective reporting (reporting bias)	High risk	No protocol (confirmed by Dr Breitbart via email correspondence)
Other bias	Low risk	The trial was a US National Institute of Mental Health funded RO1 investigator initiated 5-year funded grant

GIRARD 2018 [4]

Methods	Randomised clinical trial The trial was conducted in the US
Participants	**Age (median):** 60 years **Sex:** 57% males **Baseline disease severity: m**edian APACHE II score at ICU admission: 28.8; Median SOFA score at randomisation: 11; Informant Questionnaire on Cognitive Decline in the Elderly (Short IQCODE): 3.1 **Co-morbidities:** Median Charlson Comorbidity Index score: 2 Received assisted ventilation before randomisation: invasive 93%; non-invasive 3% Shock before randomisation: 33% Diagnosis at admission: • Adult respiratory distress syndrome: placebo 21%; haloperidol 23%; ziprazidone 18% • Sepsis: placebo 19%; haloperidol 22%; ziprazidone 17% • Airway protection: placebo 29%; haloperidol 24%; ziprazidone 23% • Chronic obstructive pulmonary disease, asthma, or other pulmonary disorder: placebo 12%; haloperidol 10%; ziprazidone 15%

	• Surgery: placebo 7%; haloperidol 7%; ziprazidone 12% • Chronic heart failure, myocardial infarction, or arrhythmia: placebo 3%; haloperidol 3%; ziprazidone 3% • Cirrhosis or liver failure: placebo 3%; haloperidol 2%; ziprazidone 2% • Seizures or neurologic disease: placebo 1%; haloperidol 2; ziprazidone 1% **Setting:** patients admitted to the ICU (28% admitted to surgical ICU). **Inclusion criteria:** adults (≥18 years old) with delirium admitted to medical and/or surgical ICUs in participating hospitals who were treated with mechanical ventilation, non-invasive positive pressure ventilation, vasopressor(s), or intraaortic balloon pump. **Exclusion criteria:** patients who, at baseline, had severe cognitive impairment; were at high risk for medication side effects because of pregnancy, breast-feeding, a history of torsades de pointes, QT prolongation, a history of neuroleptic malignant syndrome, or allergy to haloperidol or ziprasidone; were receiving ongoing treatment with an antipsychotic medication; were in a moribund state; had rapidly resolving organ failure; were blind, deaf, or unable to speak or understand English; were incarcerated; or were enrolled in another study or trial that prohibited co-enrolment. **Delirium:** Delirium was detected with the use of the Confusion Assessment Method for the ICU (CAM-ICU) that identifies delirium on the basis of an acute change or fluctuating course of mental status plus inattention and either altered level of consciousness or disorganized thinking. If delirium was not present at the time that informed consent was obtained, trained research personnel evaluated patients twice daily until delirium was present or until death, discharge from the ICU, development of an exclusion criterion, or a maximum of 5 days. Delirium characteristics at randomisation: 11% hyperactive and 89% hypoactive.
Interventions	**Experimental intervention:** patients younger than 70 years of age received 2.5 mg haloperidol per 0.5 ml and 1.25 mg of haloperidol per 0.25 ml when older than 70 years of age. Patients in the haloperidol group received a dose of up to 10 mg per administration and up to 20 mg per day. Mean (±SD) daily doses of haloperidol administered were 11.0±4.8 mg. **Control intervention:** patients younger than 70 years of age received 5 mg of ziprasidone per 0.5 ml or 2.5 mg of ziprasidone per 0.25 ml when older than 70 years of age. Patients in the ziprasidone group received a dose of up to 20 mg per administration and up to 40 mg per day. Mean (±SD) daily doses of ziprasidone administered were 20.0±9.4 mg. Patients younger than 70 years of age received 0.5 ml placebo (0.9% saline) and 0.25 ml of placebo when older than 70 years of age. Volume and dose of a trial drug or placebo were halved if a patient did not have delirium (i.e., had a negative CAM-ICU assessment) for two consecutive assessments and was not yet receiving the minimum dose. Trial drug or placebo were temporarily withheld if a patient did not have delirium for four consecutive assessments or for safety reasons. Trial drug or placebo were permanently discontinued when any of the following occurred: torsades de pointes, neuroleptic malignant syndrome, drug reaction with eosinophilia and systemic symptoms syndrome, new onset coma due to structural brain disease, or any life-threatening, serious adverse event that was related to the intervention, as determined by an independent data and safety monitoring board. **Timing:** patients were randomised when delirium was present at the time of informed consent or during the 5 days after informed consent was obtained and the corrected QT interval was less than 550 msec on a 12-lead electrocardiogram. Immediately after the trial-group assignment, the first dose of trial drug or placebo was administered. **Duration:** 14 days or at ICU discharge. The median duration of exposure to a trial drug or placebo was 4 days (interquartile range, 3 to 7). **Co-intervention:** the administration of open-label haloperidol, ziprasidone, or any other antipsychotic (except those prescribed specifically for nausea, such as compazine) were restricted during the 14-day study drug period. Open-label haloperidol was used in 15% of the patients in the placebo group and in 12% of the patients in the ziprasidone group. Approximately 90% of the patients received analgesics or sedatives. Daily rate of adherence to each of the five components of the ABCDE bundle was greater than 88%.

Outcomes	**Outcomes:** • days alive without delirium or coma (defined as the number of days that a patient was alive and free from both delirium and coma during the 14-day intervention period) • duration of delirium • time to freedom from mechanical ventilation (defined as extubation that was followed by at least a 48-hour period during which the patient was alive and free from mechanical ventilation) • time to final successful ICU discharge (defined as the last ICU discharge during the index hospitalization that was followed by at least a 48-hour period during which the patient was alive and outside the ICU) • time to ICU readmission • time to successful hospital discharge (defined as discharge that was followed by at least a 48-hour period during which the patient was alive and outside the hospital) • 30-day and 90-day survival • incidence of torsades de pointes • incidence of neuroleptic malignant syndrome • severity of extrapyramidal symptoms as measured on the modified Simpson–Angus Scale **Timing of outcome measurement:** outcome dependant
Notes	The study was a three-arm study. We have split the study into two studies: haloperidol versus placebo and haloperidol versus ziprazidone, thus, we have divided patients and events from the haloperidol group in two. 29 May 2019: E-mail sent to Dr Ely and Dr Girard asking for additional information on the outcomes. Reply was received. Link to complete dataset was received and number of events of prolonged QTc was extracted. Cognitive function was only measured as pre-existing cognitive impairment.

Risk of bias assessment

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated random numbers
Allocation concealment (selection bias)	Low risk	From protocol: The randomization scheme will be created by the study's primary biostatistician and will be distributed directly to the investigational pharmacy at each study site as a set of sealed, sequentially numbered, opaque envelopes containing tri-folded randomization assignments. Once a consented patient has become delirious and an order for blinded study drug is placed, the investigational pharmacist will open the next available envelope to establish that patient's treatment assignment. Treatment assignment will be known only by the investigational pharmacists.
Blinding of participants and personnel (performance bias)	Low risk	Colourless preparations delivered in identical bags was used
Blinding of outcome assessment (detection bias)	Low risk	The research personnel and managing clinicians
Incomplete outcome data (attrition bias)	Low risk	Less than 5% were lost to follow-up
Selective reporting (reporting bias)	Low risk	The protocol was pre-published. All outcomes were reported on except for long-term neuropsychological function, functional independence, quality of life, and posttraumatic stress disorder symptoms at 3-month and 1-year follow-up are not reported on; however, we believe these data will be published in a secondary publication.
Other bias	Low risk	The trial was supported by public grants and appeared free of other components that could put it at risk of bias

HAN 2004 [5]

Methods	Randomised clinical trial The trial was conducted in Korea
Participants	**Age (mean):** 66 years **Sex:** 54% males **Baseline disease severity:** mean Memorial Delirium Assessment Scale score 24.5

	Co-morbidities: Medical diagnoses at admission: • fractures: haloperidol 25%; risperidone 34% • cerebrovascular accident: haloperidol 25%; risperidone 17% • peritonitis: haloperidol 8%; risperidone 8% • chronic renal failure: haloperidol 8%; risperidone 17% • cancer: haloperidol 8%; risperidone 8% • cardiovascular disease: haloperidol 18%; risperidone 8% • other: haloperidol 8%; risperidone 8% **Setting:** patients admitted to the ICU (1 patient in each group was admitted to an oncology ward). **Inclusion criteria:** patients with delirium admitted to an ICU or oncology ward. **Exclusion criteria:** patients with any type of dementia or other psychiatric diagnosis, patients who had already been injected with antipsychotics or benzodiazepines in the emergency room or ICU for their disturbing behavioural problems before the arrival of the consulting psychiatrist. **Delirium:** Screening and detection of delirium were conducted with the Confusion Assessment Method (cutoff 13) and Delirium Rating Scale. Diagnosis of delirium was determined with the Structured Clinical Interview for DSM-III-R (SCID) according to DSM-III-R criteria. The Memorial Delirium Assessment Scale was used to measure delirium severity and a cut-off at 13 was used.
Interventions	**Experimental intervention:** The initial starting dose haloperidol 0.75 mg twice a day. The dosage was increased depending on the status of delirium during the 7 days. **Control intervention:** The initial starting dose of risperidone was 0.5 mg twice a day. The dosage was increased depending on the status of delirium during the 7 days. **Timing:** study medication was started after diagnosis of delirium. **Duration:** 7 days
Outcomes	**Outcomes:** • mean delirium Rating Scale scores • delirium severity assessed with Memorial Delirium Assessment Scale • delirium resolution • duration of delirium **Timing of outcome measurement:** 7 days
Notes	The study included patients admitted to four medical wards, two ICUs and two oncology wards. As more than 90% of the included patients came from ICU, we have included this study in our review. Memorial Delirium Assessment Scale scores (mean and standard deviation) were extracted from figure 1. 29 May 2019: E-mail sent to Dr Kim asking for additional information on risk of bias and outcomes. Reply was received. No additional data or clarifications were received due to loss of data.

Risk of bias assessment

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Stated that the patients were randomly assigned to the interventions, but the method of sequence generation was not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	High risk	The trial is described as double-blinded. However, tablets were not identical, thus the trial was not blinded

Blinding of outcome assessment (detection bias)	Low risk	Psychiatrist who assessed delirium symptoms had no information about to which group the patients were allocated
Incomplete outcome data (attrition bias)	High risk	Four patients (14%) dropped out (two in each group) and results did not include intention to treat data
Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	Low risk	The study was supported by a public fund
ORIC-I [6]		
Methods	Randomised clinical trial The trial was conducted in the US	
Participants	**Age (mean)**: not specified **Sex:** 63% males **Baseline disease severity:** not reported **Co-morbidities**: not reported **Setting**: mechanically ventilated patients with delirium **Inclusion criteria:** all adult (>=18 years of age) mechanically ventilated patients admitted to the medical, surgical, trauma or cardiothoracic ICUs who are expected by the ICU clinical team to require > 24 hours of mechanical ventilation **Exclusion criteria**: Baseline QTc >480 milliseconds (ms); history of Parkinsons's disease; pregnancy; history of schizophrenia or neurologic disease that could confound the delirium assessment; deafness or inability to understand English or Spanish; extubation prior to enrolment; previously enrolled in the same study; patient, family, or attending physician refusal; death before enrolment; treatment with haloperidol within 2 days prior to ICU admission; and prisoners. **Delirium:** details on delirium assessment were not specified	
Interventions	**Experimental intervention:** haloperidol 5 mg IV every 12 hours **Control intervention**: 5 mg saline placebo **Timing**: study medications were started after diagnosis of delirium **Duration**: intervention was continued until liberation from mechanical ventilation or 28 days, whichever was first **Co-intervention**: not reported	
Outcomes	**Pre-planned outcomes:** • 28-day all-cause mortality • 90-day all-cause mortality • duration of mechanical ventilation • ICU length of stay • total delirium days **Reported outcomes:** • 30-day all-cause mortality • Prolongation of QTc interval	

Notes	The study was identified when searching clinical.trials.gov. The study was terminated early due to insufficient recruitment to meet the aims. The study has not been published, but some results have been posted at clinical.trials.gov. 3 June 2019: E-mail sent to Barbara Early asking for additional information on risk of bias and outcomes. Reply was received. It was clarified that mortality was measured at day 30 and that the 4 measurements of QTc prolongation corresponds to 4 individual patients. Rescue drug was not used. No additional results were available and no clarifications for the risk of bias assessment was received.

Risk of bias assessment

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	Unclear risk	The trial was described as blinded (participants, care provider, investigator) and placebo was used in the control group. However, method of blinding was not adequately described.
Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	Low risk	No patient was lost to follow-up
Selective reporting (reporting bias)	High risk	Only results on mortality, serious adverse events and adverse events are reported
Other bias	High risk	It was unclear how the trial was funded. The study was pre-maturely terminated and only limited data are reported due to either missing data or conflicting data

SKROBIK 2004 [7]

Methods	Randomised clinical trial The trial was conducted in Canada
Participants	**Age (mean):** 65 years **Sex:** 73% males **Baseline disease severity:** mean APACHE II score 12.89 **Co-morbidities:** not reported Type of admission: • surgical elective: haloperidol 58%; olanzapine 79% • surgical urgent: haloperidol 38%; olanzapine 14% • medical: haloperidol 4%; olanzapine 7% **Setting:** patients admitted to a medical-surgical ICU. **Inclusion criteria:** patients aged 18–75 years with delirium admitted to the ICU for more than 24 h. **Exclusion criteria:** pregnant patients, those who received antipsychotic medication within 10 days prior to hospital or ICU admission, or in whom either haloperidol or olanzapine was contraindicated. Contraindications to drug administration were Parkinson's disease, oropharyngeal dysfunction, prolonged QT interval, and hepatic or renal dysfunction. Individuals with gastrointestinal dysfunction, precluding oral/enteral drug administration, or whose neurological status did not permit an adequate neuropsychiatric evaluation (e.g., stupor or coma), were also excluded. **Delirium:** patients were screened three times daily for delirium utilizing the ICU Delirium Screening Checklist, ICU-DSC. In screened patients with an ICU DSC of 4 or with clinical manifestations delirium, the diagnosis was confirmed by a physician using DSMIV criteria.

Interventions	**Experimental intervention:** haloperidol was initiated at 2.5–5 mg every 8 h. Patients over 60 years received a lower initial dosage (haloperidol 0.5–1 mg). Patients in the haloperidol group received a mean daily dose of 6.5 mg (range 1–28 mg). **Control intervention:** olanzapine was initiated at 5 mg daily. Patients over 60 years received a lower initial dosage (olanzapine 2.5 mg). Patients in the olanzapine group received a mean daily dose of 4.54 mg (range 2.5–13.5 mg). **Timing:** within two hours of the diagnosis of delirium. **Duration:** 5 days. **Co-intervention:** patients who developed agitation during the study were permitted intravenous haloperidol administration (recorded as "rescue haloperidol"). Rescue haloperidol was used in 42.2% of the patients in the haloperidol group and in 35.7% of the patients in the olanzapine group.
Outcomes	**Outcomes:** • delirium severity • use of rescue haloperidol • extrapyramidal symptoms **Timing of outcome measurement:** end of study intervention (day 5)
Notes	29 May 2019: E-mail sent to Dr Skrobik asking for additional information on risk of bias and outcomes. Reply was received. No additional info received. We extracted end scores from figure 1. SDs were not reported. We used SDs from the trial of Bakri 2015 as this trial also used ICDSC.

Risk of bias assessment

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Odd/even day randomisation was used
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	High risk	Treating physician and nurses were not blinded
Blinding of outcome assessment (detection bias)	Low risk	Objective evaluations were performed on a daily basis by a clinician or research nurse blinded to the dispensed medication.
Incomplete outcome data (attrition bias)	High risk	10% of the enrolled patients were lost to follow-up (7 patients whose allocation group was not stated) and intention to treat data were not reported
Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	High risk	The study was funded by the industry (Zyprexa fund, Eli-Lilly)

TAGARAKIS 2012 [8]

Methods	Randomised clinical trial The trial was conducted in Greece
Participants	**Age:** 71 years **Sex:** 66% males **Baseline disease severity:** not stated **Co-morbidities:** not stated **Setting:** patients after on-pump cardiac surgery

	Inclusion criteria: patients with delirium after on-pump cardiac surgery (coronary artery bypass graft surgery, aortic valve replacement surgery, mitral valve surgery or combined procedures) **Exclusion criteria:** history of severe psychiatric disease **Delirium:** For the detection and evaluation of postoperative delirium, a scale developed by Bayindir et al. was applied. This 4-point scale was rated as follows: 0: normal, 1: patient with restlessness and mild confusion but cooperative, 2: patient disorientated but cooperative, memory gaps, 3: patient disorientated and uncooperative with augmented mobility that could put him to danger, 4: patient totally disorientated, violent and aggressive, presence of hallucinations. Patients were evaluated before and 10 minutes after administration of study drug.
Interventions	**Experimental intervention:** 5 mg haloperidol iv **Control intervention:** 8 mg ondansetron iv **Timing:** study medication was started after diagnosis of delirium **Duration:** unclear **Co-intervention:** not reported
Outcomes	**Outcomes:** • Delirium severity • Delirium resolution **Timing of outcome measurement:** 10 minutes post study drug administration
Notes	29 May 2019: E-mail sent to Dr Tagarakis asking for additional information on risk of bias and outcomes. Reminder sent 5 June 2019. Reply was not received. We have calculated SDs for delirium severity from the reported numbers reported (0.1), which we assume is SEM.

Risk of bias assessment

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	The two substances were administered at a random, alternate (one but one) order
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	Unclear risk	The trial was described as double-blind, but it was unclear how the blinding was performed
Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	Low risk	No patient was lost to follow-up
Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	Unclear risk	It was unclear how the trial was funded

SELECTED EXCLUDED TRIALS

Wrong population

1. Adams F, Fernandez F, Andersson BS, Emergency pharmacotherapy of delirium in the critically ill cancer patient, Psychosomatics 1986;27(1 Suppl):33-8
2. Davies MP, Effect of Lorazepam with Haloperidol vs. Haloperidol Alone on Agitated Delirium in Patients, Journal of Pain and Symptom Management 2018;55(2):546-546
3. Goldstein BJ, Clyde DJ, Haloperidol in controlling symptoms of acute psychoses. A double-blind evaluation of haloperidol and trifluoperazine, Current Therapeutic Research-Clinical and Experimental 1966;8(5):236.
4. Hui D, Frisbee-Hume S, Wilson A, Dibaj SS, Nguyen TT, De La Cruz MG, et al, Lorazepam as an adjuvant to haloperidol for agitated delirium at the end-of-life: A double-blind randomized controlled trial, Supportive care in cancer 2017;25 (2 Supplement 1)():S192-S193
5. Jain R, Arun P, Sidana A, Sachdev A, Comparison of efficacy of haloperidol and olanzapine in the treatment of delirium, Indian journal of psychiatry 2017;59(4):451-456
6. Jin KJ, Kook LH, Un PC, Uk LC, Ho PI, Chul L, Comparison of intramuscular olanzapine and haloperidol for the treatment of delirium, European Psychiatry 2009;24:1
7. Jung HY, Lee SI, Kim SG, Hong JH, Double blind comparison on the efficacy of haloperidol versus quetiapine in the treatment of delirium, International Journal of Neuropsychopharmacology 2008;11():299-299
8. Jung W, Jung HY, Lee SI, Kim SG, Shin EY, Park J, Park M, Comparison on the efficacy of risperidone versus haloperidol in the treatment of delirium: prospective, randomized, double blind trial, 163rd annual meeting of the american psychiatric association 2010
9. Kim JJ, Lim HK, Pae CU, Lee CU, Lee C, Paik IH, Comparison of intramuscular olanzapine and haloperidol for the treatment of delirium, European neuropsychopharmacology 2008;18 (S4):S424-S425
10. Kim YR, Jung HY, Lee ISY, Kwon YJ, Cho SH, Hong JH, Double blind trial of amantadine and haloperidol for treatment of delirium, European neuropsychopharmacology 2006;16:S548-S548
11. Kumar V, Chakrabarti S, Grover S, Comparative efficacy of typical and atypical antipsychotic in delirium Indian journal of psychiatry 2010;1:S22
12. Lee CS, Rim GM, Hahn KH, Kim BI, Comparison of Efficacy between Aripiprazole and Haloperidol in the Treatment of Patients with Delirium, Korean journal of psychopharmacology 2007;18(4):240-245
13. Lee KJ, Kim H, Lee SH, Chung YC, Effects of risperidone in delirium: A comparison with haloperidol International Journal of Neuropsychopharmacology 2004;7:S384-S385
14. Lee SH, Park SJ, Comparison of efficacy and side effects between aripiprazole and haloperidol in the treatment of delirium, International psychogeriatrics 2013;1:S132-S133
15. Lee Y, Lee J, Rim HD, Kim SH, Chung US, Cho GA, Won SH, A comparative study of haloperidol and quetiapine in the treatment of delirium: a preliminary randomized open label, flexible dose trial, European neuropsychopharmacology 2006;16:S488-S489
16. Lee YI, Jung HY, Lee SI, Kim SG, Park JH, Comparison on the Efficacy of Quetiapine Versus Haloperidol in the Treatment of Delirium: prospective, Randomized Trial, Korean journal of biological psychiatry 2009;16(1):15-24
17. Lim HK, Paik IH; Oh K, Lee CU, Lee C, Comparison of the Clinical Efficacy and Safety between Intramuscular Olanzapine and Intramuscular Haloperidol Injection in the Treatment of Delirium, Korean journal of psychopharmacology 2007;18(6):423-428
18. Maneeton B, Maneeton N, Srisurapanont M, A double-blind, randomised, controlled trial of quetiapine versus haloperidol for the treatment of delirium: A preliminary report, European neuropsychopharmacology September 2011;3:S557
19. Maneeton B, Maneeton N, Srisurapanont M, Chittawatanarat K, Quetiapine versus haloperidol in the treatment of delirium: a double-blind, randomized, controlled trial, Drug design, development and therapy 2013;7():657-667
20. Pasechnik IN, Makhlai AV, Tepliakova AN, Gubaidullin RR, Sal'nikov PS, Borisov A, Berezenko MN, New approach to postoperative delirium treatment. [Russian], Khirurgiia 2015;(3):71-75
21. Peterson LG, Bongar B, Navane versus Haldol, Treatment of acute organic mental syndromes in the general hospital, General hospital psychiatry 1989;11(6):412-417
22. Ropert R, Payan C, Allard S, Sultopride versus haloperidol for the treatment of acute psychosis. Results of a multicenter double-blind controlled trial, Annales de psychiatrie 1989;4(1):92-98
23. Terhaar HW, Comparison of chlormethiazole and haloperidol in treatment of elderly patients with confusion or organic and psychogenic origin – a double-blind crossover study, Pharmatherapeutica 1977;1(9):563-569
24. Tune L, Jewart R, Egeli S, Greene Y, Pharmacologic management of acute delirium: A naturalistic, prospective comparison of atypical antipsychotics and haloperidol, Journal of the american geriatrics society 2001;49(4):S93-S93

25. Van Der Vorst M, Neefjes E, Boddaert M, Verdegaal B, Beeker A, Teunissen S, Beekman A, Wilschut, J, Berkhof J, Verheul H, Efficacy and side effect profile of olanzapine versus haloperidol for symptoms of delirium in hospitalized patients with advanced cancer: A multicenter, investigator-blinded, randomized, controlled trial (RCT), Palliative medicine 2018;32 (1 Supplement):8
26. Verachai V, Rukngan W, Chawanakrasaesin K, Nilaban S, Suwanmajo S, Thanateerabunjong R, Kaewkungwal J, Kalayasiri R, Treatment of methamphetamine-induced psychosis: a double-blind randomized controlled trial comparing haloperidol and quetiapine, Psychopharmacology 2014;231(16):3099-3108
27. Lee SH, Park SJ, Comparison of Efficacy and Side Effects between Aripiprazole and Haloperidol in the Treatment of Delirium, Journal of korean geriatric psychiatry 2011;15(1):31-37
28. Nakamura J, Uchimura N, Yamada S, Nakazawa Y, Does plasma free-3-methoxy-4-hydroxyphenyl(ethylene)glycol increase in the delirious state? A comparison of the effects of mianserin and haloperidol on delirium, International clinical psychopharmacology 1997;12(3):147-152
29. Duprey M, Al-Qadheeb N, Roberts R, Skrobik Y, Schumaker G, Devlin J, QTC interval prolongation with low-dose iv haloperidol: post hoc analysis of a placebo control trial, Critical care medicine 2016;44(12):290
30. Reade MC, O'Sullivan K, Bates S, Goldsmith D, Ainslie WRSTJ, Bellomo R, Dexmedetomidine vs. haloperidol in delirious, agitated, intubated patients: A randomised open-label trial, Critical Care 2009;13: R75
31. Hu H, Deng W, Yang H, et al, A prospective random control study: comparison of olanzapine and haloperidol in senile delirium, Chongqing Medical Journal 2004; 33(8): 234-37
32. Lliu XB, Tan QR, Wang ZZ, Olanzapine vs Haloperidol for Delirium: A Prospective Randomized Controlled Trial, South China Journal of National Defence Medicine 2008; 22(6): 1-3
33. Zou WS, Wang JF, Tang W, Randomly controlled study of olanzapine and haloperidol in the treatment of delirium in elderly, Chinese Journal of Geriatric Heart Brain and Vessel Disease 2010;12(3):245-47

Wrong indication

1. Eremenko AA; Chernova EV, Dexmedetomidine use for intravenous sedation and delirium treatment during early postoperative period in cardio-surgical patients, Anesteziologiia i reanimatologiia 2013;(5):4-8
2. Girard TD, Pandharipande PP, Carson SS, Schmidt GA, Wright PE, Canonico AE, Pun BT, Thompson JL, Shintani AK, Meltzer HY, et al, Feasibility, efficacy, and safety of antipsychotics for intensive care unit delirium: the MIND randomized, placebo-controlled trial, Critical care medicine 2010;38(2):428-437
3. Page VJ, Ely EW, Gates S, Zhao XB, Alce T, Shintani A, Jackson J, Perkins GD, McAuley DF, Effect of intravenous haloperidol on the duration of delirium and coma in critically ill patients (Hope-ICU): a randomised, double-blind, placebo-controlled trial, The Lancet Respiratory medicine 2013;1(7):515-523
4. Page VJ, Marti J, McAuley DF, Casarin A, Alce, T, Zhao XB, Health evaluation and cost-effectiveness analysis from a randomized trial of haloperidol in the management of delirium in the critically ill (Hope-ICU Trial), American journal of respiratory and critical care medicine 2015;191:A4022
5. Litvineneko IV, Odinak MM, Khlystov LUV, Perstnev SV, Fedorov BB, Efficacy and safety of rivastigmine (exelon) in the confusion syndrome in the acute phase of ischemic stroke, Zhurnal nevrologii i psikhiatrii imeni S.S. Korsakova 2010;110(11 Pt 2):36-41

Wrong intervention

1. Eremenko, Aa; Chernova, Ev, Treatment of delirium in the early postoperative period after cardiac surgery, Anesteziologiia i reanimatologiia 2014;(3):30-34
2. Graham BB, Overdier K, Douglas IS, A prospective randomized study of haloperidol in addition to standard sedation in delirious and intubated patients: preliminary safety analysis, American thoracic society international conference 2008 (a816)
3. Graham BB, Douglas IS, A prospective randomized study of haloperidol in addition to standard sedation in delirious and intubated patients: Preliminary safety analysis, Critical care medicine 2006;34(12):A160
4. Khan BA, Perkins AJ, Campbell NL, Gao S, Farber MO, Wang S, Khan SH, Zarzaur BL, Boustani MA, Pharmacological Management of Delirium in the Intensive Care Unit: A Randomized Pragmatic Clinical Trial, Journal of the american geriatrics society 2019; 67(5):1057–65.
5. van Eijk MM, Roes KC, Honing ML, Kuiper MA, Karakus A, van der Jagt M, Spronk PE, van Gool WA, van der Mast RC, Kesecioglu J, Slooter AJ, Effect of rivastigmine as an adjunct to usual care with haloperidol on duration of delirium and mortality in critically ill patients: a multicentre, double-blind, placebo-controlled randomised trial, Lancet 2010;376:1829-37

ONGOING AND TERMINATED TRIALS

Trial id	Trial name	Country	Status	Registration year	Target sample size	Experimental intervention	Control
IRCT20180911040998N1	Comparison the Effect of Quetiapine and Haloperidol on the Treatment of delirium in ICU	Iran	Recruiting	Oct/18	60	haloperidol	Quetiapine
IRCT20121231011956N10	Comparison the Effectiveness of Haloperidol and quetiapine for delirium in the Emergency department and intensive Care Unit	Iran	Recruiting	Oct/18	100	haloperidol	quetiapine
NCT03392376	Agents Intervening Against Delirium in Intensive Care Unit (AID-ICU)	Denmark	Recruiting	Jan/18	1000	haloperidol	placebo
NCT03628391	Efficacy of Haloperidol to Decrease the Burden of Delirium in Adult Critically Ill Patients (EuRIDICE): a Prospective Randomised Multi-center Double-blind Placebo-controlled Clinical Trial	Netherlands	Recruiting	Aug/18	742	haloperidol	placebo
NCT02343575	Valproic Acid for Treatment of Hyperactive or Mixed Delirium in ICU	US	Terminated	Jan/15	3 enrolled (terminated)	haloperidol	placebo vs valproic acid
NCT02345902	Treatment of Hypoactive Delirium and Outcome Measures (THDOM)	Mexico	Unknown, was recruiting	Jan/15	60	haloperidol	placebo vs non-pharm
NCT01811459	Trial Comparing Haloperidol, Quetiapine and Placebo in the Pharmacological Treatment of Delirium (Haloquet)	Canada	Completed	Mar/13	107 enrolled (completed)	haloperidol	Quetiapine vs placebo
NCT01140529	Dexmedetomidine for the Treatment of Delirium After Heart Surgery (DexinDelir)	Sweden	Terminated, slow recruitment	Jun/10	3 enrolled (terminated)	haloperodol	dexmedetomidine vs placebo
NCT00833300	Haloperidol vs Olanzapine for the Management of ICU Delirium	Canada	Terminated	Feb/09	200	haloperodol	olanzapine

NCT00599287	Methylphenidate, Rivastigmine or Haloperidol in Hypoactive Delirium in Intensive Care Patients	Netherlands	Terminated. Incl rate too low due to a lack of eligible patients and difficulties obtaining informed consent	Jan/08		80	haloperidol	methylphenidate va rivastigmine
ACTRN12606000085572	Dexmedetomidine and Haloperidol for the management of emergence delirium in intensive care (DeHedic)	Australia	Terminated early	Feb/06		60	haloperidol	dexmedetomidine
Trial NL495 (NTR537)	Delirium treatment at the surgical ward (DELTa S) Treatment of delirium: rivastigmine or haloperidol as primary treatment for delirium in elderly patients with a fractured hip. A randomized placebo-controlled study.		Terminated	Nov/05	target 100		haloperidol	rivastigmine

ADDITIONAL ANALYSES ON THE OUTCOMES

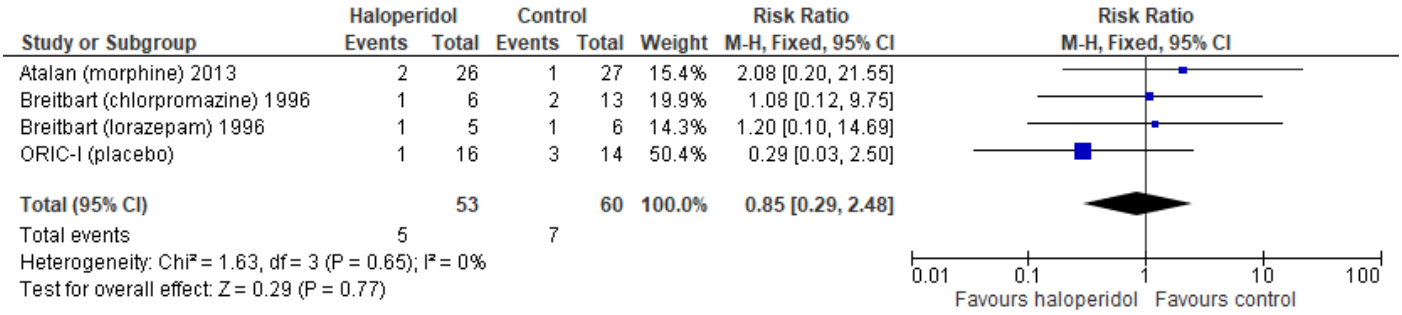
All-CAUSE MORTALITY, EXCLUDING TRIALS USING RESCUE HALOPERIDOL

Bayes Factors

Excluding trials using rescue haloperidol.
A Bayes factor of 1.05 was calculated based on an a priori RR of 1.20 and the meta-analysis result (RR 1.01) supporting that the result is likely 1.05 times more compatible with the null-hypothesis of a RRR = 0% than the alternative hypothesis of a RRR = 20%, for an effect of haloperidol versus control for treatment of delirium on all-cause mortality.

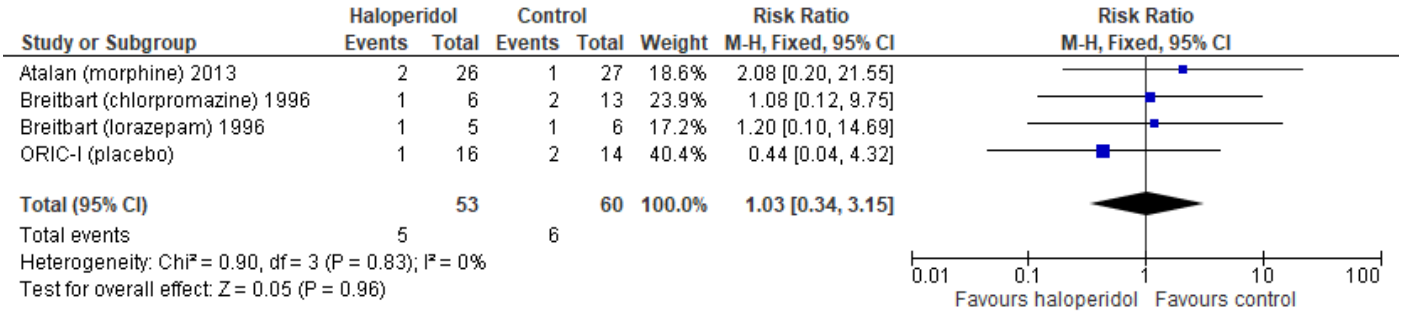
A Bayes factor of 1.08 was calculated based on an a priori RR of 0.80 and the meta-analysis result (RR 1.01) supporting that the result is likely 1.08 times more compatible with the null-hypothesis of a RRR = 0% than the alternative hypothesis of a RRR = 20%, for an effect of haloperidol versus control for the treatment of delirium on all-cause mortality.

Forest plot of best-worst case scenario sensitivity analysis for missing data on all-cause mortality



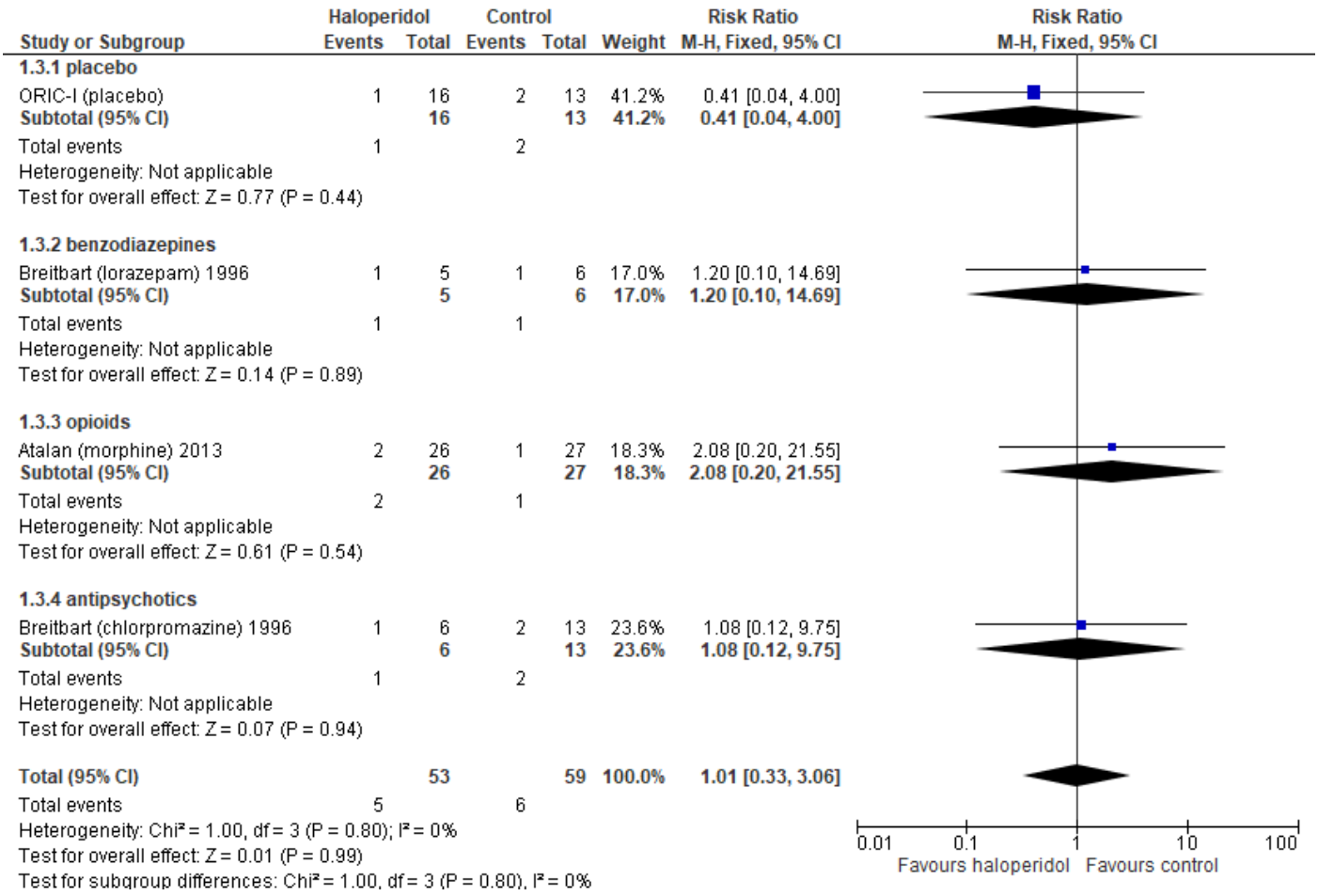
Excluding trials using rescue haloperidol

Forest plot of worst-best case scenario sensitivity analysis for missing data on all-cause mortality



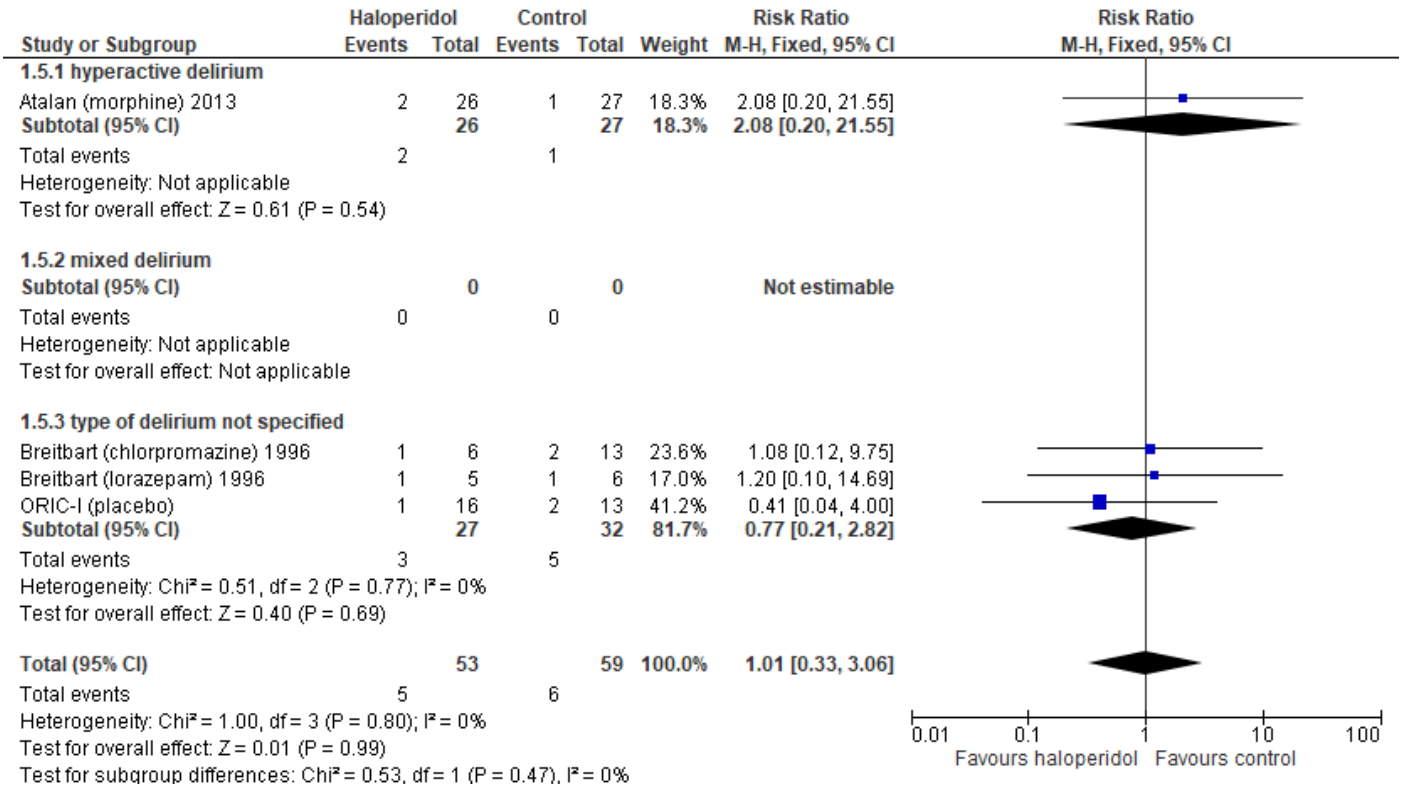
Excluding trials using rescue haloperidol

Forest plot of all-cause mortality stratified by used control intervention



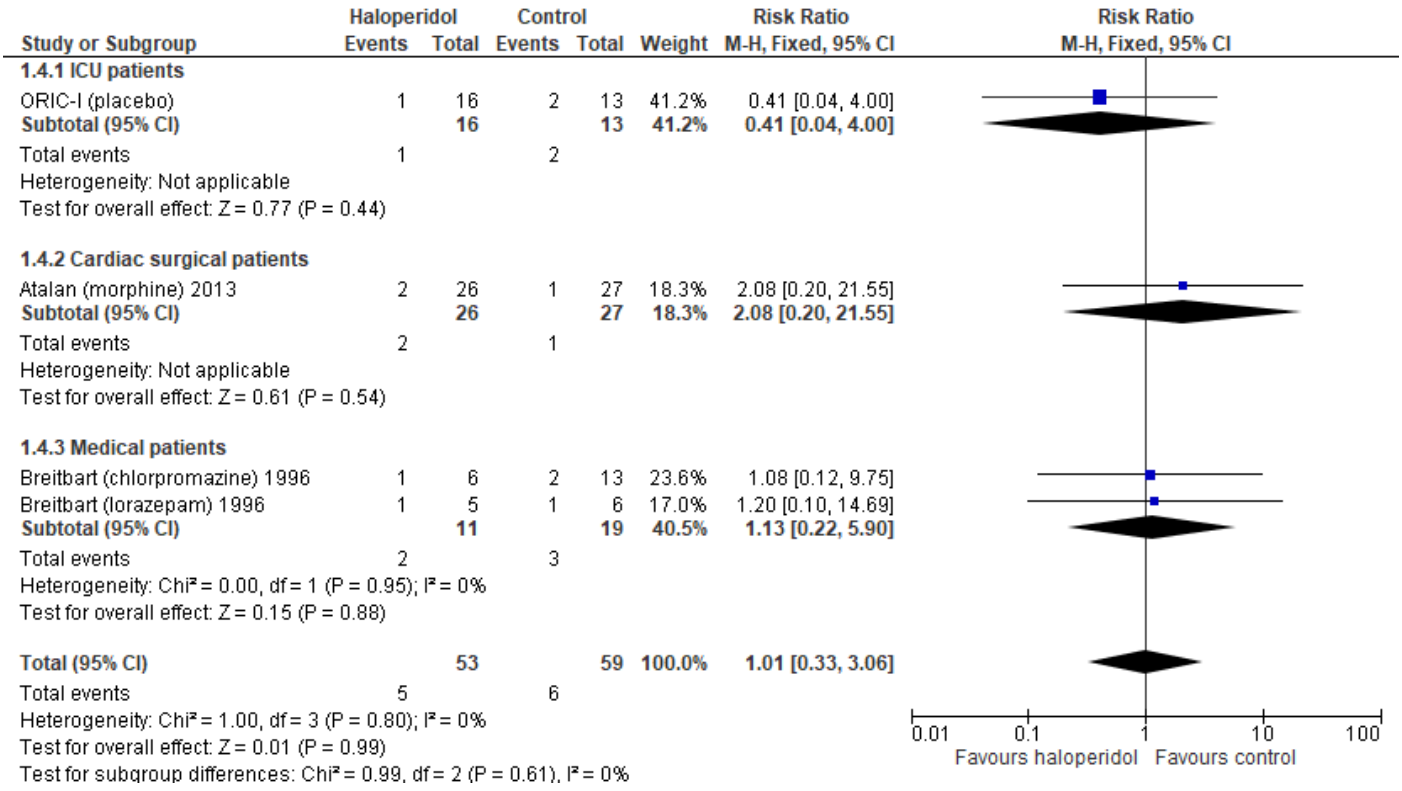
Excluding trials using rescue haloperidol

Forest plot of all-cause mortality stratified by type of delirium



Excluding trials using rescue haloperidol

Forest plot of all-cause mortality stratified by patient population



Excluding trials using rescue haloperidol

SENSITIVITY ANALYSES: AII-CAUSE MORTALITY, INCLUDING TRIALS USING RESCUE HALOPERIDOL

Sensitivity analysis including the trial using rescue haloperidol, a total of 31.8% of the participants in the haloperidol group versus 30.9% of the participants in the control group died. Meta-analysis of all four trials (six comparisons) regardless of risk of bias showed no evidence of a difference of haloperidol versus control for the treatment of delirium when assessing mortality (fixed effect model RR 1.10; 95% CI 0.88-1.37; I^2=0%; 678 participants; 4 trials; 6 comparisons).

Trial Sequential Analysis showed that with an anticipated RRI of 20%, a mortality in the control group of 30.9%, a type 1 error of 3.3%, a type 2 error of 10% and a diversity of 0%, the required information size was 2732 participants; thus only 25% of the required information size had been reached. The cumulative Z-curve did not cross any boundaries for benefit or harm, nor trial sequential monitoring boundaries for futility, indicating that, considering sparse data and repetitive testing, evidence was insufficient to confirm or refute a 20% RRI of haloperidol versus control on all-cause mortality. The TSA adjusted CI was 0.65-1.89. The TSA sensitivity analyses did not have the potential to influence the results.

Bayes Factors

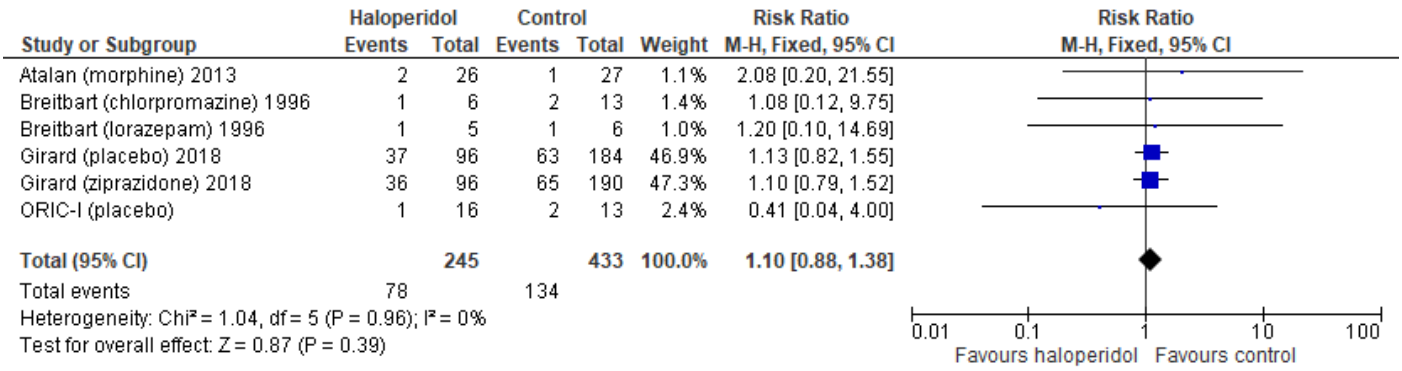
A Bayes factor of 0.9 was calculated based on an a priori RR of 1.20 and the meta-analysis result (RR 1.10) supporting that the result is likely 0.9 times more compatible with the null-hypothesis of a RRR = 0% than the alternative hypothesis of a RRR = 20%, for an effect of haloperidol versus control for the treatment of delirium on all-cause mortality.

A Bayes factor of 26 was calculated based on an a priori RR of 0.80 and the meta-analysis result (RR 1.10) supporting that the result is likely 26 times more compatible with the null-hypothesis of a RRR = 0% than the alternative hypothesis of a RRR = 20%, for an effect of haloperidol versus control for the treatment of delirium on all-cause mortality.

The sensitivity analyses on missing data indicated that incomplete outcome data did not have the potential to influence the results (best-worst case scenario and worst-best case scenario.

The subgroup analysis excluding trials at overall high risk of bias could not be performed as only one trial was overall low risk of bias [4]. We found no evidence of a difference in subgroup analyses stratified by overall risk of bias, used control intervention, patient population and type of delirium.

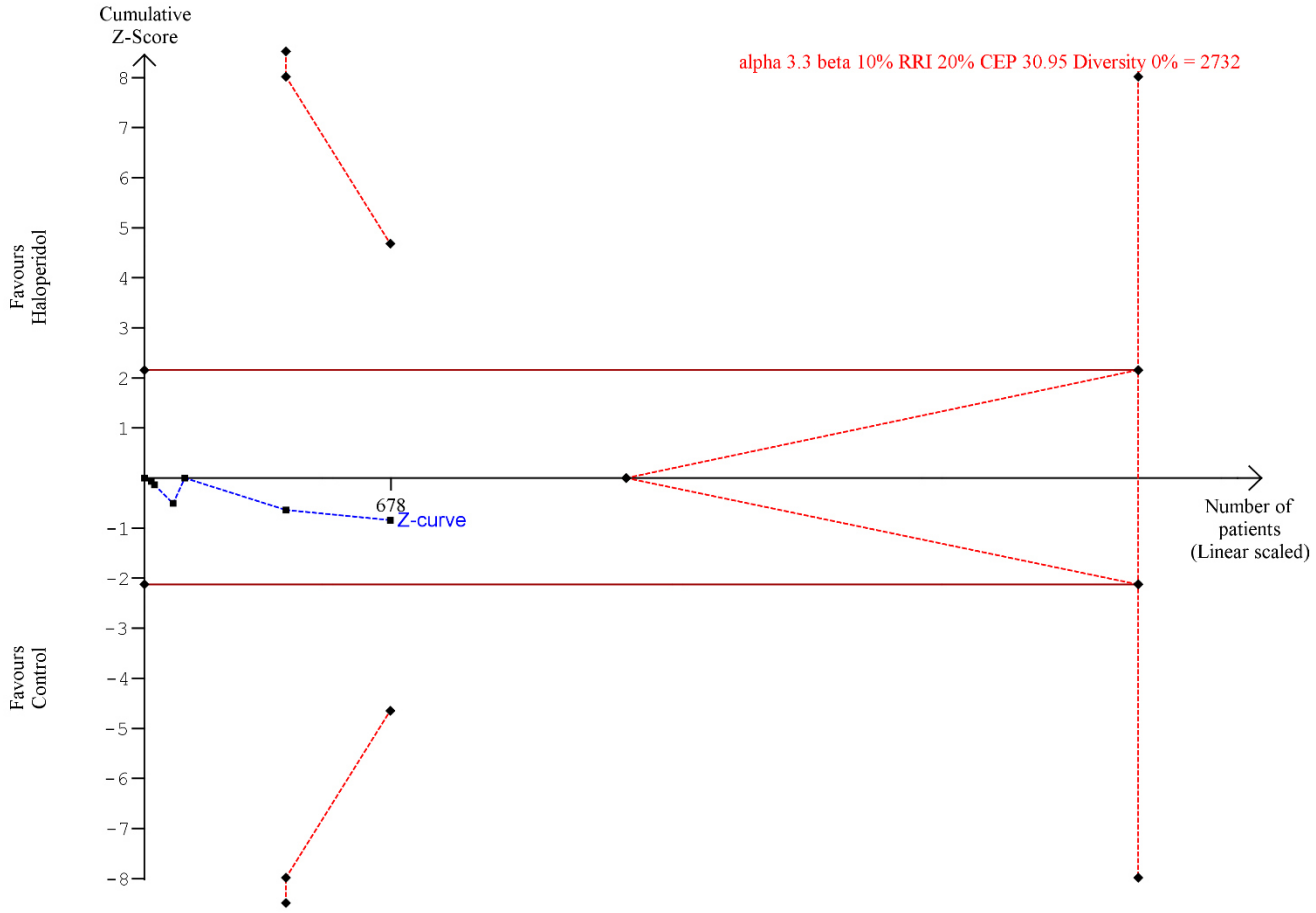
Forest plot of all-cause mortality



Including trials using rescue haloperidol.

Forest plot of all-cause mortality in all trials regardless of used control intervention and overall risk of bias stratified by control intervention. Size of squares for risk ratio reflects weight of trial in pooled analysis. Horizontal bars represent 95% confidence intervals.

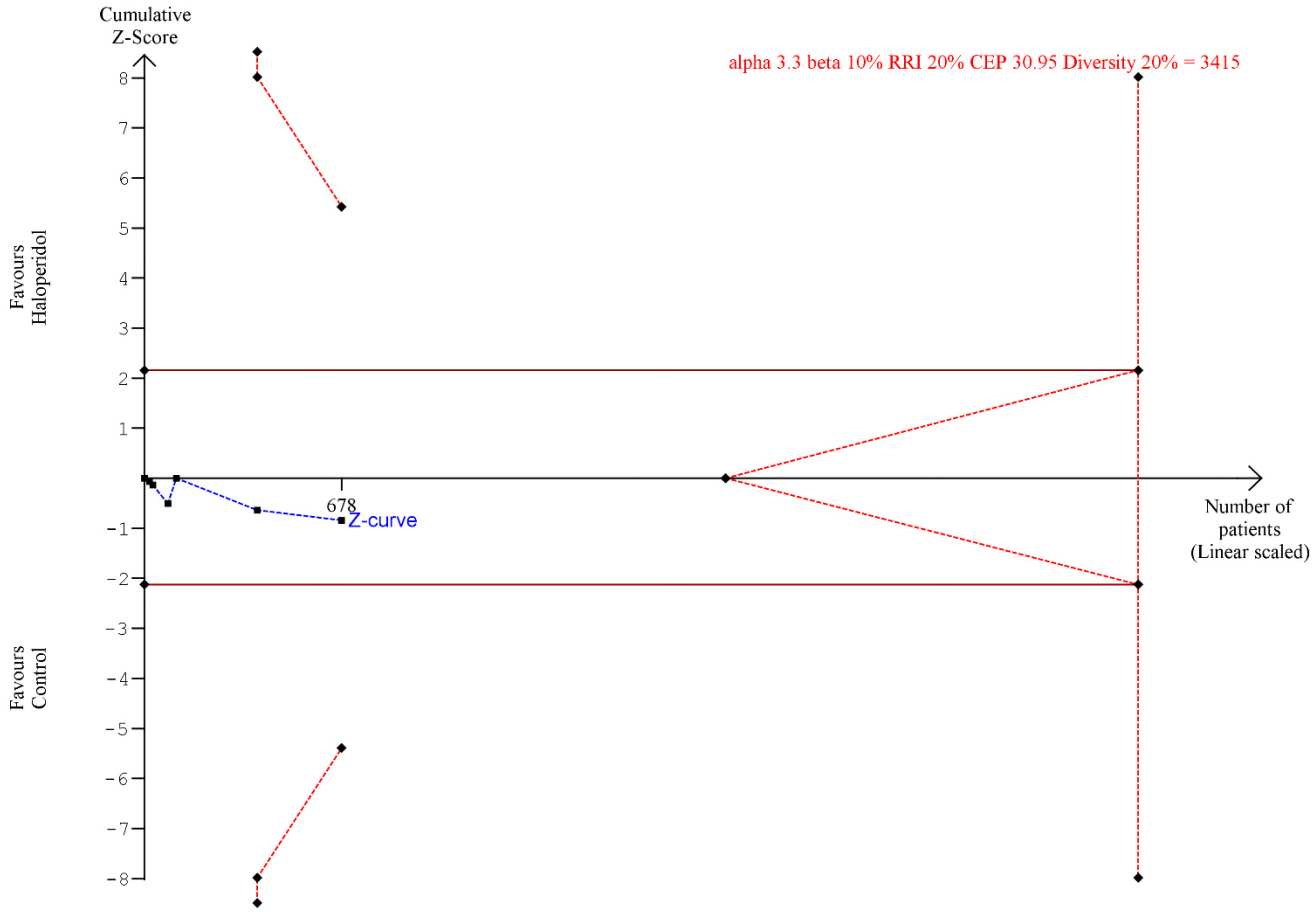
TSA sensitivity analysis: all-cause mortality



Including trials using rescue haloperidol.

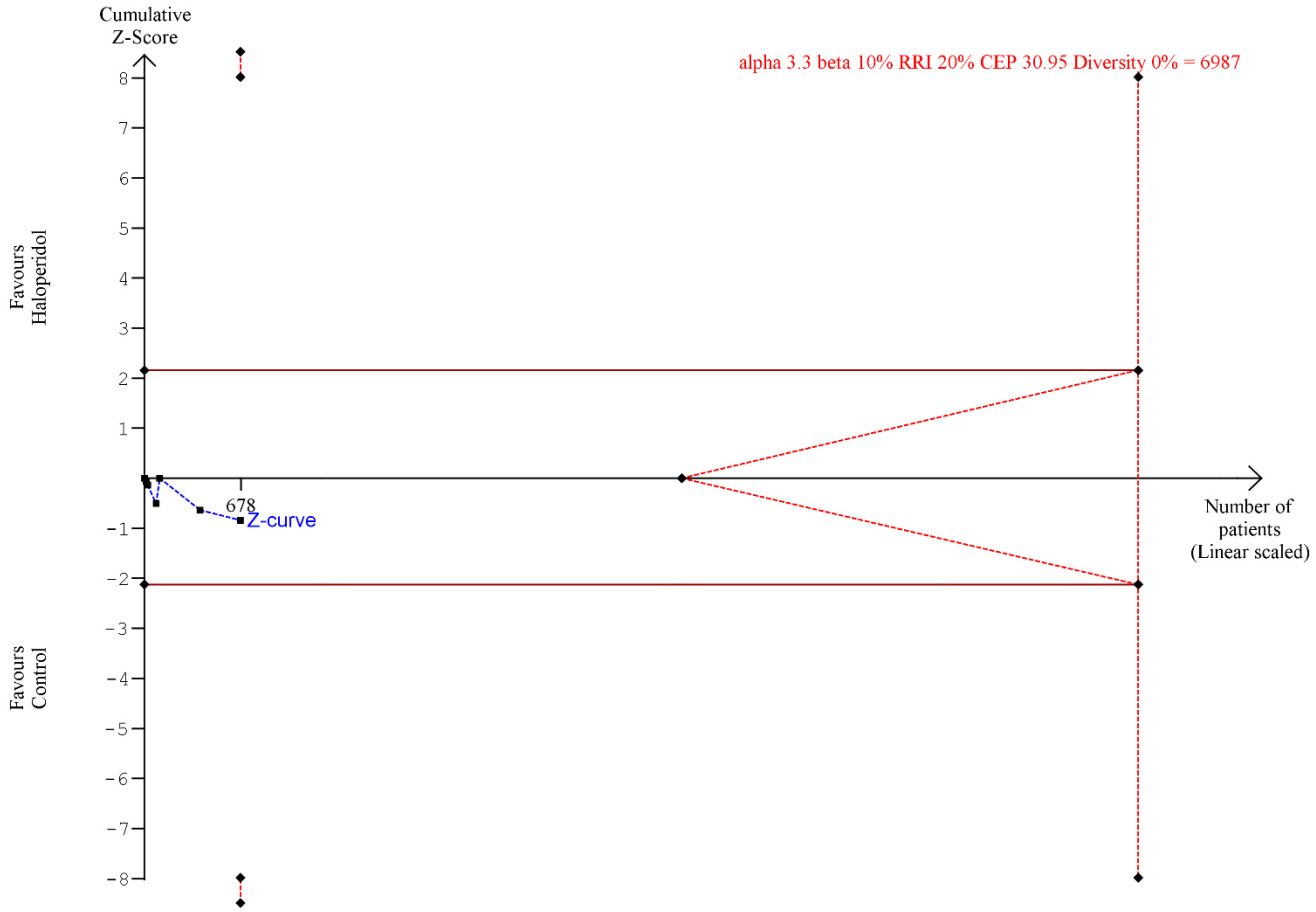
Trial Sequential Analysis for all-cause mortality in all trials regardless of used control intervention and overall risk of bias. We used an alpha of 3.3%, a power of 90%, a relative risk increase of 20%, a control event proportion of 30.95% (from the included trials) and a diversity (D^2) of 0%. The risk ratio was 1.10 with a TSA-adjusted confidence interval 0.65-1.89. As the cumulative Z-curve does not reach the trial sequential monitory boundaries, futility area or required information size, we not have enough information to detect or reject a 20% relative risk increase or reduction.

TSA sensitivity analysis: all-cause mortality, with a diversity of 20%



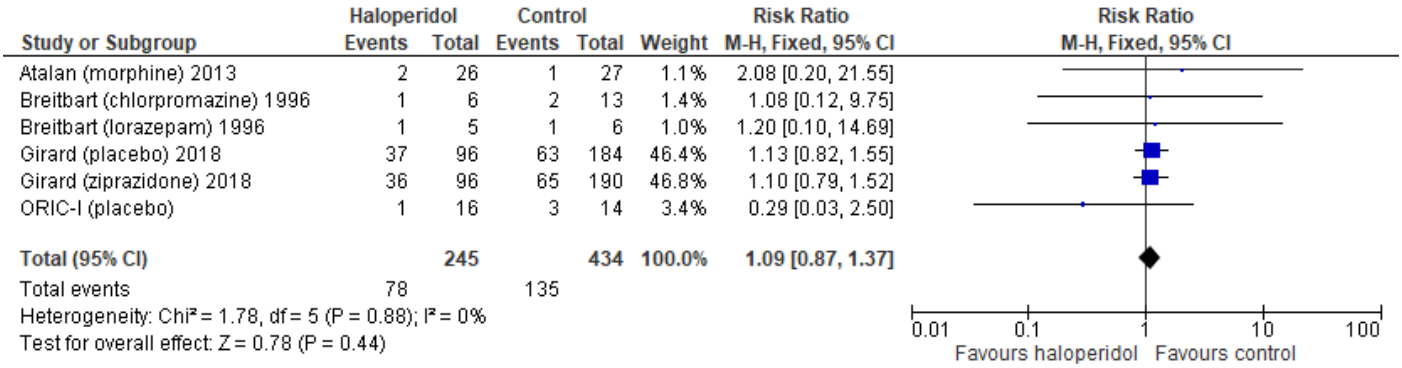
Including trials using rescue haloperidol.
We used a diversity of 20% if the actual measured heterogeneity was in fact zero because in this case heterogeneity will most likely increase when further trials are added until the required information size is reached. TSA-adjusted confidence interval changed from 0.65-1.89 to 0.61-1.99 when changing diversity from 0% to 20%.

TSA sensitivity analysis: all-cause mortality, for a 12% RRR



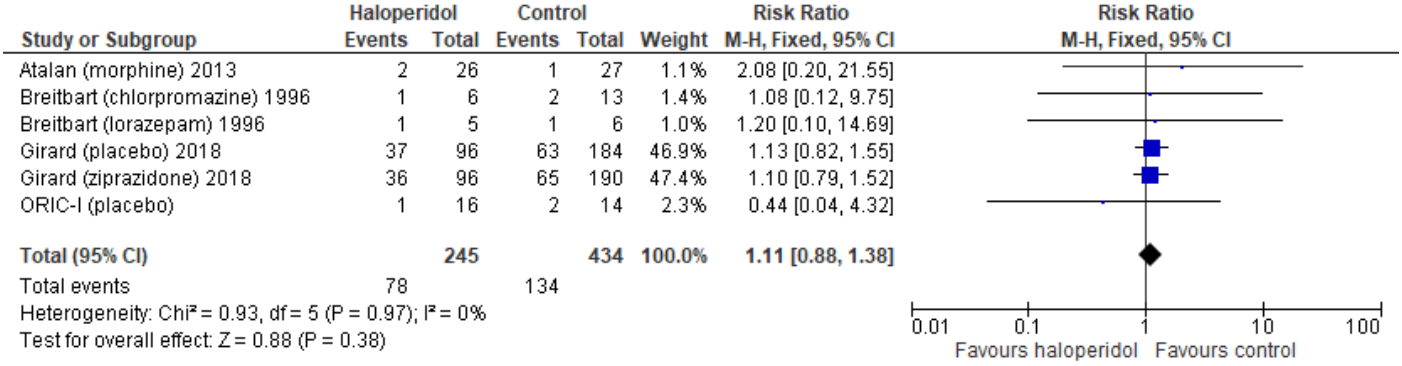
Including trials using rescue haloperidol.
We pre-planned to conduct a TSA using the RRR or RRI based on the confidence limit closets to null effect in the 95% CI in the traditional analysis. The 95% CI was 0.88-1.38, thus, we used a 12% RRR.
TSA-adjusted CI 0.44-2.77

Forest plot of best-worst case scenario sensitivity analysis for missing data on mortality



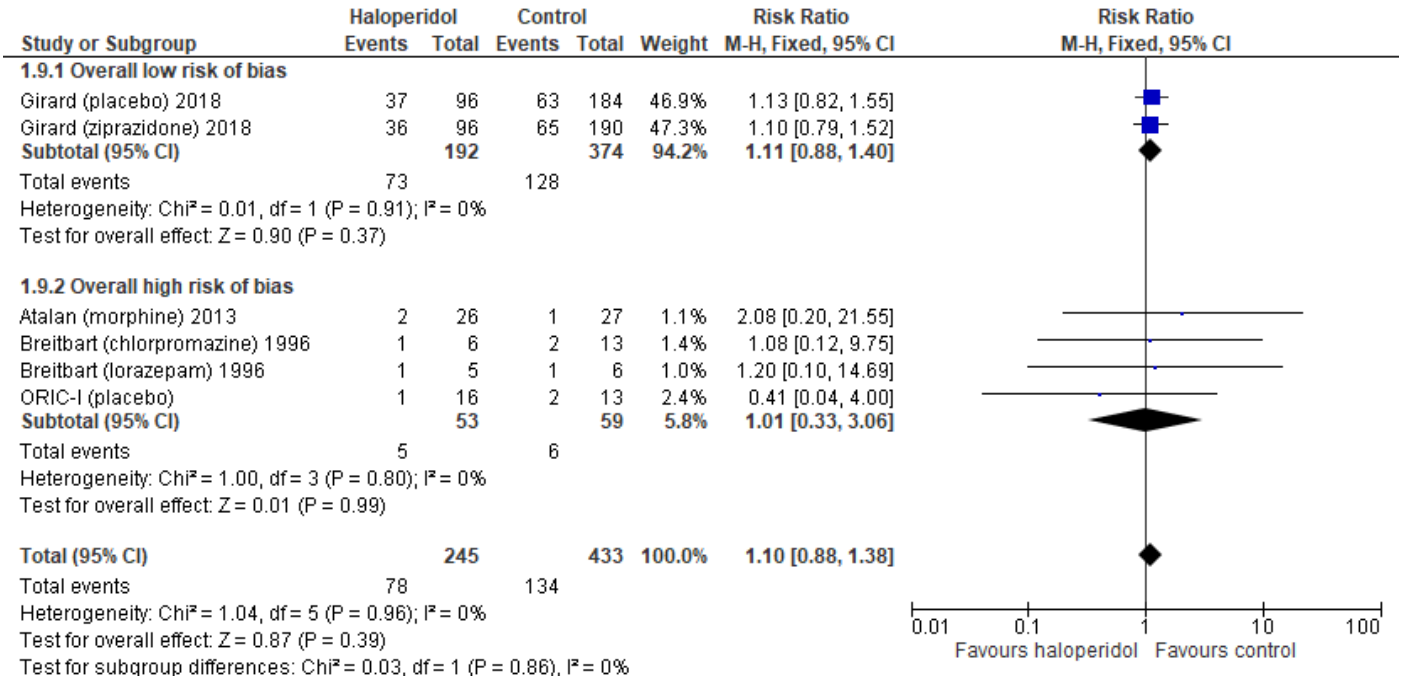
Including trials using rescue haloperidol.
In this analysis it is assumed that all participants lost to follow-up in the haloperidol group have survived and all those with missing outcomes in the control group have died. The primary meta-analysis result (RR 1.10, 95% CI 0.88-1.38) and the result from this sensitivity analysis (RR 1.09, 95% CI 0.87-1.37) show similar P values and CIs.

Forest plot of worst-best case scenario sensitivity analysis for missing data on mortality



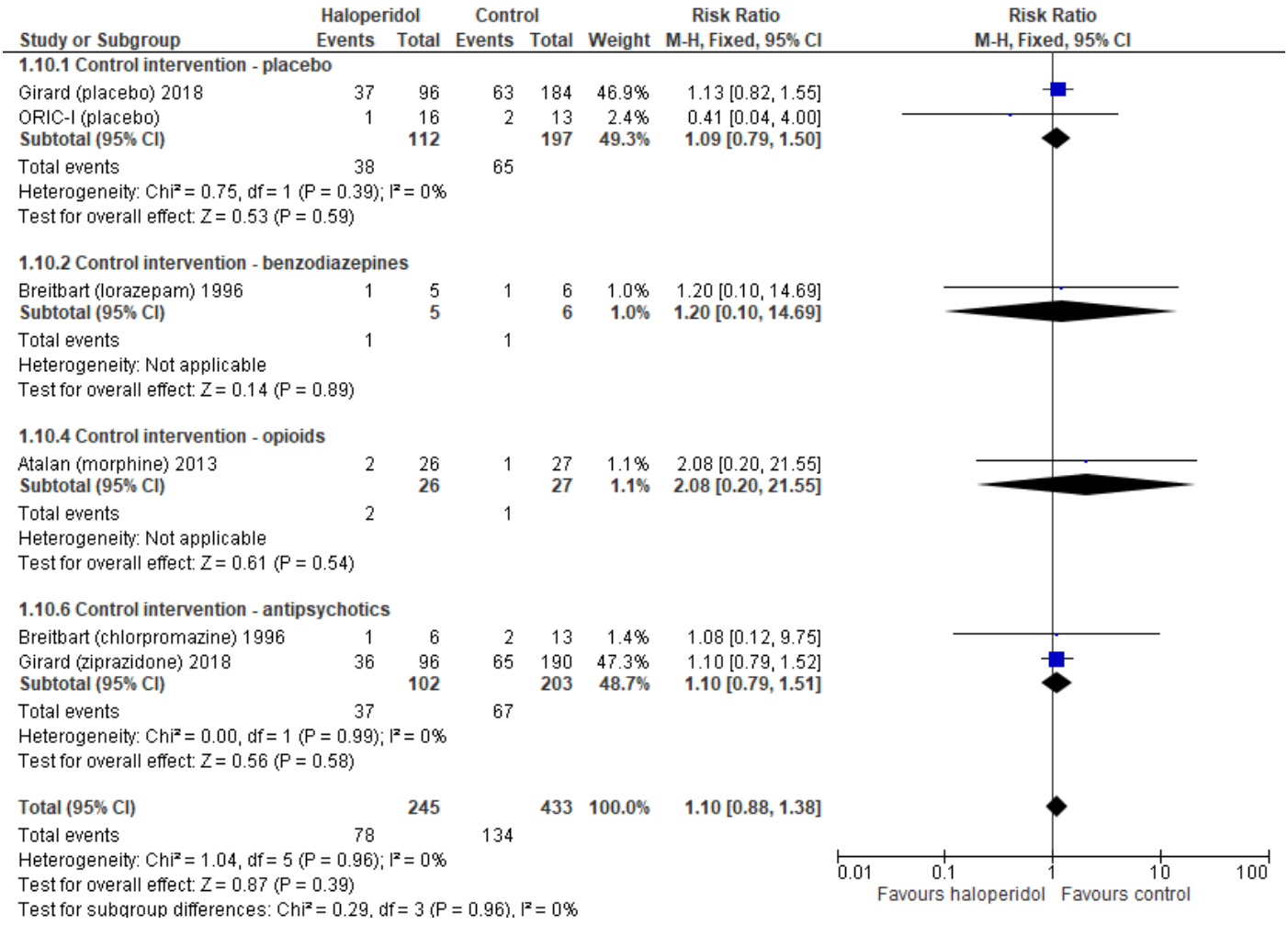
Including trials using rescue haloperidol.
In this analysis it is assumed that all participants lost to follow-up in the haloperidol group died and all those with missing outcomes in the control group have survived. The primary meta-analysis result (RR 1.10, 95% CI 0.88-1.38) and the result from this sensitivity analysis (RR 1.11, 95% CI 0.88-1.38) show similar P values and CIs.

Forest plot of mortality stratified by overall risk of bias



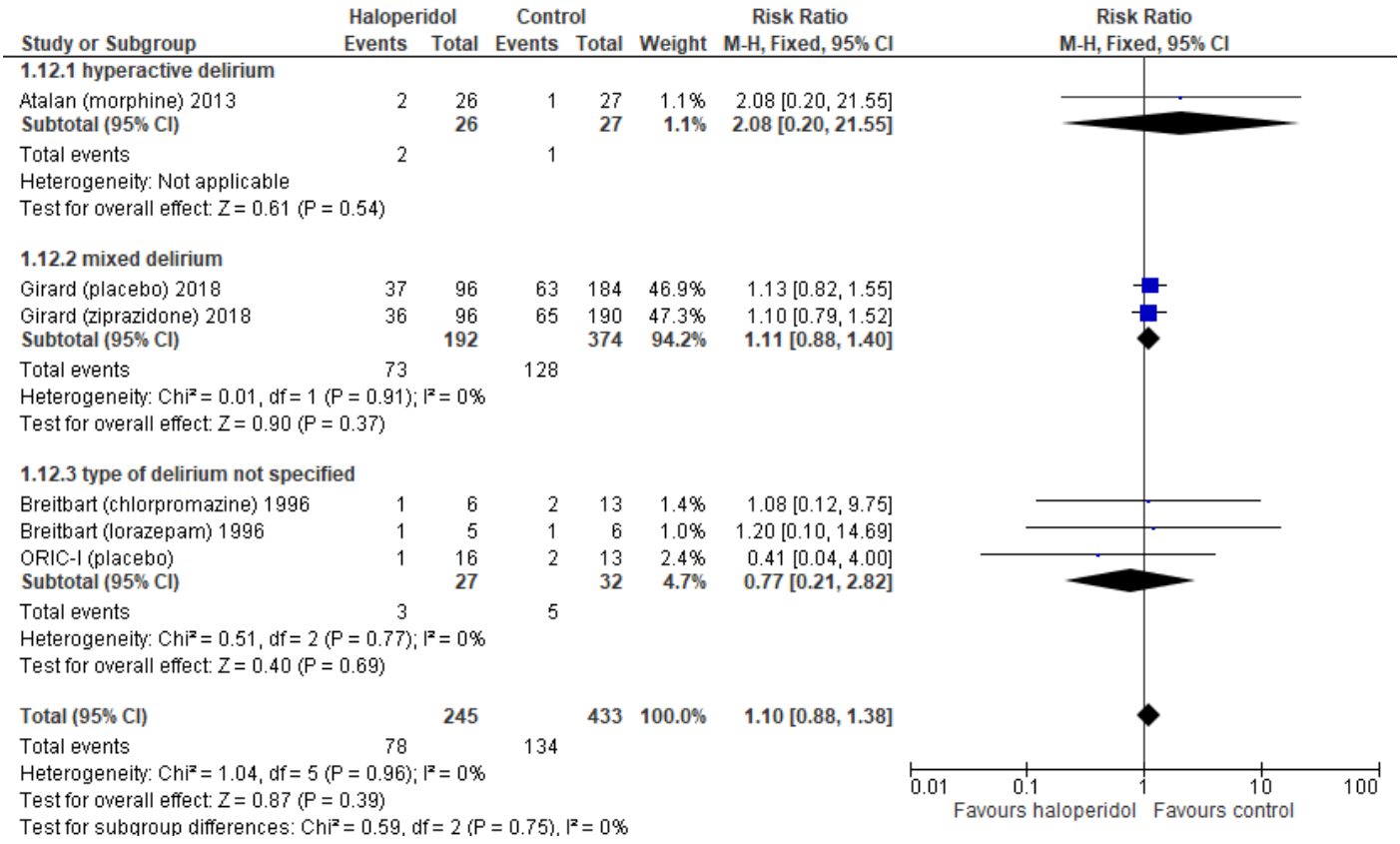
Including trials using rescue haloperidol.
Subgroup analysis stratified according to overall risk of bias.

Forest plot of all-cause mortality stratified by used control intervention



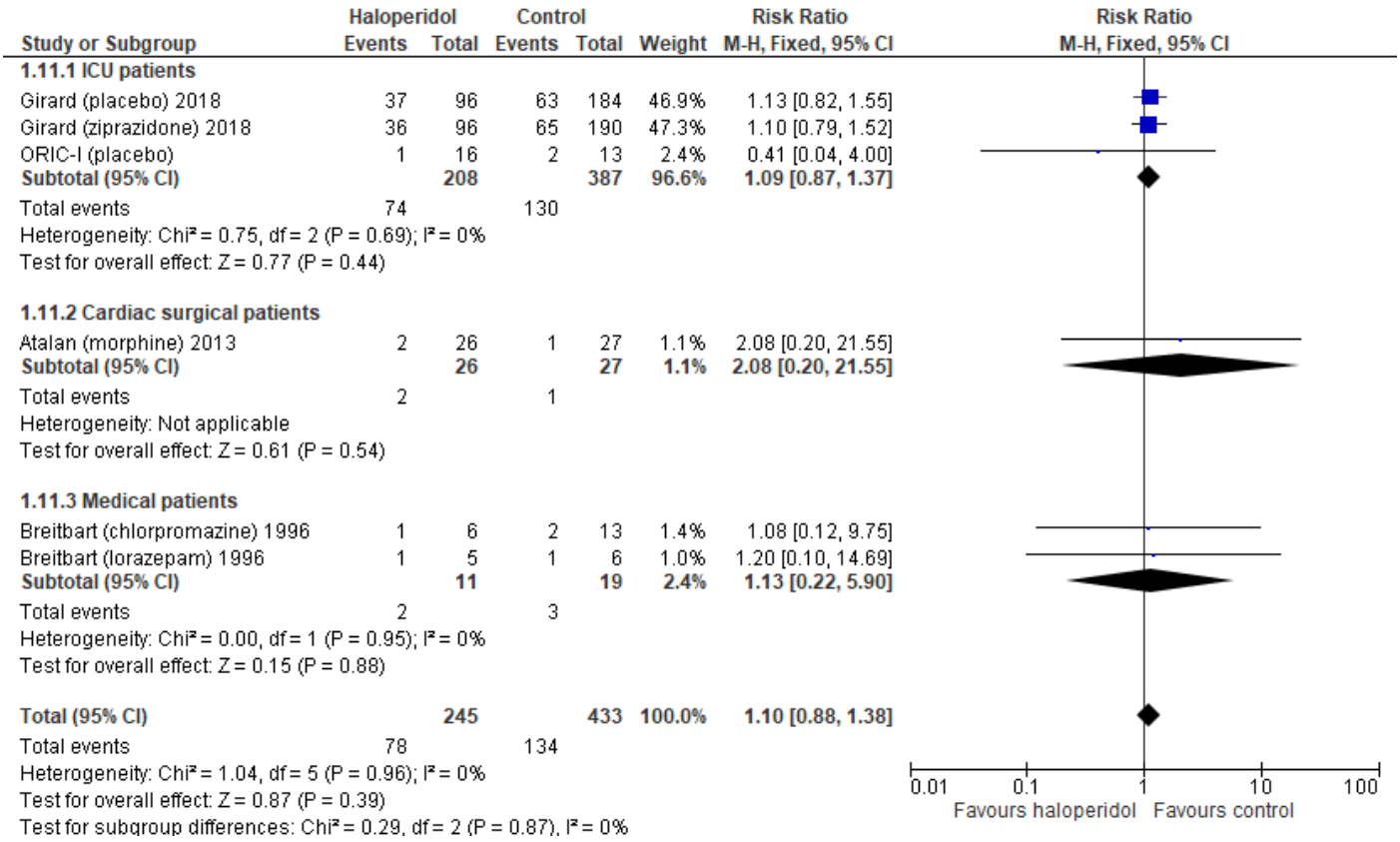
Including trials using rescue haloperidol.
Subgroup analysis stratified according to used control intervention.

Forest plot of all-cause mortality stratified by type of delirium



Including trials using rescue haloperidol.
Subgroup analysis stratified according to type of delirium.

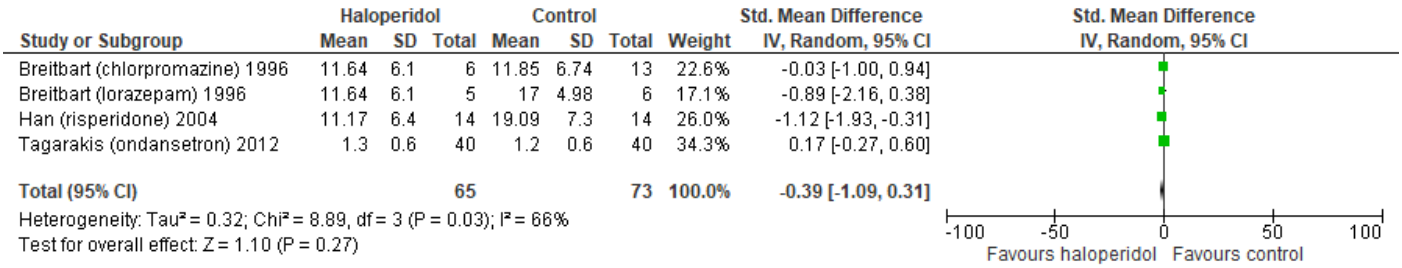
Forest plot of all-cause mortality stratified by patient population



Including trials using rescue haloperidol.
Subgroup analysis stratified according to patient population.

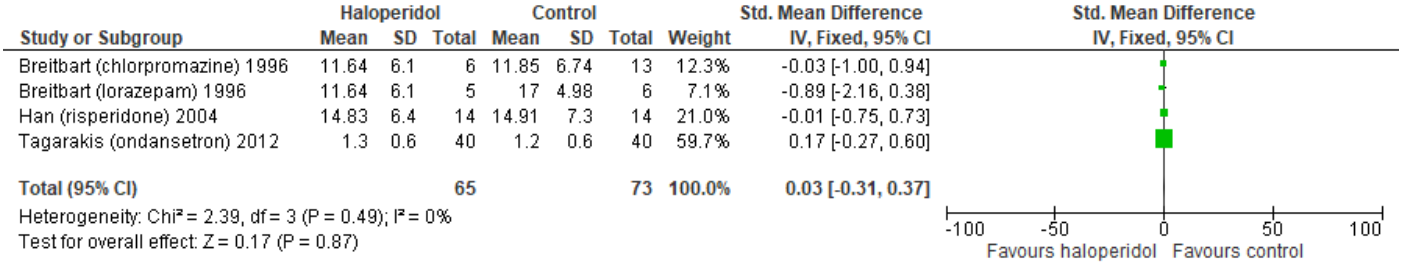
SEVERITY OF DELIRIUM, EXCLUDING TRIALS USING RESCUE HALOPERIDOL

Forest plot of best-worst case scenario sensitivity analysis for missing data on delirium severity



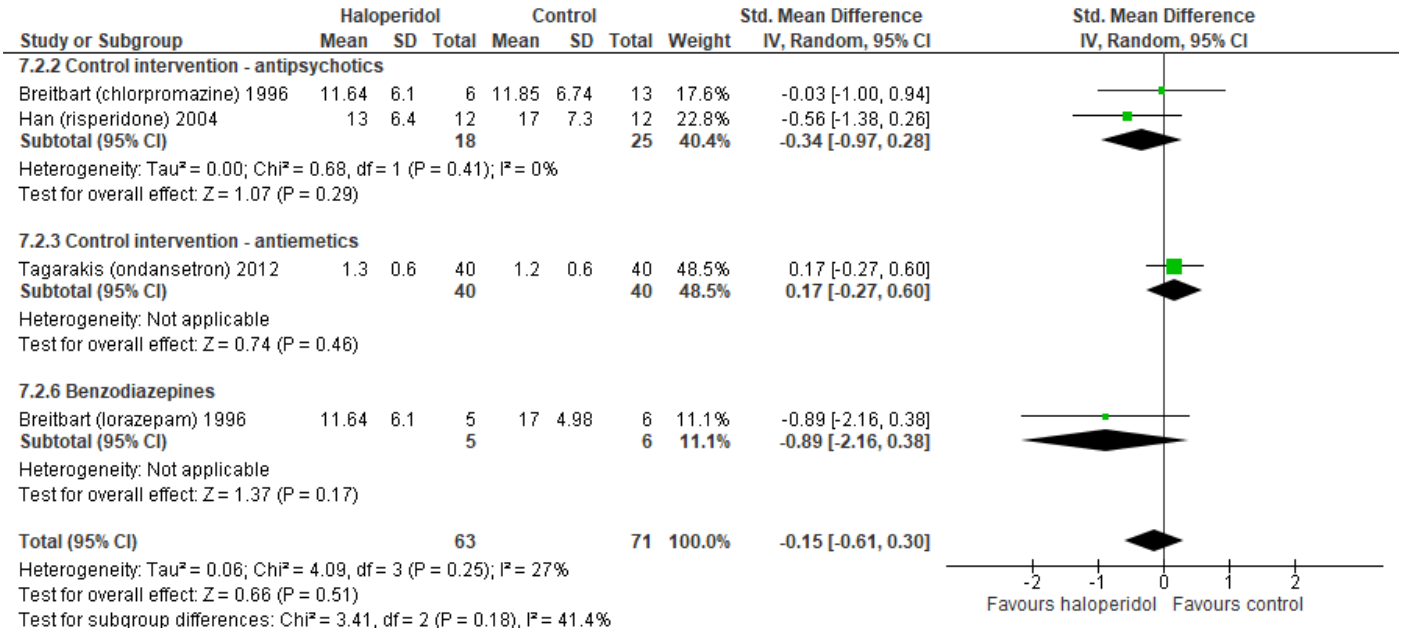
Excluding trials using rescue haloperidol.
In this analysis it is assumed that all participants lost to follow-up in the haloperidol group have survived and all those with missing outcomes in the control group have died. The primary meta-analysis result (RR -0.15, 95% CI -0.61-0.30) and the result from this sensitivity analysis (RR -0.39, 95% CI -1.09-0.31) show similar P values and CIs.

Forest plot of worst-best case scenario sensitivity analysis for missing data on delirium severity



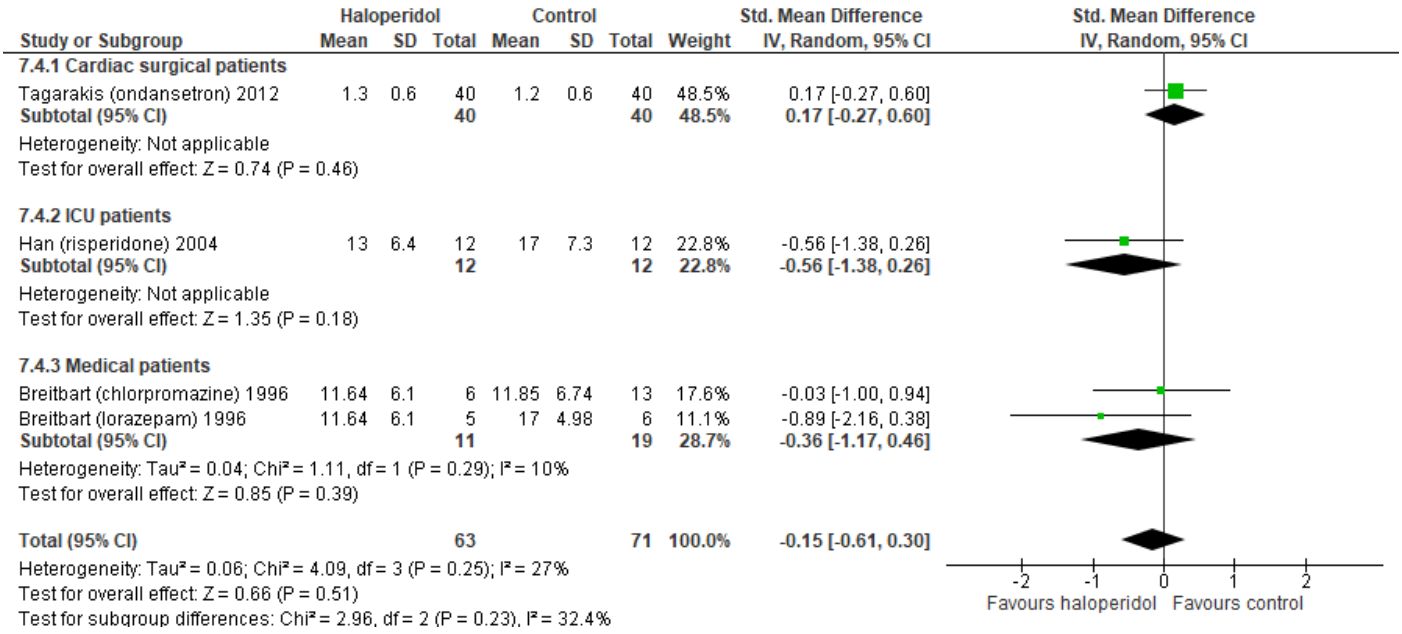
Excluding trials using rescue haloperidol.
In this analysis it is assumed that all participants lost to follow-up in the haloperidol group died and all those with missing outcomes in the control group have survived. The primary meta-analysis result (RR -0.15, 95% CI -0.61-0.30) and the result from this sensitivity analysis (RR 0.03, 95% CI -0.31-0.37) show similar P values and CIs.

Forest plot of delirium severity stratified by used control intervention



Excluding trials using rescue haloperidol.
Subgroup analysis stratified according to used control intervention.

Forest plot of delirium severity stratified by patient population



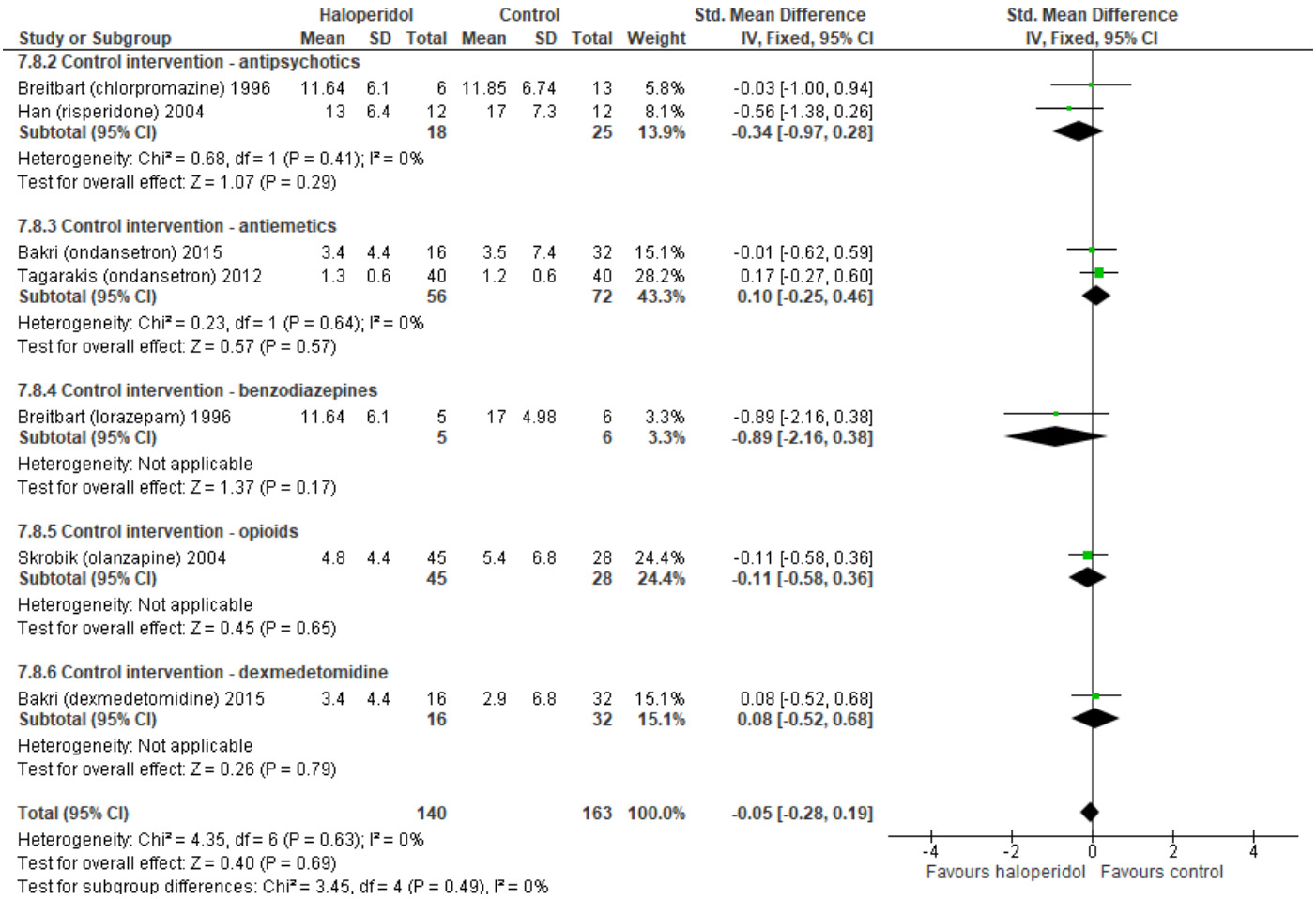
Excluding trials using rescue haloperidol.
Subgroup analysis stratified according to patient population.

SENSITIVITY ANALYSES: SEVERITY OF DELIRIUM, INCLUDING TRIALS USING RESCUE HALOPERIDOL

Sensitivity analysis including the trials using rescue haloperidol indicated no evidence of a difference of haloperidol versus control for the treatment of delirium when assessing delirium severity (fixed effect model SMD -0.05; 95% CI -0.28-0.19; I^2=0; 303 participants; 5 trials; 7 comparisons).

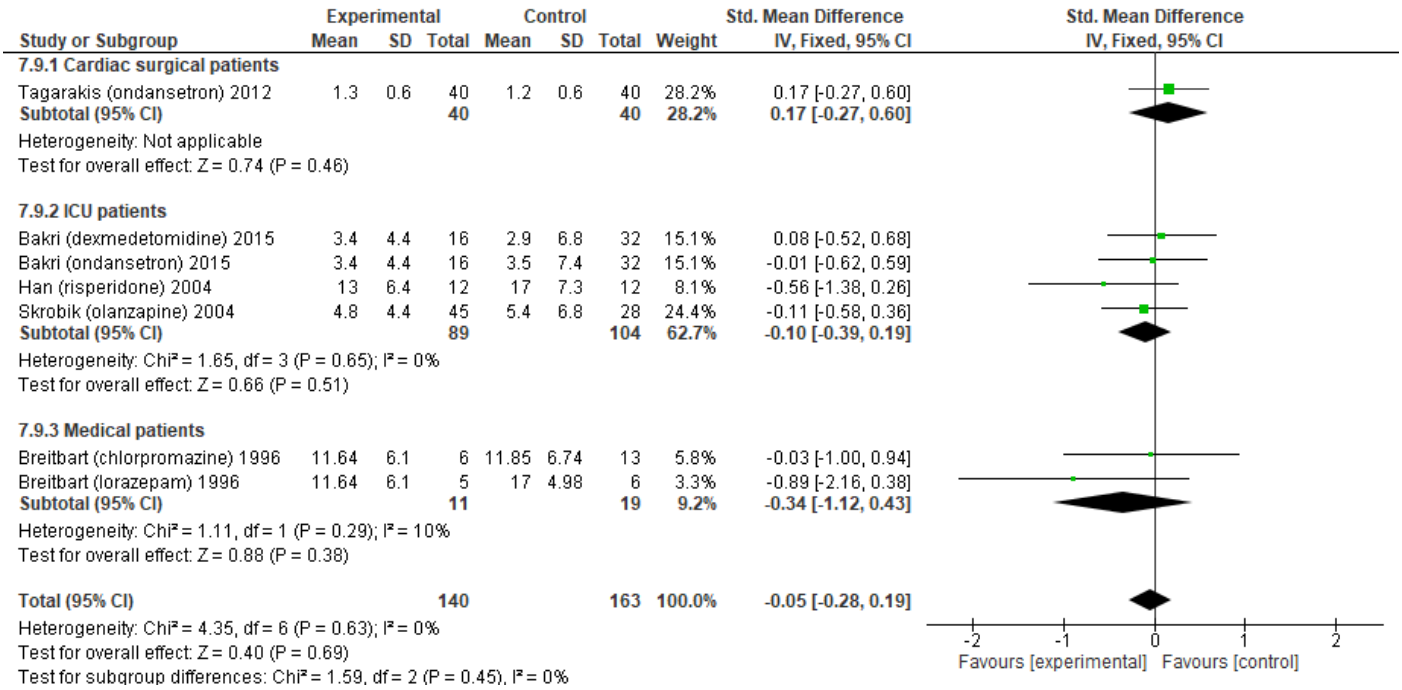
The subgroup analysis excluding trials at overall high risk of bias could not be performed as no trial was overall low risk of bias. Subgroup analysis on delirium type could not be performed as none of the four trials specified type of delirium. We found no evidence of a difference in subgroup analyses stratified to used control intervention and patient population.

Forest plot of delirium severity



Including trials using rescue haloperidol.
Forest plot of delirium severity in all trials regardless of used control intervention and overall risk of bias stratified by control intervention. No trials were overall low risk of bias. Size of squares for standardised mean difference reflects weight of trial in pooled analysis. Horizontal bars represent 95% confidence intervals.

Forest plot of delirium severity stratified by patient population



Including trials using rescue haloperidol.
Subgroup analysis stratified according to patient population.

SENSITIVITY ANALYSES: QTc PROLONGATION, INCLUDING TRIALS USING RESCUE HALOPERIDOL

Sensitivity analysis including the trials using haloperidol as rescue drug, a total of 6.67% of the participants in the haloperidol group versus 6.87% of the participants in the control group had QTc prolongation. Meta-analysis of all three trials (five comparisons), regardless of risk of bias, showed no evidence of a difference of haloperidol versus control when assessing QTc prolongation (random effects model RR 0.97; 95% CI 0.48-1.94; I^2=16%; 691 participants; three trials, five comparisons).

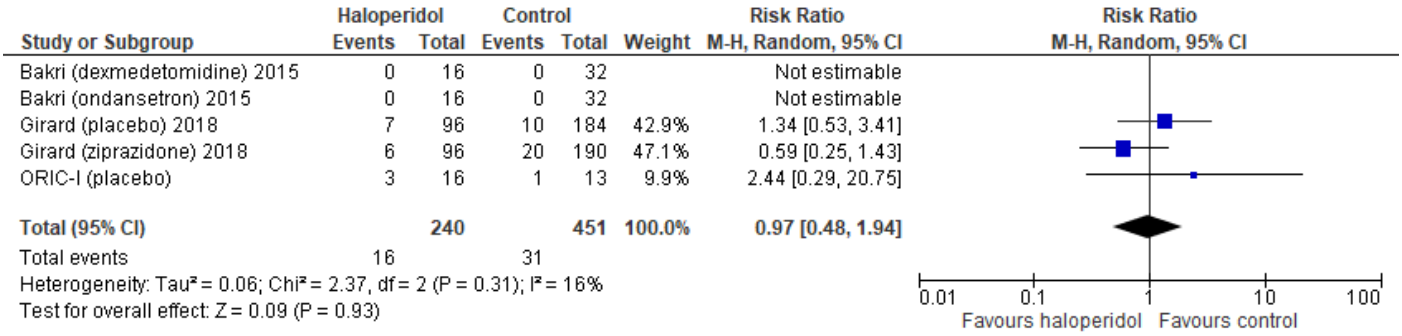
As only 2.53% of the required information size had been reached TSA adjusted-CI could not be calculated.

A Bayes factor of 1.29 was calculated based on an a priori RR of 1.20 and the meta-analysis result (RR 0.97) supporting that the result is likely 1.29 times more compatible with the null-hypothesis of a RRR of 0% than the alternative hypothesis of a RRR = 20%, for an effect of haloperidol versus control for the treatment of delirium on QTc prolongation.

A Bayes factor of 1.22 was calculated based on an a priori RR of 0.80 and the meta-analysis result (RR 0.97) supporting that the result is likely 1.22 times more compatible with the null-hypothesis of a RRR = 0% than the alternative hypothesis of a RRR = 20%, for an effect of haloperidol versus control for the treatment of delirium on all-cause mortality.

The sensitivity analyses on missing data indicated that incomplete outcome data did not have the potential to influence the results (best-worst case scenario and worst-best case scenario.
The subgroup analysis excluding trials at overall high risk of bias could not be performed as only one trial was overall low risk of bias. We found no evidence of a difference in subgroup analysis stratified to used control intervention. Subgroup analyses on patient population and type of delirium could not be performed due to all patients being ICU patients and due to lack of trials reporting on delirium type.

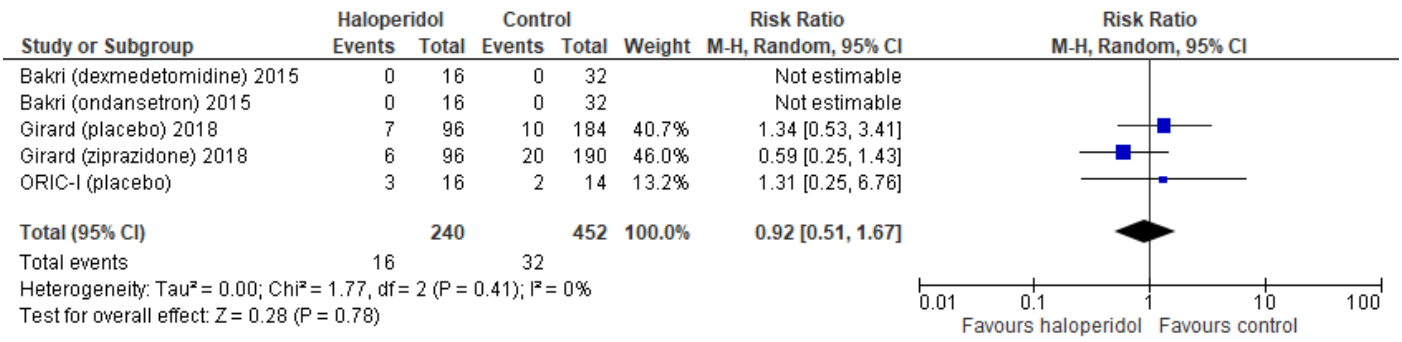
Forest plot of QTc prolongation



Including trials using rescue haloperidol.

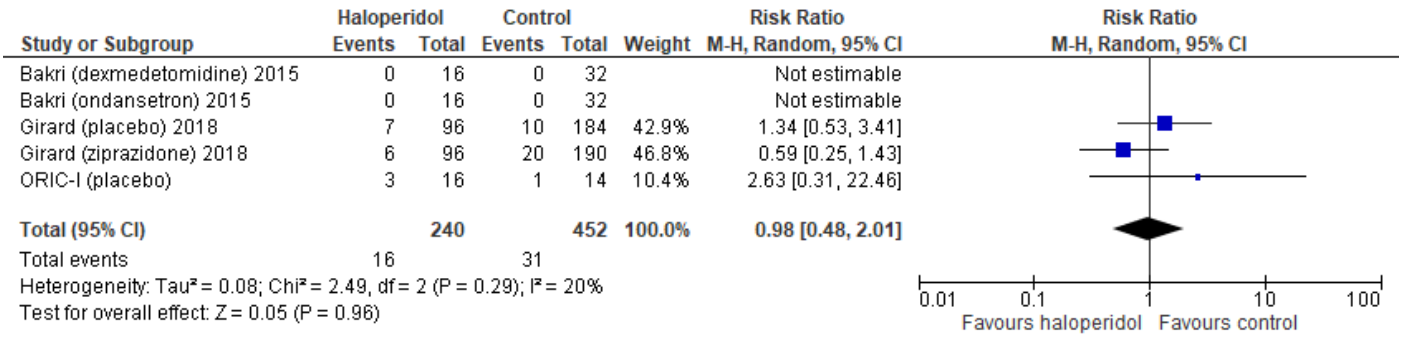
Forest plot of QTc prolongation in all trials regardless of used control intervention and overall risk of bias stratified by control intervention. Size of squares for standardised mean difference reflects weight of trial in pooled analysis. Horizontal bars represent 95% confidence intervals.

Forest plot of best-worst case scenario sensitivity analysis for missing data on QTc prolongation



Including trials using rescue haloperidol.
In this analysis it is assumed that all participants lost to follow-up in the haloperidol group have survived and all those with missing outcomes in the control group have died. The primary meta-analysis result (RR 0.97, 95% CI 0.48-1.94) and the result from this sensitivity analysis (RR 0.92, 95% CI 0.51-1.67) show similar P values and CIs.

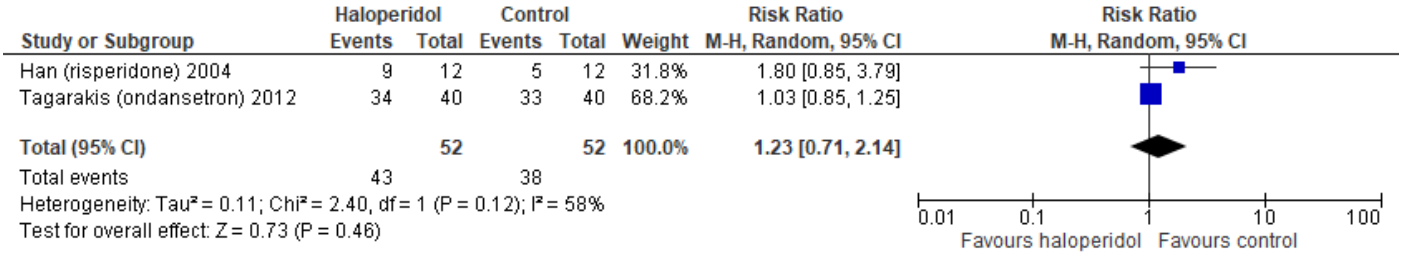
Forest plot of worst-best case scenario sensitivity analysis for missing data on QTc prolongation



Including trials using rescue haloperidol.
In this analysis it is assumed that all participants lost to follow-up in the haloperidol group died and all those with missing outcomes in the control group have survived. The primary meta-analysis result (RR 0.97, 95% CI 0.48-1.94) and the result from this sensitivity analysis (RR 0.98, 95% CI 0.48-2.01) show similar P values and CIs.

DELIRIUM RESOLUTION, EXCLUDING TRIALS USING RESCUE HAOPERIDOL

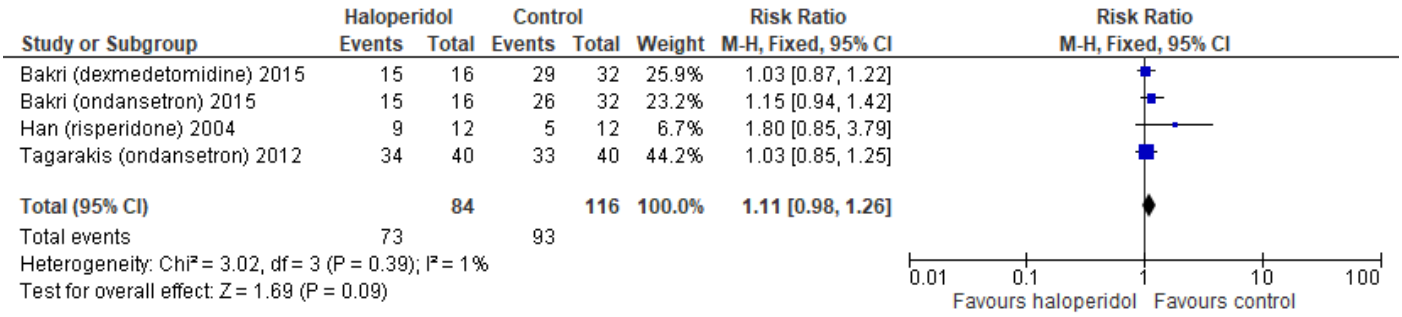
Forest plot of delirium resolution



Excluding trials using rescue haloperidol.
Forest plot of delirium resolution regardless of used control intervention. Meta-analysis showed that haloperidol versus control for the treatment does not reduce/increase delirium resolution. All trials were overall high risk of bias.

SENSITIVITY ANALYSES: DELIRIUM RESOLUTION, INCLUDING TRIALS USING RESCUE HAOPERIDOL

Forest plot of delirium resolution



Including trials using rescue haloperidol.

Meta-analysis including trials using haloperidol as rescue drug. Meta-analysis showed that haloperidol versus control for the treatment does not reduce/increase delirium resolution. All trials were overall high risk of bias.

EXTRAPYRAMIDAL SYMPTOMS, EXCLUDING TRIALS USING RESCUE HALOPERIDOL

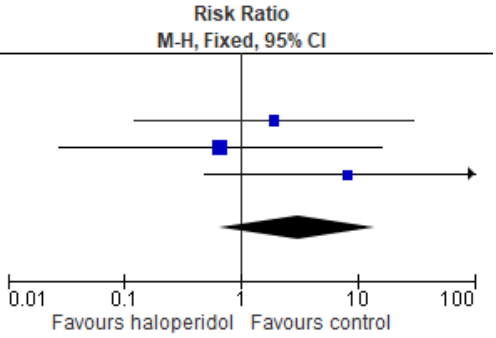
Only one trial that did not use rescue haloperidol reported on extrapyramidal symptoms [3]. Thus, meta-analysis was not conducted.

SENSITIVITY ANALYSES: EXTRAPYRAMIDAL SYMPTOMS, INCLUDING TRIALS USING RESCUE HALOPERIDOL

Forest plot of extrapyramidal symptoms

	Experimental		Control			Risk Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Fixed, 95% CI
Breitbart (chlorpromazine) 1996	0	6	0	13		Not estimable
Breitbart (lorazepam) 1996	0	5	0	6		Not estimable
Girard (placebo) 2018	1	96	1	184	29.7%	1.92 [0.12, 30.31]
Girard (ziprazidone) 2018	0	96	1	190	43.8%	0.66 [0.03, 15.96]
Skrobik (olanzapine) 2004	6	45	0	28	26.6%	8.20 [0.48, 140.09]
Total (95% CI)		**248**		**421**	**100.0%**	**3.03 [0.65, 14.20]**
Total events	7		2			

Heterogeneity: Chi² = 1.46, df = 2 (P = 0.48); I² = 0%
Test for overall effect: Z = 1.41 (P = 0.16)



Including trials using rescue haloperidol.
Meta-analysis showed that haloperidol versus control for the treatment does not reduce/increase extrapyramidal symptoms.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

1. In our protocol, we stated that we would search for eligible trials in Allied and Complementary Medicine Database (AMED). We could, however, not access AMED, and no search was conducted.

2. We used a power of 90%, and not 80% as pre-defined in the review protocol [18], as meta-analysis should use higher (or same) power as its included trials, to be able to communicate the best available evidence.

3. In our protocol, we stated that we would use a statistical significance level of 2% and 98% CIs for each of the four co-secondary outcomes. However, as no trials reported on HRQoL, we only report on three co-secondary outcomes and therefore we used a significance level of 2.5% and 97.5% CIs.

4. In our protocol, we defined the control group as those receiving placebo or any type of pharmacological (besides haloperidol) or non-pharmacological intervention. As we identified 3 trials using haloperidol as rescue medication, we conducted post-hoc sensitivity analyses where we included the three trials using rescue haloperidol.

REFERENCE LIST

1. Atalan N, Efe Sevim M, Akgun S, Fazliogullari O, Basaran C. Morphine is a reasonable alternative to haloperidol in the treatment of postoperative hyperactive-type delirium after cardiac surgery. J Cardiothorac Vasc Anesth 2013;27(5):933-8.
2. Bakri MH IE, Ibrahim A Comparison of dexmedetomidine or ondansetron with haloperidol for treatment of postoperative delirium in trauma patients admitted to intensive care unit: randomized controlled trial. Anaesth, pain & intensive care 2015;19(2):118-23.
3. Breitbart W, Marotta R, Platt MM, Weisman H, Derevenco M, Grau C, Corbera K, Raymond S, Lund S, Jacobson P. A double-blind trial of haloperidol, chlorpromazine, and lorazepam in the treatment of delirium in hospitalized AIDS patients. Am J Psychiatry 1996;153(2):231-7.
4. Girard TD, Exline MC, Carson SS, Hough CL, Rock P, Gong MN, Douglas IS, Malhotra A, Owens RL, Feinstein DJ, Khan B, Pisani MA, Hyzy RC, Schmidt GA, Schweickert WD, Hite RD, Bowton DL, Masica AL, Thompson JL, Chandrasekhar R, Pun BT, Strength C, Boehm LM, Jackson JC, Pandharipande PP, Brummel NE, Hughes CG, Patel MB, Stollings JL, Bernard GR, Dittus RS, Ely EW. Haloperidol and Ziprasidone for Treatment of Delirium in Critical Illness. N Engl J Med 2018;379(26):2506-2516.
5. Han CS, Kim YK. A double-blind trial of risperidone and haloperidol for the treatment of delirium. Psychosomatics 2004;45(4):297-301.
6. NCT00300391 ORIC-I: Optimizing Recovery From Intensive Care: Mechanical Ventilation and Delirium, https://clinicaltrials.gov/ct2/show/NCT00300391?term=oric-I&rank=1 (accessed 4 June 2019).
7. Skrobik YK, Bergeron N, Dumont M, Gottfried SB. Olanzapine vs haloperidol: treating delirium in a critical care setting. Intensive Care Med 2004;30(3):444-9.
8. Tagarakis GI, Voucharas C, Tsolaki F, Daskalopoulos ME, Papaliagkas V, Parisis C, Gogaki E, Tsagalas I, Sataitidis I, Tsolaki M, Tsilimingas NB. Ondasetron versus haloperidol for the treatment of postcardiotomy delirium: a prospective, randomized, double-blinded study. J Cardiothorac Surg 2012;7:25.

ELECTRONICALLY SUPLEMENTARY MATERIAL (ESM)

This supplementary has been provided by the authors to give readers additional information about their work.

Supplement to: Barbateskovic M, Marker S, Granholm A, et al. Stress ulcer prophylaxis with proton pump inhibitors or histamin-2 receptor antagonists in adult intensive care patients: a systematic review with meta-analysis and Trial Sequential Analysis

Supplementary material for
Stress ulcer prophylaxis with proton pump inhibitors or histamin-2 receptor antagonists in adult intensive care patients: a systematic review with meta-analysis and Trial Sequential Analysis

Table S1. PRISMA Checklist

Section/topic	#	Checklist item	Line number(s)
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1-3
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	43-60
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	74-88
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	86-88, 97-98
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	90 (published protocol), 91-92 (PROSPERO)
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	96-101, 124-133
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	102-109
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	ESM
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	111-114
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	116-122, 161-162
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	118-122, S2

Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	136-143
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	149-150
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	150-176
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	144-145
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	162-170, 178-213
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	221-223, Fig.1.
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	S3-S4, ESM
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	Fig. 2, S4.
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	Fig. 3+4, S20, S25+26, S30, S32, S37
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	236-323
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	248, 270, 288, 309
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	240-323
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	325-332, Table 1
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	340-358
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	375-392
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	403-405

Search strategies

Cochrane Central Register of Controlled Trials 11.10.2018 (issue 10 of 12, October 2018)

#1 MeSH descriptor: [Peptic Ulcer] explode all trees and with qualifier(s): [Drug therapy - DT, Prevention & control - PC, Therapy - TH]
#2 MeSH descriptor: [Gastrointestinal Hemorrhage] explode all trees and with qualifier(s): [Drug therapy - DT, Prevention & control - PC, Therapy - TH]
#3 ((stress or stomach or peptic) near/2 ulcer)
#4 gastrointestinal bleeding
#5 MeSH descriptor: [Proton Pumps] explode all trees
#6 MeSH descriptor: [Proton Pump Inhibitors] explode all trees
#7 (PPI or PPIs or (proton near/3 pump near/3 inhibitor$))
#8 (dexlansoprazole or kapidex or dexilant)
#9 (esomeprazole or nexium or esotrex or alenia or escz or esofag or nexiam)
#10 (lansoprazole or lanzoprazole or agopton or bamalite or Inhibitol or Levant or Lupizole or lanzor or monolitum or ogast or ogastro or opiren or prevacid or prezal or pro ulco or promeco or takepron or ulpax or zoton)
#11 (omeprazole or losec or nexium or prilosec or rapinex or zegerid or ocid or Lomac or omepral or omez)
#12 (rabeprazole or aciphex or dexrabeprazole or pariet or zechin or rabecid or nzole-d or rabeloc)
#13 (pantoprazole or protium or protonix or pantotab or pantopan or pantozol or pantor or pantoloc or astropan or controloc or pantecta or inipomp or somac or pantodac or zurcal or zentro)
#14 MeSH descriptor: [Histamine H2 Antagonists] explode all trees
#15 ((h2 or histamine) near/2 (blocker$ or agonist$ or receptor$))
#16 (burimamide or cimetidine or famotidine or metiamide or nizatidine or ranitidine)
#17 (#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16)
#18 MeSH descriptor: [Critical Illness] explode all trees
#19 MeSH descriptor: [Critical Care] explode all trees
#20 (critical$ near/2 (ill$ or care))
#21 MeSH descriptor: [Intensive Care Units] explode all trees
#22 MeSH descriptor: [Respiration, Artificial] explode all trees
#23 (ICU or intensive care or high dependency unit$ or HDU or intermediate care or mechanical ventilation)
#24 MeSH descriptor: [Neurosurgery] explode all trees
#25 MeSH descriptor: [Brain] explode all trees and with qualifier(s): [Surgery - SU]
#26 MeSH descriptor: [Craniocerebral Trauma] explode all trees and with qualifier(s): [Surgery - SU]
#27 (neurosurgery or neurosurgical or brain surgery)
#28 MeSH descriptor: [Thorax] explode all trees
#29 MeSH descriptor: [Thoracic Surgery] explode all trees
#30 ((cardiothoracic or thorax or thoraracic or chest) near/2 (surgical or surgery or operation))
#31 MeSH descriptor: [Abdomen] explode all trees and with qualifier(s): [Surgery - SU]
#32 major abdominal surgery
#33 MeSH descriptor: [Vascular Diseases] explode all trees and with qualifier(s): [Surgery - SU]
#34 (vascular near/2 surgery)
#35 MeSH descriptor: [Pelvis] explode all trees and with qualifier(s): [Surgery - SU]
#36 ((pelvis or pelvic) near/2 surgery)
#37 MeSH descriptor: [Hip] explode all trees and with qualifier(s): [Surgery - SU]
#38 MeSH descriptor: [Arthroplasty, Replacement, Hip] explode all trees
#39 (hip near/2 (surgery or replacement or implantation$))
#40 MeSH descriptor: [Organ Transplantation] explode all trees
#41 ((organ or heart or heart-lung or kidney or liver or lung or pancreas) near/2 transplantation)
#42 MeSH descriptor: [Burns] explode all trees
#43 (burn injury or burn unit or thermal injury)
#44 MeSH descriptor: [Heart Arrest] explode all trees
#45 MeSH descriptor: [Myocardial Infarction] explode all trees
#46 (coronary care unit or CCU or cardiac intensive care unit CICU or cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct$)
#47 MeSH descriptor: [Hematologic Neoplasms] explode all trees
#48 ((hematologic$ or hematopoietic) near/2 (malignanc$ or neoplasm$ or illness))
#49 MeSH descriptor: [Acute Kidney Injury] explode all trees
#50 ((acute kidney or acute renal) near/2 (injur$ or failure or insufficienc$))
#51 MeSH descriptor: [Liver Failure] explode all trees
#52 ((hepatic or liver) near/2 failure)
#53 MeSH descriptor: [Sepsis] explode all trees
#54 sepsis
#55 MeSH descriptor: [Steroids] explode all trees and with qualifier(s): [Therapeutic use - TU]
#56 (steroid$ near/2 (treatment or therap$))
#57 (high near/2 (dose or dosis))
#58 (#55 or #56)
#59 (#57 and #58)
#60 (#18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39 or #40 or #41 or #42 or #43 or #44 or #45 or #46 or #47 or #48 or #49 or #50 or #51 or #52 or #53 or #54 or #59)
#61 (#17 and #60)

Medline (Ovid) 1946 to 11.10.2018

1. exp Peptic Ulcer/dt, pc, th [Drug Therapy, Prevention & Control, Therapy]
2. exp Gastrointestinal Hemorrhage/dt, pc, th [Drug Therapy, Prevention & Control, Therapy]
3. ((stress or stomach or peptic) adj2 ulcer).mp.
4. gastrointestinal bleeding.mp.
5. exp Proton Pumps/
6. exp Proton Pump Inhibitors/
7. (PPI or PPIs or (proton adj3 pump adj3 inhibitor$)).mp.
8. (dexlansoprazole or kapidex or dexilant).mp.
9. (esomeprazole or nexium or esotrex or alenia or escz or esofag or nexiam).mp.
10. (lansoprazole or lanzoprazole or agopton or bamalite or Inhibitol or Levant or Lupizole or lanzor or monolitum or ogast or ogastro or opiren or prevacid or prezal or pro ulco or promeco or takepron or ulpax or zoton).mp.
11. (omeprazole or losec or nexium or prilosec or rapinex or zegerid or ocid or Lomac or omepral or omez).mp.
12. (rabeprazole or aciphex or dexrabeprazole or pariet or zechin or rabecid or nzole-d or rabeloc).mp.
13. (pantoprazole or protium or protonix or pantotab or pantopan or pantozol or pantor or pantoloc or astropan or controloc or pantecta or inipomp or somac or pantodac or zurcal or zentro).mp.
14. exp Histamine H2 Antagonists/
15. ((h2 or histamine) adj2 (blocker$ or agonist$ or receptor$)).mp.
16. (burimamide or cimetidine or famotidine or metiamide or nizatidine or ranitidine).mp.
17. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16
18. exp Critical Illness/
19. exp Critical Care/
20. (critical$ adj2 (ill$ or care)).mp.
21. exp Intensive Care Units/
22. exp Respiration, Artificial/
23. (ICU or intensive care or high dependency unit$ or HDU or intermediate care or mechanical ventilation).mp.
24. exp Neurosurgery/
25. exp Brain/su [Surgery]
26. exp Craniocerebral Trauma/su [Surgery]
27. (neurosurgery or neurosurgical or brain surgery).mp.
28. exp Thorax/su [Surgery]
29. exp Thoracic Surgery/
30. ((cardiothoracic or thorax or thoraracic or chest) adj2 (surgical or surgery or operation)).mp.
31. exp Abdomen/su [Surgery]
32. major abdominal surgery.mp.
33. exp Vascular Diseases/su [Surgery]
34. (vascular adj2 surgery).mp.
35. exp Pelvis/su [Surgery]
36. ((pelvis or pelvic) adj2 surgery).mp.
37. exp Hip/su [Surgery]
38. exp Arthroplasty, Replacement, Hip/
39. (hip adj2 (surgery or replacement or implantation$)).mp.
40. exp Organ Transplantation/
41. ((organ or heart or heart-lung or kidney or liver or lung or pancreas) adj2 transplantation).mp.
42. exp Burns/
43. (burn injury or burn unit or thermal injury).mp.
44. exp Heart Arrest/
45. exp Myocardial Infarction/
46. (coronary care unit or CCU or cardiac intensive care unit CICU or cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct$).mp.
47. exp Hematologic neoplasms/
48. ((hematologic$ or hematopoietic) adj2 (malignanc$ or neoplasm$ or illness)).mp.
49. exp Acute kidney injury/
50. ((acute kidney or acute renal) adj2 (injur$ or failure or insufficienc$)).mp.
51. exp Liver failure/
52. ((hepatic or liver) adj2 failure).mp.
53. exp Sepsis/
54. sepsis.mp.
55. exp Steroids/dt, tu, th [Drug Therapy, Therapeutic Use, Therapy]
56. (steroid$ adj2 (treatment or therap$)).mp.
57. (high adj2 (dose or dosis)).mp.
58. 55 or 56
59. 57 and 58
60. (18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 59)
61. (17 and 60)
62. randomized controlled trial.pt.
63. controlled clinical trial.pt.
64. randomized.ab.
65. placebo.ab.
66. clinical trial.sh.
67. randomly.ab.
68. trial.ti.
69. 62 or 63 or 64 or 65 or 66 or 67 or 68
70. humans.sh.
71. 69 and 70
72. 61 and 71

Embase (Ovid) 1974 to 11.10.2018

1. exp peptic ulcer/dt, pc, th [Drug Therapy, Prevention, Therapy]
2. exp gastrointestinal hemorrhage/dt, pc, th [Drug Therapy, Prevention, Therapy]
3. ((stress or stomach or peptic) adj2 ulcer).mp.
4. gastrointestinal bleeding.mp.
5. exp proton pump/
6. exp proton pump inhibitor/
7. (PPI or PPIs or (proton adj3 pump adj3 inhibitor$)).mp.
8. (dexlansoprazole or kapidex or dexilant).mp.
9. (esomeprazole or nexium or esotrex or alenia or escz or esofag or nexiam).mp.
10. (lansoprazole or lanzoprazole or agopton or bamalite or Inhibitol or Levant or Lupizole or lanzor or monolitum or ogast or ogastro or opiren or prevacid or prezal or pro ulco or promeco or takepron or ulpax or zoton).mp.
11. (omeprazole or losec or nexium or prilosec or rapinex or zegerid or ocid or Lomac or omepral or omez).mp.
12. (rabeprazole or aciphex or dexrabeprazole or pariet or zechin or rabecid or nzole-d or rabeloc).mp.
13. (pantoprazole or protium or protonix or pantotab or pantopan or pantozol or pantor or pantoloc or astropan or controloc or pantecta or inipomp or somac or pantodac or zurcal or zentro).mp.
14. exp histamine H2 receptor antagonist/
15. ((h2 or histamine) adj2 (blocker$ or agonist$ or receptor$)).mp.
16. (burimamide or cimetidine or famotidine or metiamide or nizatidine or ranitidine).mp.
17. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16
18. exp critical illness/
19. exp intensive care/
20. (critical$ adj2 (ill$ or care)).mp.
21. exp intensive care unit/
22. exp artificial ventilation/
23. (ICU or intensive care or high dependency unit$ or HDU or intermediate care or mechanical ventilation).mp.
24. *neurosurgery/
25. brain/su [Surgery]
26. head injury/su [Surgery]
27. (neurosurgery or neurosurgical or brain surgery).mp.
28. thorax/su [Surgery]
29. *thorax surgery/
30. ((cardiothoracic or thorax or thoraracic or chest) adj2 (surgical or surgery or operation)).mp.
31. abdomen/su [Surgery]
32. *abdominal surgery/
33. vascular disease/su [Surgery]
34. (vascular adj2 surgery).mp.
35. pelvis/su [Surgery]
36. ((pelvis or pelvic) adj2 surgery).mp.
37. hip/su [Surgery]
38. *hip replacement/
39. (hip adj2 (surgery or replacement or implantation$)).mp.
40. *organ transplantation/
41. ((organ or heart or heart-lung or kidney or liver or lung or pancreas) adj2 transplantation).mp.
42. *burn/
43. (burn injury or burn unit or thermal injury).mp.
44. *heart arrest/
45. *heart infarction/
46. (coronary care unit or CCU or cardiac intensive care unit CICU or cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct$).mp.
47. *hematologic malignancy/
48. ((hematologic$ or hematopoietic) adj2 (malignanc$ or neoplasm$ or illness)).mp.
49. *acute kidney failure/
50. ((acute kidney or acute renal) adj2 (injur$ or failure or insufficienc$)).mp.
51. *liver failure/
52. ((hepatic or liver) adj2 failure).mp.
53. *sepsis/
54. sepsis.mp.
55. steroid/dt, th [Drug Therapy, Therapy]
56. (steroid$ adj2 (treatment or therap$)).mp.
57. (high adj2 (dose or dosis)).mp.
58. (55 or 56)
59. (57 and 58)
60. (18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 59)
61. (17 and 60)
62. CROSSOVER PROCEDURE.sh.
63. DOUBLE-BLIND PROCEDURE.sh.
64. SINGLE-BLIND PROCEDURE.sh.
65. (crossover* or cross over*).ti,ab.
66. placebo*.ti,ab.
67. (doubl* adj blind*).ti,ab.
68. allocat*.ti,ab.
69. trial.ti.
70. RANDOMIZED CONTROLLED TRIAL.sh.
71. random*.ti,ab.
72. (62 or 63 or 64 or 65 or 66 or 67 or 68 or 69 or 70 or 71)
73. (exp animal/ or exp invertebrate/ or animal.hw. or nonhuman/) not (exp human/ or human cell/ or (human or humans or man or

men or wom?n).ti.)
74. (72 not 73)
75. (61 and 74)

Science Citation Index – Expanded (web of science) 1900 to 11.10.2018

#32 (#31 AND #30)
#31 TS=(random* OR control* OR RCT OR placebo OR group* OR trial*)
#30 (#29 AND #12)
#29 (#28 OR #27 OR #26 OR #25 OR #24 OR #23 OR #22 OR #21 OR #20 OR #19 OR #18 OR #17 OR #16 OR #15 OR #14 OR #13)
#28 TS=(steroid$ NEAR (treatment OR therap$)) AND TS=(high NEAR (dose OR dosis))
#27 TS=(sepsis)
#26 TS=((hepatic OR liver) NEAR failure)
#25 TS=((acute) NEAR (kidney OR renal) NEAR (injur$ OR failure OR insufficienc$))
#24 TS=((hematologic$ OR hematopoietic) NEAR (malignanc$ OR neoplasm$ OR illness))
#23 TS=(coronary care unit OR CCU OR cardiac intensive care unit OR CICU OR cardiac arrest OR cardiac failure OR CPR OR heart arrest OR heart failure OR myocardial infarct$)
#22 TS=(burn injury OR burn unit OR thermal injury)
#21 TS=((organ OR heart OR heart-lung OR kidney OR liver OR lung OR pancreas) NEAR transplantation)
#20 TS=(hip NEAR (surgery OR replacement OR implantation$))
#19 TS=((pelvis or pelvic) NEAR surgery)
#18 TS=(vascular NEAR surgery)
#17 TS=(major abdominal surgery)
#16 TS=((cardiothoracic OR thorax OR thoraracic OR chest) NEAR (surgical OR surgery or operation))
#15 TS=(neurosurgery or neurosurgical or brain surgery)
#14 TS=(ICU or intensive care or high dependency unit$ or HDU or intermediate care or mechanical ventilation)
#13 TS=(critical$ NEAR (ill$ OR care))
#12 (#11 OR #10 OR #9 OR #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1)
#11 TS=(burimamide or cimetidine or famotidine or metiamide or nizatidine or ranitidine)
#10 TS=((h2 OR histamine) NEAR (blocker$ OR agonist$ OR receptor$))
#9 TS=(pantoprazole or protium or protonix or pantotab or pantopan or pantozol or pantor or pantoloc or astropan or controloc or pantecta or inipomp or somac or pantodac or zurcal or zentro)
#8 TS=(rabeprazole or aciphex or dexrabeprazole or pariet or zechin or rabecid or nzole-d or rabeloc)
#7 TS=(omeprazole or losec or nexium or prilosec or rapinex or zegerid or ocid or Lomac or omepral or omez)
#6 TS=(lansoprazole or lanzoprazole or agopton or bamalite or Inhibitol or Levant or Lupizole or lanzor or monolitum or ogast or ogastro or opiren or prevacid or prezal or pro ulco or promeco or takepron or ulpax or zoton)
#5 TS=(esomeprazole or nexium or esotrex or alenia or escz or esofag or nexiam)
#4 TS=(dexlansoprazole or kapidex or dexilant)
#3 TS=(PPI) OR TS=(PPIs) OR TS=(proton NEAR pump NEAR inhibitor$)
#2 TS=(gastrointestinal bleeding)
#1 TS=((stress OR stomach OR peptic) NEAR ulcer)

BIOSIS Previews (web of science) 1969 to 11.10.2018

#32 (#31 AND #30)
#31 TS=(random* OR control* OR RCT OR placebo OR group* OR trial*)
#30 (#29 AND #12)
#29 (#28 OR #27 OR #26 OR #25 OR #24 OR #23 OR #22 OR #21 OR #20 OR #19 OR #18 OR #17 OR #16 OR #15 OR #14 OR #13)
#28 TS=(steroid$ NEAR (treatment OR therap$)) AND TS=(high NEAR (dose OR dosis))
#27 TS=(sepsis)
#26 TS=((hepatic OR liver) NEAR failure)
#25 TS=((acute) NEAR (kidney OR renal) NEAR (injur$ OR failure OR insufficienc$))
#24 TS=((hematologic$ OR hematopoietic) NEAR (malignanc$ OR neoplasm$ OR illness))
#23 TS=(coronary care unit OR CCU OR cardiac intensive care unit OR CICU OR cardiac arrest OR cardiac failure OR CPR OR heart arrest OR heart failure OR myocardial infarct$)
#22 TS=(burn injury OR burn unit OR thermal injury)
#21 TS=((organ OR heart OR heart-lung OR kidney OR liver OR lung OR pancreas) NEAR transplantation)
#20 TS=(hip NEAR (surgery OR replacement OR implantation$))
#19 TS=((pelvis or pelvic) NEAR surgery)
#18 TS=(vascular NEAR surgery)
#17 TS=(major abdominal surgery)
#16 TS=((cardiothoracic OR thorax OR thoraracic OR chest) NEAR (surgical OR surgery or operation))
#15 TS=(neurosurgery or neurosurgical or brain surgery)
#14 TS=(ICU or intensive care or high dependency unit$ or HDU or intermediate care or mechanical ventilation)
#13 TS=(critical$ NEAR (ill$ OR care))
#12 (#11 OR #10 OR #9 OR #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1)
#11 TS=(burimamide or cimetidine or famotidine or metiamide or nizatidine or ranitidine)
#10 TS=((h2 OR histamine) NEAR (blocker$ OR agonist$ OR receptor$))
#9 TS=(pantoprazole or protium or protonix or pantotab or pantopan or pantozol or pantor or pantoloc or astropan or controloc or pantecta or inipomp or somac or pantodac or zurcal or zentro)
#8 TS=(rabeprazole or aciphex or dexrabeprazole or pariet or zechin or rabecid or nzole-d or rabeloc)
#7 TS=(omeprazole or losec or nexium or prilosec or rapinex or zegerid or ocid or Lomac or omepral or omez)
#6 TS=(lansoprazole or lanzoprazole or agopton or bamalite or Inhibitol or Levant or Lupizole or lanzor or monolitum or ogast or ogastro or opiren or prevacid or prezal or pro ulco or promeco or takepron or ulpax or zoton)
#5 TS=(esomeprazole or nexium or esotrex or alenia or escz or esofag or nexiam)

#4 TS=(dexlansoprazole or kapidex or dexilant)
#3 TS=(PPI) OR TS=(PPIs) OR TS=(proton NEAR pump NEAR inhibitor$)
#2 TS=(gastrointestinal bleeding)
#1 TS=((stress OR stomach OR peptic) NEAR ulcer)

PubMed 1966 to 11.10.2018

Search ((((((((((((((((((((Peptic Ulcer[MeSH Terms]) OR Gastrointestinal Hemorrhage[MeSH Terms]) OR (((stress or stomach or peptic) and ulcer))) OR gastrointestinal bleeding) OR Proton Pumps[MeSH Terms]) OR Proton Pump Inhibitors[MeSH Terms]) OR (((PPI or PPIs)) OR (proton and pump and inhibitor$))) OR ((dexlansoprazole or kapidex or dexilant))) OR ((esomeprazole or nexium or esotrex or alenia or escz or esofag or nexiam))) OR ((lansoprazole or lanzoprazole or agopton or bamalite or Inhibitol or Levant or Lupizole or lanzor or monolitum or ogast or ogastro or opiren or prevacid or prezal or pro ulco or promeco or takepron or ulpax or zoton))) OR ((omeprazole or losec or nexium or prilosec or rapinex or zegerid or ocid or Lomac or omepral or omez))) OR ((rabeprazole or aciphex or dexrabeprazole or pariet or zechin or rabecid or nzole-d or rabeloc))) OR ((pantoprazole or protium or protonix or pantotab or pantopan or pantozol or pantor or pantoloc or astropan or controloc or pantecta or inipomp or somac or pantodac or zurcal or zentro))) OR Histamine H2 Antagonists[MeSH Terms]) OR (((h2 or histamine)) AND (blocker$ or agonist$ or receptor$))) OR ((burimamide or cimetidine or famotidine or metiamide or nizatidine or ranitidine)))) AND ((Critical Illness[MeSH Terms]) OR Critical Care[MeSH Terms]) OR ((critical$) AND (ill$ or care))) OR Intensive Care Units[MeSH Terms]) OR Respiration, Artificial[MeSH Terms]) OR ((ICU or intensive care or high dependency unit$ or HDU or intermediate care or mechanical ventilation))) OR Neurosurgery[MeSH Terms]) OR ((Brain[MeSH Terms]) AND Surgery[MeSH Terms])) OR ((Craniocerebral Trauma[MeSH Terms]) AND Surgery[MeSH Terms])) OR ((neurosurgery or neurosurgical or brain surgery))) OR ((Thorax[MeSH Terms]) AND surgery[MeSH Terms])) OR Thoracic Surgery[MeSH Terms]) OR (((cardiothoracic or thorax or thoraracic or chest)) AND (surgical or surgery or operation))) OR ((Abdomen[MeSH Terms]) AND surgery[MeSH Terms])) OR major abdominal surgery) OR ((Vascular Diseases[MeSH Terms]) AND surgery[MeSH Terms])) OR ((vascular) AND surgery)) OR ((Pelvis[MeSH Terms]) AND surgery[MeSH Terms])) OR (((pelvis or pelvic)) AND surgery)) OR ((Hip[MeSH Terms]) AND surgery[MeSH Terms])) OR Arthroplasty, Replacement, Hip[MeSH Terms]) OR ((hip) AND (surgery or replacement or implantation$))) OR Organ Transplantation[MeSH Terms]) OR (((organ or heart or heart-lung or kidney or liver or lung or pancreas)) AND transplantation)) OR Burns[MeSH Terms]) OR ((burn injury or burn unit or thermal injury))) OR Heart Arrest[MeSH Terms]) OR Myocardial Infarction[MeSH Terms]) OR ((coronary care unit or CCU or cardiac intensive care unit CICU or cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct$))) OR Hematologic neoplasms[MeSH Terms]) OR (((hematologic$ or hematopoietic)) AND (malignanc$ or neoplasm$ or illness))) OR Acute kidney injury[MeSH Terms]) OR (((acute kidney or acute renal)) AND (injur$ or failure or insufficienc$))) OR Liver failure[MeSH Terms]) OR (((hepatic or liver)) AND failure)) OR Sepsis[MeSH Terms]) OR sepsis) OR ((((Steroids[MeSH Terms]) OR ((steroid$) AND (treatment or therap$)))) AND ((high) AND (dose or dosis)))))) AND (((((((((randomized controlled trial [pt]) OR controlled clinical trial [pt]) OR randomized [tiab]) OR placebo [tiab]) OR clinical trials as topic [mesh: noexp]) OR randomly [tiab]) OR trial [ti])) NOT ((animals [mh] NOT humans [mh])))

Trial registries and websites

We searched the following trial registers:

- ClinicalTrials.gov https://clinicaltrials.gov/
- WHO International Clinical Trials Registry Platform (ICTRP); http://apps.who.int/trialsearch/
- EU clinical trial register https://www.clinicaltrialsregister.eu/
- Australian New Zealand Clinical Trials Registry (ANZCTR) http://www.anzctr.org.au/
- AstraZeneca https://astrazenecagrouptrials.pharmacm.com/ST/Submission/Search
- Baver HealthCare https://clinicaltrials.bayer.com/
- Bristol-MyersSquibb https://www.bms.com/researchers-and-partners/clinical-trials-and-research.html
- GlaxoSmithKline https://www.gsk-clinicalstudyregister.com/
- Janssen https://globaltrialfinder.janssen.com/
- Lilly https://www.lillytrialguide.com/en-US
- Pfizer https://www.pfizer.com/science/find-a-trial/

In addition, we searched for unpublished trials on the websites of:

- U.S. Food and Drug Administration (FDA) https://fda.opentrials.net/search
- European Medicines Agency (EMA) https://www.ema.europa.eu/

Fig. S2. Data Collection Form

General					Interventions				Patient information					Randomisation and follow-up (best worst/worst best)							
Trial id	Year	Publ. Type	Protocol	OVERALL RoB Trial	Intervention	Control	Intervention period	Max follow-up	Type of ICU pts	Sub-pop 1 shock	Sub-pop 2 dialysis	Sub-pop 3 Mech vent	GI Bleeding (all or clin imp)	E: No random (PPI or H2RA)	C: No random (placebo/no proph)	Total randomised	*E: Lost to follow-up*	*C: Lost to follow-up*	E: No analysed	C: No analysed	Total analysed

All-cause mortality							Prop. of pts with CLIN IMP GI bleeding							Prop. of pts with ANY GI bleeding						
E: Number of events	E: Numbers analysed	C: Number of events	C: Numbers analysed	Blind pt+pers	Blind outcome assessor	Incomepl outcome data	E: Number of events	E: Numbers analysed	C: Number of events	C: Numbers analysed	Blind pt+pers	Blind outcome assessor	Incomepl outcome data	E: Number of events	E: Numbers analysed	C: Number of events	C: Numbers analysed	Blind pt+pers	Blind outcome assessor	Incomepl outcome data

Serious adverse events							Hospital-acquired pneumonia						Prop. of patients with Cl. Difficile						
E: Number of events	E: Numbers analysed	C: Number of events	C: Numbers analysed	Blind pt+pers	Blind outcome assessor	Incomepl outcome data	E: Numbers analysed	C: Number of events	C: Numbers analysed	Blind pt+pers	Blind outcome assessor	Incomepl outcome data	E: Number of events	E: Numbers analysed	C: Number of events	C: Numbers analysed	Blind pt+pers	Blind outcome assessor	Incomepl outcome data

Prop. of pts with myocardial ischemia							Quality of life							Comments
E: Number of events	E: Numbers analysed	C: Number of events	C: Numbers analysed	Blind pt+pers	Blind outcome assessor	Incomepl outcome data	E: Number of events	E: Numbers analysed	C: Number of events	C: Numbers analysed	Blind pt+pers	Blind outcome assessor	Incomepl outcome data	

Details on the assessment of risk of bias

We classified each trial and outcome result according to the domains below.

Random sequence generation

- Low risk: If sequence generation was achieved using computer, random number generator or a random numbers table. Drawing lots, tossing a coin, shuffling cards and throwing dice are also being considered adequate if performed by an independent adjudicator.
- Unclear risk: If the method of randomization was not specified.
- High risk: If the allocation sequence was not random.

Allocation sequence concealment

- Low risk: If the allocation of patients was performed by a central independent unit, on-site locked computer, identically looking numbered sealed opaque envelopes, drug bottles or containers prepared by an independent investigator. There must have been no risk of the investigator knowing the sequence.
- Unclear risk: If the trial was classified as randomized but the allocation concealment process is not described.
- High risk: If the allocation sequence was known to the investigators who assigned participants.

Blinding of participants and personnel

- Low risk: If the participants and the personnel were blinded to treatment allocation and this is described.
- Unclear risk: If the procedure of blinding was insufficiently described or not described at all.
- High risk: If blinding of participants and personnel was not performed.

Blinding of outcome assessment

- Low risk: If the trial investigators performing the outcome assessments, analyses and calculations were blinded to the intervention.
- Unclear risk: If the procedure of blinding was insufficiently described or not described at all.
- High risk: If blinding of outcome assessment was not performed.

Incomplete outcome data

- Low risk: (1) There are no dropouts or withdrawals for all outcomes, or (2) the numbers and reasons for the withdrawals and dropouts for all outcomes were clearly stated and could be described as being similar in both groups. As a general rule the trial was judged as at a low risk of bias due to incomplete outcome data if the number of dropouts was less than five per cent. However, the five per cent cut off was not definitive.
- Unclear risk: The numbers and reasons for withdrawals and dropouts were not clearly stated.
- High risk: The pattern of dropouts could be described as being different in the two intervention groups or the trial used improper methodology in dealing with the missing data, e.g., last observation carried forward.

Selective outcome reporting

- Low risk: A protocol was published, or a trial had been registered in a trial register (e.g. clinincaltrials.gov) before or at the time the trial is begun, and the outcome called for in the protocol or trial registration was reported on.
- Unclear risk: If there was no protocol and the outcome was not reported on.
- High risk: If the outcomes which are called on in a protocol were not reported on.

Other bias risk

- Low risk of bias: The trial appeared to be free of other components (for example, academic bias or for-profit bias) that could put it at risk of bias.
- Unclear risk of bias: The trial may or may not be free of other components that could have put it at risk of bias.
- High risk of bias: There were other factors in the trial that could put it at risk of bias (for example for-profit bias etc.)

Overall risk of bias

We classified all trials as:

- Overall low risk of bias: The trial was classified as overall 'low risk of bias' only if all of the bias domains described in the above paragraphs were classified as low risk of bias.
- Overall high risk of bias: The trial was classified 'high risk of bias' if any of the bias risk domains described in the above were classified as 'unclear' or 'high risk of bias'.

In addition, we assessed all domains for each outcome. Only ‘blinding of outcome assessment’, ‘incomplete outcome data’, and ‘selective outcome reporting’ varies between outcomes. Thus, we assessed the bias risk for each outcome assessed in addition to each trial.

Table S3. Characteristics of the included trials

Study	Year	Sample size	Setting/ICU		Intervention	Comparator	Duration of intervention	Maximum follow-up	Outcomes
Alhazzani [4]	2017	91	Mixed ICU	PPI	Pantoprazole 40 mg iv once daily	Placebo	Until GI bleeding, extubation or death in the ICU	Until hospital discharge	Mortality Clinically important bleeding Any GI bleeding Hospital-acquired pneumonia Cl. difficile enteritis
Apte [43]	1992	34	Medical ICU	H2RA	Ranitidine (50 mg. iv every 6 h)	No prophylaxis	Studied daily until 48 h after tracheal extubation	Studied daily until 48 h after tracheal extubation	Mortality Any GI bleeding Hospital-acquired pneumonia
Basso [44]	1981	116	Surgical ICU	H2RA	Cimetidine 200 mg every 6 h iv or orally	No prophylaxis	At least 10 days	10 days	Any bleeding
Benmenachem [45]	1994	200	Medical ICU	H2RA	Continuous iv Cimetidine titrated to maintain gastric ph at 4.0	No prophylaxis	Intervention was maintained until the occurrence of clinically severe hemorrhage, onset of drug-related complications, death, or discharge from medical ICU	NR	Mortality Clinically important bleeding Hospital-acquired pneumonia
Berg [46]	1985	34	Mixed ICU	H2RA	Cimetidine 20 mg/kg weight per 24 h	Placebo	At least 3 days, extubation or after 14 days in patients	NR	Any GI bleeding

							with prolonged assisted ventilation		
Burgess [47]	1995	34	Surgical ICU	H2RA	6.25 mg/h continuous intravenous ranitidine infusion	Placebo	72 h	NR	Mortality Any GI bleeding
Cartier [48]	1980	121	Not specified	H2RA	Cimetidine 1,2 g/day iv	Placebo	7 days	7 days	Any GI bleeding
Chan [49]	1995	101	Surgical ICU	H2RA	Ranitidine 6 mg every 6 h	Placebo	NR	6 months	Clinically important bleeding Hospital-acquired pneumonia Serious adverse events
Darlong [50]	2003	31	Mixed ICU	H2RA	Ranitidine 50 mg every 8 h	No prophylaxis	NR	NR	Any GI bleeding
Domingues [51]	1985	30	Mixed ICU	H2RA	Ranitidine, 50 mg every 6 h	No prophylaxis	NR	NR	
El-Kersh [17]	2018	124	Medical ICU	PPI	40 mg IV pantoprazole	Placebo	Until discharge from ICU or cessation of oral feeds	Until discharge from ICU or cessation of oral feeds	Mortality Any GI bleeding Clinically important bleeding Cl. difficile enteritis
Friedman [52]	1982	25	Medical ICU	H2RA	Cimetidine 300 mg iv q 6 h	Placebo	Until GI bleeding, weaning from ventilator, or death	Until GI bleeding, was weaned from ventilator, or died.	Any GI bleeding
Groll [53]	1986	221	Mixed ICU	H2RA	Cimetidine 300 mg	Placebo	Until GI bleeding, discharge from ICU or death	Until GI bleeding, discharge from ICU or death.	Mortality Any GI bleeding

Gundogan [54]	2017	158	Not specified	PPI	Pantoprazole (dose not specified)	No prophylaxis	NR	Until discharge from ICU or cessation of enteral nutrition up to four weeks	Any GI bleeding
Gursoy [55]	2008	75	Mixed ICU	PPI	Omeprazole 20 mg capsule, or pantoprazole 40 mg tablet, or esomeprazole 20 mg tablet, or rabeprazole 20 mg tablet	Placebo	One dose	ICU discharge	Mortality
Halloran [56]	1980	50	Surgical ICU	H2RA	300 mg cimetidine iv every 4 h	Placebo	Max 3 weeks	NR	Mortality Any GI bleeding
Hanish [57]	1998	114	Surgical ICU	H2RA	Ranitidine 3 x 50 mg iv	Placebo	NR	NR	Mortality Clinically important bleeding Hospital-acquired pneumonia
Hummer-Sigiel [58]	1986	22	Surgical ICU	H2RA	Ranitidine 0.2 mg.kg-1.h	Placebo	NR	NR	Any GI bleeding Hospital-acquired pneumonia
Jakob [59]	2005	43	Not specified	H2RA	Ranitidine 50 mg every 8 h	Placebo	24 h	ICU discharge	Mortality Any GI bleeding
Kam [60]	2011	80	Not specified	PPI	Not specified	No prophylaxis	NR	NR	Any GI bleeding
Kantotova H2RA [11]	2004	108	Surgical ICU	H2RA	Famotidine 40 mg	Placebo	NR	NR	Mortality

					twice a day at 12 h intervals by slow iv				Any GI bleeding Clinically important bleeding Hospital-acquired pneumonia Serious adverse events
Kantorova PPI [11]	2004	110	Surgical ICU	PPI	Omeprazole 40 mg iv once daily	Placebo	NR	NR	Mortality Any GI bleeding Clinically important bleeding Hospital-acquired pneumonia Serious adverse events
Karlstadt [61]	1990	87	Mixed ICU	H2RA	Initially 300 mg dose of cimetidine followed by a continuous infusion at a rate of 50 mg/h	Placebo	NR	NR	Mortality Clinically important bleeding Hospital-acquired pneumonia Serious adverse events
Koelz [62]	1987	67	Mixed ICU	H2RA	Ranitidine 50 mg iv every 8 h, or 25 mg in patients with a serum creatinine concentration exceeding 360 µmol/l.	Placebo	7 days	NR	Clinically important bleeding
Krag and Marker [3]	2018	3298	Mixed ICU	PPI	Pantoprazole 40 mg x 1 iv	Placebo	Until GI bleeding, ICU discharge or death	Max 90 days	Mortality Any GI bleeding Clinically important bleeding

									Hospital-acquired pneumonia Myocardial ischemia Cl. difficile enteritis
Larson [63]	1989	31	Surgical ICU	H2RA	Continuous infusion of iv ranitidine (6.25 mg/h, 150 mg/day)	Placebo	3 days	3 days	Any GI bleeding
Lin [16]	2016	120	Mixed ICU	PPI	Lanzoprazole OD 30 mg once daily	No prophylaxis	14 days	14 days GI bleeding and 30 days mortality	Mortality Any GI bleeding Clinically important bleeding Hospital-acquired pneumonia
Liu H2RA [15]	2013	80	Surgical ICU	H2RA	Cimetidine 300 mg iv every 6 h	Placebo	7 days	30 days	Mortality Any GI bleeding Hospital-acquired pneumonia
Liu PPI [15]	2013	85	Surgical ICU	PPI	Omeprazole 40 mg iv every 12 h	Placebo	7 days	30 days	Mortality Any GI bleeding Hospital-acquired pneumonia
Luk [64]	1982	123	Not specified	H2RA	Cimetidine 300 mg iv for 6 h	Placebo	NR	NR	Any GI bleeding
MacDougall [65]	1977	62	Medical ICU	H2RA	Metiamide or cimetidine 150 mg iv	No prophylaxis	Until recovery or mortality	Hospital discharge	Mortality Any GI bleeding
Martin [66]	1993	131	Mixed ICU	H2RA	Cimetidine loading	Placebo	7 days	30 days	Mortality

					dose 300 mg and then 50 mg/h				Any GI bleeding Hospital-acquired pneumonia
Metz [67]	1993	167	Surgical ICU	H2RA	Ranitidine 6.25 mg/h continuous infusion	Placebo	5 days	NR	Any GI bleeding Hospital-acquired pneumonia
Nielsen [68]	1989	25	Mixed ICU	H2RA	Ranitidine 50 mg every 6 h	No prophylaxis	8 days	90 days	Mortality
Peura [69]	1985	39	Medical ICU	H2RA	Cimetidine 300 mg iv every 6 h	Placebo	14 days	NR	Mortality Clinically important bleeding
Powell H2RA [70]	1993	16	Surgical ICU	H2RA	Ranitidine 50 mg iv every 8 h	Placebo	6 days	6 days	Mortality Any GI bleeding
Powell PPI [70]	1993	25	Surgical ICU	PPI	Omeprazole 80 mg iv loading dose then 40 mg every 8 h by iv bolus or infusion	Placebo	NR	NR	Mortality Any GI bleeding
Rigaud [71]	1988	12	Medical ICU	H2RA	Ranitidine 0.25µg/kg/h	Placebo	3 days	3 days	
Rohde [72]	1980	28	Surgical ICU	H2RA	Cimetidine (dose not specified)	Placebo	NR	NR	Mortality Any GI bleeding
Ruiz-Santana [73]	1991	49	Mixed ICU	H2RA	Ranitidine 50 mg iv every 6 h	No prophylaxis	Until weaning, GI bleeding, or death	ICU discharge	Mortality Any GI bleeding
Selvanderan [5]	2016	216	Mixed ICU	PPI	Pantoprazole 40 mg iv daily	Placebo	Max 14 days	Max 21 days, Cl. difficile until hospital discharge, mortality 90 days.	Mortality Any GI bleeding Clinically important bleeding Hospital-acquired

									pneumonia Cl. difficile enteritis
Spapen [74]	1995	30	Mixed ICU	H2RA	Cimetidine 1200 mg or ranitidine 200 mg iv during study period	Placebo	NR	7 days or until discharge	Mortality
Vlatten [75]	1998	60	Mixed ICU	PPI	Omeprazole 2 x 40 mg	No prophylaxis	NR	NR	Any GI bleeding
Zinner [76]	1981	226	Surgical ICU	H2RA	300 milligram iv every 6 h	No prophylaxis	During entire stay in the ICU	NR	Mortality Any GI bleeding Clinically important bleeding

C.: Clostridium; GI: gastrointestinal; h: hour; H2RA: histamin-2 receptor antagonist; ICU: intensive care unit; iv: intravenously; NR: not reported; PPI: proton pump inhibitor

Table S4. Details of included trials and risk of bias assessment

Alhazzani 2017		
Methods	Randomised clinical trial	
Participants	Sample size: 91 (experimental 49; placebo 42) Sex: females: 44.9% in exp group; 40.5% in control group Age: median 59,4 Country: Canada, Saudi Arabia, Australia Setting: mixed ICU Inclusion criteria: 1) adults (≥ 18 yr) who were admitted to the ICU; 2) were anticipated to receive invasive mechanical ventilation for greater than or equal to 48 hours. Exclusion criteria: 1) invasive mechanical ventilation for greater than or equal to 72 hours prior to randomisation; 2) the use of PPIs due to active bleeding or increased risk of bleeding; 3) the use of dual antiplatelet therapy prior to randomisation; 4) palliative care or decision to withdraw advanced life support; 5) previous enrollment in this or a related trial; 6) pregnancy; 7) ICU physician, patient, or substitute decision maker (SDM) declined trial participation; and 8) receipt of greater than or equal to two "daily dose equivalents" of prophylaxis with H2RA or PPI in the current ICU admission.	
Interventions	Experimental: pantoprazole 40 mg in 0.9% NaCl 50 mL (0.8 mg/mL) iv once daily Control: placebo (0.9% NaCl, 50 mL) iv once daily Co-intervention: not reported Duration: until GI bleeding, extubation or death in the ICU The study drugs were administered iv while patients were mechanically ventilated until GI bleeding or death in the ICU	
Outcomes	• Mortality • Clinically important bleeding (defined as the presence of overt GI bleeding (i.e., hematemesis, frank blood or coffee ground nasogastric aspirate, melena, or hematochezia) plus one of these features in the absence of other causes: a spontaneous drop of systolic or diastolic blood pressure of greater than or equal to 20 mm Hg within 24 hours of upper GI bleeding, an orthostatic increase in pulse rate of greater than or equal to 20 beats/min and a decrease in systolic blood pressure of greater than or equal to 10 mm Hg, a decrease in hemoglobin of greater than or equal to 2 g/dL (20 g/L) in 24 hours or transfusion of greater than or equal to two units of packed RBCs within 24 hours of bleeding) • Any GI bleeding (definition not specified) • Hospital-acquired pneumonia • *Cl. difficile*	
Notes		
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Web-based system.
Allocation concealment (selection bias)	Low risk	Randomisation was concealed.
Blinding of participants and personnel (performance bias)	Low risk	A drug-matched placebo was used.
Blinding of outcome assessment (detection bias)	Low risk	Data analyst were blinded to allocation for the trial duration and the adjudication and analysis period.
Incomplete outcome data (attrition bias)	Low risk	No lost to follow up.
Selective reporting (reporting bias)	Low risk	Study protocol was published after patient enrolment was initiated. However, the trial was registered on clinicaltrials.gov (NCT02290327) prior to randomisation. 'GI bleeding requiring invasive intervention' and 'any GI bleeding' were reported on in trial report but was not pre-specified in either published protocol or on clinical.trial.gov. Furthermore, 'length of ICU and hospital stay' and 'length of invasive mechanical ventilation' was not pre-specified on clinicaltrials.gov.
Other bias	Low risk	The trial was funded by public grants.
Apte 1992		
Methods	Randomised clinical trial	
Participants	Sample size: 34 (experimental 16; control 18) Sex (M/F): experimental: 12/4; control: 11/7. Age (yr): experimental: 27, control: 26 Country: India Setting: medical ICU Inclusion criteria: tracheotomized patients who were admitted to the medical ICU with tetanus Exclusion criteria: patients who had pneumonia before tracheostomy or who had received ranitidine before randomisation were excluded	
Interventions	Experimental: iv ranitidine (50 mg. 6 hourly) Control: no prophylaxis	

	Co-intervention: intermittent nasogastric feeding (300-400 mL, 4 hourly) Duration: until 48 hours after tracheal extubation	
Outcomes	• Mortality • Any GI bleeding (definition not specified) • Hospital-acquired pneumonia	
Notes	Contact information was not identified, thus e-mail was not sent.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	The method of random sequence was not described.
Allocation concealment (selection bias)	Unclear risk	Allocation concealment not described.
Blinding of participants and personnel (performance bias)	High risk	Unblinded.
Blinding of outcome assessment (detection bias)	High risk	Unblinded.
Incomplete outcome data (attrition bias)	Low risk	No patients were lost to follow up.
Selective reporting (reporting bias)	Unclear risk	Protocol was not published and trial was not registered, no lost to follow up.
Other bias	Unclear risk	Torrent Pharmaceuticals provided the ranitidine.
Basso 1981		
Methods	Randomised clinical trial	
Participants	Sample size: 116 (168 in total in three groups, antacid group excluded). 60 in cimetidine group and 56 in no treatment group. Sex: not reported Age: not reported Country: Italy Setting: population: surgical ICU (high risk plastic- and neurosurgery) Inclusion criteria: patients in high risk of GI bleeding. Exclusion criteria: patients were excluded from the study if they had any evidence of gross upper gastrointestinal bleeding before or during the 12 hours after the onset of the study, if they had had a gastric or oesophageal operation, if they were aged 12 years or less, or if they presented with coagulopathy.	
Interventions	Experimental: cimetidine 200 mg every 6 hours iv or orally. Control: no treatment Co-intervention: not reported Multi-arm study: this was a three arm trial. Cimetidine vs antacid (Maalox 10 ml/hours by nasogastric tube or orally) vs no treatment. We excluded the antacid group from our analysis.	
Outcomes	• Any bleeding (definition not specified)	
Notes	16 died before end of trial intervention. However total death is not reported, and in which group the deaths appeared was not reported. Number of patients analysed is assessed from looking at Figure 1, as this number is not reported in trial report. E-mail was sent to Dr. Basso 25.07.18. E-mail re-sent to Dr. Basso 27.08.18.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	List with randomised values
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	High risk	Single blinded
Blinding of outcome assessment (detection bias)	Low risk	Outcome assessor was blinded to treatment groups.
Incomplete outcome data (attrition bias)	High risk	The number of dropouts in each group is not presented - only total (31). It was unclear whether these 31 patients were analysed as intention to treat.
Selective reporting (reporting bias)	Unclear risk	There was no pre-published protocol
Other bias	Low risk	The trial was supported by public grants
Benmenachem 1994		
Methods	Randomised clinical trial	
Participants	Sample size: 200 (100 in each group) Sex: males 51% in all three groups Age (mean): control 59,6; experimental 59,0 Country: US Setting: medical intensive care unit Inclusion criteria: admission to medical ICU and above 18 years of age.	

	Exclusion criteria: 1) expected stay of 24 hours or less; 2) evidence of GI bleeding at the time of admission to ICI; 3) treatment with antacids, H2RA, sucralfate or omeprazole during the 24 hours before entering ICU; 4) use of NSAIDs, systemic anti-coagulant, or thrombolytic agents during the 7 previous days; 5) surgery requiring general anaesthesia during the previous 2 weeks; 6) closed head injury or clinical evidence of increased intracranial pressure; 7) grade 4 hepatic encephalopathy; 8) oesophageal or gastric surgery in the previous year; 9) history of GI bleeding during the previous year; 10) pregnancy or lactation.
Interventions	Experimental: continuous iv cimetidine titrated to maintain gastric pH at 4.0. Control: no prophylaxis Co-intervention: not reported Multi-arm study: this was a three arm trial. The third arm included 100 patients who received 1 g. sucralfate orally every 6 hours (excluded from our analyses) Duration: intervention was maintained until the occurrence of clinically severe haemorrhage, onset of drug-related complications, death, or discharge from medical ICU.
Outcomes	• Mortality • Clinically important bleeding (defined as substantial gastrointestinal haemorrhage requiring the presence of any of the following: I) persistent hematemesis (red blood or guaiac-positive "coffee grounds" that did not clear with 1.5 saline lavage; 2) 3-point decrease in haematocrit during 24 hours accompanied by red blood or guaiac-positive "coffee grounds," material that cleared with lavage, or melena, or three guaiac positive stools without evidence of lower gastrointestinal bleeding; 3) any unexplained 6-point decrease in haematocrit during 48-hour period) • Hospital-acquired pneumonia
Notes	E-mail was sent to Dr. Benmenachem and Dr. Bresalier 25.07.18 and reply was received.

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Permuted block randomisation was used.
Allocation concealment (selection bias)	Low risk	Sealed envelopes
Blinding of participants and personnel (performance bias)	High risk	The trial was described as blinded, but control group received no intervention.
Blinding of outcome assessment (detection bias)	Low risk	Investigators were blinded to therapy.
Incomplete outcome data (attrition bias)	Low risk	No patient was lost to follow up.
Selective reporting (reporting bias)	High risk	Protocol was not pre-registered or published (clarification received by e-mail). Recurring haemorrhage and total transfusion requirement was not reported, as stated in the methods section.
Other bias	Low risk	This study was supported by public grants.

Berg 1985

Methods	Randomised clinical trial
Participants	Sample size: 34 (experimental 17; control 17), only 28 analysed Sex: males 50% Age (mean): experimental 43,9, control 48,4 Country: the Netherlands Setting: medical and surgical ICU patients on mechanical ventilation Inclusion criteria: medical and surgical ICU patients on mechanical ventilation Exclusion criteria: patients with previous oesophageal or gastric operations and patients with upper GI bleeding on admission were excluded from the study, as were patients in whom a period of assisted ventilation of less than 3 days was expected.
Interventions	Experimental: cimetidine 20 mg/kg weight per 24 hours Control: placebo (normal saline) Duration: treatment was given for at least 3 days and ended on extubation of the patient or after 14 days in patients with prolonged assisted ventilation. Co-intervention: not reported
Outcomes	• Any GI bleeding (definition not specified)
Notes	E-mail was sent to Dr. van Blankenstein 25.07.18. E-mail re-sent 27.08.18.

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	Low risk	The trial was described as double blinded, and placebo was used in the control group.
Blinding of outcome assessment (detection bias)	Unclear risk	The trial was described as double blinded, but it was unclear how blinding was maintained.
Incomplete outcome data	High risk	6 patients dropped out and were not included in ITT analysis

(attrition bias)		
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Unclear risk	It was unclear how the trial was funded.
Burgess 1995		
Methods	Randomised clinical trial	
Participants	Sample size: 34 (experimental 16; control 18) Sex: experimental male 69%; control male 78% Age (mean): experimental 38,4; control 34,5 Country: US Setting: surgical ICU (adults with severe head injury and a Glasgow coma scale (GCS) (13) score <10) Inclusion criteria: adults with severe head injury and a Glasgow coma scale (GCS) (13) score <10 admitted to the University of Louisville surgical intensive care unit (SICU) Exclusion criteria: patients with concomitant peptic ulcer disease, other gastrointestinal injury, receiving anti-ulcer therapy, or having any oral intake were excluded	
Interventions	Experimental: 6.25 mg/hours continuous intravenous ranitidine infusion Control: saline placebo infusion Co-intervention: not reported Duration: the treatment period was complete when the patient was withdrawn from the study or had received 72 hours of study drug	
Outcomes	• Mortality • Any GI bleeding (defined as 5% decrease from baseline in haematocrit occurring at least 8 hr after study drug initiation plus any of these signs: hematemesis, haematochezia, bright red blood per NG tube or "coffee ground" NG tube aspirates)	
Notes	E-mail was sent to Dr. Burgess 25.07.18. E-mail re-sent 27.08.18.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated randomisation scheme
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	Low risk	Placebo was used.
Blinding of outcome assessment (detection bias)	Unclear risk	It is stated that the principal investigator (PI) had no access to the pH data. It is unclear whether PI was blinded to other outcomes.
Incomplete outcome data (attrition bias)	Low risk	No patients were lost to follow up
Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	High risk	The trial was supported by Glaxo Inc. Research Institute
Cartier 1980		
Methods	Randomised clinical trial	
Participants	Sample size: 121 (experimental 58; control 63). Sex: not reported Age: not reported Country: France Setting: not specified Inclusion criteria: ICU patients (critically ill patients having risk factors for GI bleeding) Exclusion criteria: not specified	
Interventions	Experimental: cimetidine 1,2 g/day iv Control: placebo	
Outcomes	• Any GI bleeding (definition not specified)	
Notes	Error in text; it is written that 61 patients were randomised to the placebo group. However, according to Table, 9 had a bleeding and 54 did not have a bleeding (n=63). In our analysis we noted that 9 bleeds out of 63 patients in the placebo group. Contact information was not identified; thus, e-mail was not sent.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	Low risk	The trial was described as double blinded using a placebo.
Blinding of outcome assessment (detection	Unclear risk	Not described.

bias)		
Incomplete outcome data (attrition bias)	Low risk	No patients were lost to follow-up.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Unclear risk	It was unclear how the trial was funded.

Chan 1995	
Methods	Randomised clinical trial
Participants	Sample size: 101 (experimental 52; control 49) Sex: experimental male 53%; control male 54% Mean age: experimental 61 (17-84); control 61 (32-89) Country: Hong Kong, China Setting: intervention administered preoperatively to patients having non-traumatic neurosurgical lesions (before emergency neurosurgery). Population: high-risk neurosurgical patients. Inclusion criteria: non-traumatic neurosurgical lesions with two or more risk factors underwent operations Exclusion criteria: 1) failure to obtain consent (four patients); 2) presence of gastroduodenal bleeding before neurosurgery (10 patients); 3) past history of chronic gastroduodenal diseases or chronic ulcers, identified at endoscopy (nine patients); and 4) concomitant major medical illnesses such as heart, lung, kidney, haematological, and liver problems (seven patients).
Interventions	Experimental: ranitidine 6 mg every 6 hour Control: placebo (normal saline) Co-intervention: concomitant medications included dexamethasone, 4 mg every 6 hours, and a single dose of ceftriaxone (1 g), which was given intravenously as prophylaxis with the first dose of ranitidine or placebo. Subsequent antibiotic medications were administered only for treatment of culture - proven infections. Those patients who required anticonvulsant therapy or prophylaxis received phenytoin (100 mg every 8 hours). Duration: The medications were administered intravenously every 6 hours and were started on call to the operating theatre and continued into the postoperative period. Twice daily doses of oral ranitidine (150 mg) or placebo were commenced when patients were considered ready for enteric feeding.
Outcomes	• Clinically important bleeding (defined as bleeding requiring blood transfusion and/or surgery) • Hospital-acquired pneumonia • Serious adverse events
Notes	E-mail was sent to Dr. Chan and Dr. Yu WC 25.07.18. E-mail re-sent 27.08.18.

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	Low risk	A drug-matched placebo was used.
Blinding of outcome assessment (detection bias)	Unclear risk	At 6 months follow-up outcomes were assessed by an independent observer. It was not mentioned who assessed the other outcomes.
Incomplete outcome data (attrition bias)	Low risk	No patients were lost to follow up.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Low risk	The study was supported by public grants.

Darlong 2004	
Methods	Randomised clinical trial
Participants	Sample size: 31 (experimental 24; control 7) Sex ratio (male:female): experimental 11:13; control 3:4 Mean age: experimental: 43,95 +- 18,46; control: 39,16 +- 19,52 Country: India Setting: general (mixed) ICU Inclusion criteria: patients who had been mechanically ventilated and who were likely to last for more than 24 hours. Exclusion criteria: patients with active upper GI haemorrhage who had received antacids, H2 receptor antagonists or sucralfate in the previous 24 hours were excluded. Patients on anticoagulants or those with coagulopathy ware also excluded.
Interventions	Experimental: ranitidine 50 mg every 8 hours Control: no prophylaxis Co-intervention: not specified Duration: 3 days Multi-arm study: 3 arm trial. 52 randomised. ranitidine: 24 patients, control: 7, sucralfate: 21 patients. We only included results on ranitidine and control.
Outcomes	• Any GI bleeding (defined as observation of fresh blood or blood of coffee ground colour in the gastric

	aspirates)
Notes	E-mail was sent to Dr. Darlong and Dr. Tandon 25.07.18. E-mail re-sent 27.08.18.

Risk of bias assessment

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Insufficient information given (patients were allocated at random to three groups).
Allocation concealment (selection bias)	Unclear risk	Insufficient information given.
Blinding of participants and personnel (performance bias)	High risk	The trial was unblinded.
Blinding of outcome assessment (detection bias)	Low risk	Gastroenterologist who performed the endoscopic evaluation was blinded to the randomisation.
Incomplete outcome data (attrition bias)	Low risk	No patients were lost to follow-up.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Unclear risk	It was unclear how the trial was funded.

Domingues 1985

Methods	Randomised clinical trial
Participants	Sample size: 30 (experimental 15; control 15) Sex: experimental: 8 males + 7 females; control: 8 males + 7 females Mean age: experimental group 50; control 53 Country: Brazil Setting: mixed ICU Inclusion criteria: not reported Exclusion criteria: not reported
Interventions	Experimental: ranitidine 50 mg every 6 hours Control: no prophylaxis intervention Duration: not specified
Outcomes	• No relevant outcomes reported
Notes	Contact information was not identified; thus, e-mail was not sent.

Risk of bias assessment

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Authors stated that intervention groups were randomly formed, but the method of random sequence was not described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	High risk	Non-blinded.
Blinding of outcome assessment (detection bias)	Low risk	Endoscopist was blinded to intervention groups.
Incomplete outcome data (attrition bias)	Low risk	No patients were lost to follow up.
Selective reporting (reporting bias)	Unclear risk	There was no pre-published protocol.
Other bias	Unclear risk	It was unclear how the trail was funded.

El-Kersh 2018

Methods	Randomised clinical trial
Participants	Sample size: 124 (experimental 62; control 62) Sex males: experimental 55%; control 60% Age: experimental 62 (49,5-68); control 58 (40,5-66,5) Country: US Setting: medical ICU. Population: Mechanically ventilated ICU patients. Inclusion criteria: patients who were 18 years or older who were expected to need mechanical ventilation for N48 hours with no contraindications to EN within the first 24 hours after admission to the ICU. Exclusion criteria: exclusion criteria included 1) evidence of GI bleeding during the hospitalization period prior to study enrolment, 2) admission to the ICU with primary diagnosis of burn injury, 3) closed head injury or increased intracranial pressure, 4) history of partial or complete gastrectomy, and 5) pregnancy, or lactation.
Interventions	Experimental: 40 mg iv pantoprazole. Control: placebo Co-intervention: early enteral nutrition. Duration: until discharge from ICU or cessation of oral feeds
Outcomes	• Mortality • Clinically important bleeding (defined by a 3-point decrease in haematocrit within a 24-hour period with clinical signs of overt GI bleeding, or by an unexplained 6-point decrease in haematocrit in a 48-hour

	period) • Any GI bleeding (defined as overt bleeding by the presence of coffee-ground aspirate in nasogastric tube or coffee-ground emesis, bloody secretions in nasogastric tube or hematemesis, melena or haematochezia) • Cl. difficile	
Notes	E-mail was sent to Dr. El-Kersh 25.07.18 and reply was received. Data on myocardial ischemia and quality of life was not collected.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer-generated random numbers
Allocation concealment (selection bias)	Low risk	Sealed opaque envelopes
Blinding of participants and personnel (performance bias)	Low risk	A drug-matched placebo (made at the pharmacy) was used.
Blinding of outcome assessment (detection bias)	Low risk	Reserach personnel were blinded to treatment assignment.
Incomplete outcome data (attrition bias)	High risk	Number of patients excluded after randomisation are high and unequal; 11% in experimental group and 24% in control group. Trial authors have not not used ITT analysis.
Selective reporting (reporting bias)	Low risk	All outcomes stated in the protocol (NCT01477320) were reported or argued why they were not reported (cost analysis)
Other bias	Low risk	The trial was supported by Abbott Nutrition. They had no active role in the design, methodology, data collections, analysis, preparation of the manuscript or decision to submit the manuscript for publication.
Friedman 1982		
Methods	Randomised clinical trial	
Participants	Sample size: 25 (experimental 11, control 14) Sex: not specified Age: not specified Country: US Setting: medical ICU. Population: Mechanically ventilated patients. Inclusion criteria: patients who had been receiving mechanical ventilation for less than 12 hours were eligible to enter the study. Exclusion criteria: patients were excluded from the study if they had renal insufficiency (creatinine > 3 mg/dl), active gastrointestinal bleeding at the time of initiation of ventilatory support, had received antacids and/or cimetidine immediately before ventilation, or were pregnant.	
Interventions	Experimental: cimetidine 300 mg iv q 6 h Control: placebo Co-intervention: not specified Duration: interventions were continued until a patient developed GI bleeding, was weaned from ventilation or died. Multi-arm study: the trial included a third arm: antacid (Mylanta II)	
Outcomes	• Any GI bleeding (overt upper GI bleeding was defined as the presence of fresh or old blood in the nasogastric aspirate which failed to clear with saline lavage in 15 min, or as melena. Occult GI bleeding was defined as a drop in the haematocrit of 5 or more points, associated with positive tests for stool occult blood for 3 consecutive days without obvious non-upper GI bleeding)	
Notes	Contact information was not identified; thus, e-mail was not sent	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised but the method of sequence generation was not described.
Allocation concealment (selection bias)	Unclear risk	Insufficient information given
Blinding of participants and personnel (performance bias)	Low risk	A drug-matched placebo was used.
Blinding of outcome assessment (detection bias)	Unclear risk	Not described.
Incomplete outcome data (attrition bias)	Low risk	Numbers and reasons for withdrawals are stated and can be described as being similar in both groups.
Selective reporting (reporting bias)	Unclear risk	There was no pre-published protocol.
Other bias	Unclear risk	Study drug was provided by Smith Kline & French Laboratories, but it was unclear whether they had any active role in the design, methodology, data collections, analysis, preparation of the manuscript or decision to submit the manuscript for publication.
Groll 1986		

Methods	Randomised clinical trial	
Participants	Sample size: 221 (experimental 114; placebo 107) Sex: male/female ratio: experimental 75/39; control 68/39 Age mean: experimental 58 (16-90); control 57 (15-88) Country: Canada Setting: mixed ICU Inclusion criteria: mixed ICU patients Exclusion criteria: bleeding on admission to the ICU, pregnancy, renal failure requiring haemodialysis or peritoneal dialysis, drug overdosage, acute myocardial infarction, use of antacids, stay in the unit less than 24 hours	
Interventions	Experimental: cimetidine 300 mg in 20 ml normal Control: placebo Co-intervention: not specified Duration: the trial was terminated when either bleeding occured or the patient was discharged from the unit.	
Outcomes	• Mortality • Any GI bleeding (defined as (i) frank haematemesis or gastric aspirate of >50 ml fresh blood, (ii) melaena or fresh blood per rectum with an upper source of haemorrhage verified by endoscopy if the gastric aspirate was clear, (iii) a fall in haemoglobin level >2 g/dl in a 24 hour period associated with either 4+ occult blood in the stools or coffee ground gastric drainage of at least 100 ml)	
Notes	E-mail was sent to Dr. Groll and Dr. Depew 25.07.18. E-mail re-sent 27.08.18.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	It was stated that the trial was randomised, but the method of sequence of generation was not described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	Low risk	A drug matched placebo was used.
Blinding of outcome assessment (detection bias)	Unclear risk	It was not mentioned if the outcome assessor were blinded.
Incomplete outcome data (attrition bias)	Low risk	No patients dropped out or were lost to follow-up.
Selective reporting (reporting bias)	Unclear risk	There was no pre-published protocol.
Other bias	Unclear risk	The trial was supported by Smith Kline and French Canada Ltd, but it was unclear whether they had any active role in the design, methodology, data collections, analysis, preparation of the manuscript or decision to submit the manuscript for publication.
Gundogan 2017		
Methods	Randomised clinical trial	
Participants	Sample size: 158 (experimental 80; control 78) Sex: not specified Age: 61.3 +- 17.6 Country: Turkey Setting: medical ICU. Inclusion criteria: critically ill patients. Exclusion criteria: not specified.	
Interventions	Experimental: pantoprazole (dosis not specified) and oral/enteral nutrition Control: only oral/enteral nutrition Co-intervention: not specified Duration: 3 days	
Outcomes	• Any GI bleeding (definition not specified)	
Notes	E-mail was sent to Dr. Kurat Gundogan and Dr. Murat Sungur 25.07.18. E-mail re-sent 27.08.18. Reply received 29/08/19, however, only with minor clarifications. 14% of all included patients died and numbers for each group were not provided.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but method of random sequence generation was mot described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	High risk	The trial is stated as an open-label trial on clinicaltrials.gov.
Blinding of outcome assessment (detection bias)	High risk	The trial is described as open label.
Incomplete outcome data	Low risk	No patients were lost to follow up (clarification received by e-mail).

(attrition bias)		
Selective reporting (reporting bias)	High risk	The protocol (NCT03098537) was published retrospectively, and not all outcomes specified in protocol are reported in abstract.
Other bias	Unclear risk	Funding by industry (Nestle), but it was unclear whether they had any active role in the design, methodology, data collections, analysis, preparation of the manuscript or decision to submit the manuscript for publication.
Gursoy 2008		
Methods	Randomised clinical trial	
Participants	Sample size: 75 (experimental 60; control 15) Sex (F/M): Group C 9/6, Group O 8/7, Group P 6/9, Group E 8/7, Group R 7/8 Age (y) Group C: 58.00 ± 16.20, Group O: 58.00 ± 16.05, Group P: 57.67 ± 21, Group E: 59.07 ± 17.98, Group R: 54.67 ± 19.89 Country: Turkey Setting: mixed ICU. Population: Adult trauma, general surgical and medical patients requiring mechanical ventilatory support. Inclusion criteria: requirement for nasogastric intubation and arterial catheter, and (ii) haemodynamic stability prior to the study, defined as: no transient hypotension episodes (arterial systolic blood pressure <100 mmHg) for 24 hours and no changes in pulse rate (± 10 beats/min). Exclusion criteria: an expected time of nasogastric intubation of <24 hours, gastrointestinal bleeding or recent surgery to the upper gastrointestinal tract.	
Interventions	Experimental: 60 patients who were divided into groups that received the following treatments: Group O (n = 15), omeprazole 20 mg capsule (Erbolin, Biofarma, Istanbul, Turkey); Group P (n = 15), pantoprazole 40 mg tablet (Panto, Ilsan, Kocaeli, Turkey); Group E (n = 15) esomeprazole 20 mg tablet (Nexium®, AstraZeneca, Istanbul, Turkey); Group R (n = 15) rabeprazole 20 mg tablet (Pariet®, Johnson and Johnson, Istanbul, Turkey). Control: saline 100 mL. Co-intervention: not specified. Duration: one dose.	
Outcomes	• Mortality	
Notes	We pooled the four PPI groups into one group. E-mail was sent to Dr. Dilek Memis and Dr. Sut 25.07.18. E-mail re-sent 27.08.18.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	Low risk	The drug was infused to all patients by a nurse who had no knowledge of the study protocol.
Blinding of outcome assessment (detection bias)	Unclear risk	Not described.
Incomplete outcome data (attrition bias)	Low risk	No patients dropped out or were lost to follow up.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Low risk	The trial appeared to be free of other components that could put it at risk of bias.
Halloran 1980		
Methods	Randomised clinical trial	
Participants	Sample size: 50 (experimental 26; control 24) Sex (male:female): experimental 23:3; control 18:6 Age: experimental 29.6 (15-54); control 30.6 (8-62) Country: US Setting: surgical ICU. Population: patients with severe head injury. Inclusion criteria: the minimal neurologic criterion for entry into this study was the inability of the patient to obey simple commands after closed head injury. Exclusion criteria: apnoeic patients with fixed dilated pupils and no motor response to painful stimuli on arrival were excluded. Additionally, patients were excluded if they were known to have peptic ulcer disease, were pregnant, had concomitant injury of the upper gastrointestinal tract or had severe hepatic or renal disease.	
Interventions	Experimental: 300 mg cimetidine iv every 4 hours Control: placebo (content not specified) Duration: 3 weeks Co-intervention: steroids (dexamethasone or methylprednisolone) and prophylactic anticonvulsant therapy. Muscle relaxants were used to facilitate artificial respiratory support" assessed patients to be mechanical ventilated.	
Outcomes	• Mortality • Any GI bleeding (defined as bright red blood or a 4 plus positive guaiac in the gastric aspirate for three consecutive 8 hour periods (exclusive of the 1st day after injury) if no oropharyngeal source of bleeding was present)	
Notes	GI bleeding is reported (separately) on both mild to moderate (overt) and those were transfusion were needed	

	(clinically important) Contact information was not identified; thus, e-mail was not sent.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described.
Allocation concealment (selection bias)	Unclear risk	Insufficient information given.
Blinding of participants and personnel (performance bias)	Low risk	A drug-matched placebo was used.
Blinding of outcome assessment (detection bias)	Low risk	Coded medication was stopped at the discretion of the physician responsible for the patients' care when bleeding occurred.
Incomplete outcome data (attrition bias)	Low risk	No patient was lost to follow up.
Selective reporting (reporting bias)	Unclear risk	No protocol was identified.
Other bias	Unclear risk	The trial was supported by Smith, Kline and French Laboratories, but it was unclear whether the company had any active role in the design, methodology, data collections, analysis, preparation of the manuscript or decision to submit the manuscript for publication.
Hanisch 1998		
Methods	Randomised clinical trial	
Participants	Sample size: 114 (experimental 57; control 57) Sex: not specified. Age: experimental 55 (22-88); control: 58 (22-88) Country: Germany Setting: surgical ICU Inclusion criteria: all patients referred to the intensive care unit of the surgical department of the Johann Wolfgang Goethe-University Frankfurt/Main were considered for the study Exclusion criteria: exclusion criteria were patients with an active peptic ulcer disease and a concomitant ulcer medication; patients with upper gastrointestinal bleeding; patients ,18 years; transplanted patients (kidney, liver, heart); and patients with pre-existing pneumonia and gastric resection.	
Interventions	Experimental: ranitidine 3 x 50 mg iv Control: placebo Duration: not specified Multi-arm study: the trial included a third arm, pirenzepine (44 patients were included in this arm) and excluded from our study.	
Outcomes	• Mortality • Clinically important bleeding (defined as bright red blood via gastric tube or melena combined with hemodynamic changes (systolic blood pressure <100 mm Hg, tachycardia>100 beats per minute) • and requirement of blood transfusion (fall in haemoglobin > 2 g/dL within 24 hours) and endoscopic identification of bleeding site and activity) • Hospital-acquired pneumonia	
Notes	E-mail was sent to Dr. Windolf 25.07.18. E-mail re-sent 27.08.18.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	A complete and balanced randomisation schedule was generated by the institute by the Institute of Biomethematics of the University of Frankfurt.
Allocation concealment (selection bias)	Low risk	At the time of entering the ICU patients were assigned to a consecutive study number, and the application of the blinded drug regimen was started.
Blinding of participants and personnel (performance bias)	Low risk	A drug-matched placebo was used.
Blinding of outcome assessment (detection bias)	Low risk	Data documentation for each patient was performed by different staff not involved in the treatment of patients and in the randomised assignment of drugs.
Incomplete outcome data (attrition bias)	Low risk	No patients were lost to follow up.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Unclear risk	It was unclear how the trial was funded.
Hummer Siegel 1986		
Methods	Randomised clinical trial	
Participants	Sample size: 22 (experimental 11; control 11) Sex (M/F): experimental 7/4; control 10/1 Age (mean): experimental 23; control 26 Country: France	

	Setting: surgical ICU. Population: Severe traumatic brain injury Inclusion criteria: severe traumatic brain injury, under sedation for hypometabolizing cerebral purpose. Exclusion criteria: patients who have undergone surgery on the upper digestive tract or whose condition requires emergency surgery on the upper digestive tract; patients who present a digestive ulceration. and / or gastrointestinal bleeding at the entry into the trial as well as lesions related to the presence of a digestive probe; patients with significant blood-crushed disorders, as well as those who have received antacid therapy, salicylic acid or derivatives, high dose corticosteroids since the initial accident of cimetidine.	
Interventions	Experimental: ranitidine 0.2 $mg.kg^{-1}.h$ Control: placebo Co-intervention: not specified Duration: not specified.	
Outcomes	• Any GI bleeding (definition not specified) • Hospital-acquired pneumonia	
Notes	Contact information was not identified; thus, e-mail was not sent.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described.
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	Low risk	A drug matched placebo was used.
Blinding of outcome assessment (detection bias)	Unclear risk	It was not mentioned whether the outcome assessors were blinded.
Incomplete outcome data (attrition bias)	Low risk	No patient was lost to follow-up.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	Unclear risk	It was unclear how the trial was funded.
Jakob 2005		
Methods	Randomised clinical trial	
Participants	Sample size: 46 (experimental 24; control 22) Sex (M/F): experimental 13/7; control 11/9 Age: experimental 62+-16; Control: 62+-15 Country: Finland Setting: medical ICU. Population: Mechanically ventilated patients. Inclusion criteria: (1) emergency admission due to acute circulatory or respiratory failure, and (2) requirement for a naso-gastric tube. Exclusion criteria: (1) expected time on mechanical ventilation less than 24 h, (2) gastrointestinal bleeding, (3) recent surgery of the upper gastrointestinal tract, (4) treatment limitation because of bad prognosis, and (5) age less than 18 years.	
Interventions	Experimental: ranitidine 50mg every 8 hours for 24 hours. Control: placebo Duration: 24 hours	
Outcomes	• Mortality • Any GI bleeding (definition not specified)	
Notes	E-mail was sent to Dr. Stephan 25.07.18 and reply was received. E-mail was sent to Dr. Parviainen 26.07.18 and reply was received 15.08.18.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Block randomisation.
Allocation concealment (selection bias)	High risk	Numbered sealed envelopes. However, block sizes are identical, which gives a high risk of bias in allocation concealment.
Blinding of participants and personnel (performance bias)	Low risk	A drug-matched placebo was used.
Blinding of outcome assessment (detection bias)	Low risk	Outcome assessor was blinded (clarification received by e-mail).
Incomplete outcome data (attrition bias)	High risk	6 patients were not included in analysis. 3 of these patients received the intervention, but these data were not included in the analysis.
Selective reporting (reporting bias)	High risk	Protocol was not pre-registered or published (clarification received by e-mail).
Other bias	Low risk	The trial was supported in part by Intrumentarium Corp, who had no active role in the design, methodology, data collections, analysis, preparation of the manuscript or decision to submit the manuscript for publication (clarified by e-mail).

Kam 2011		
Methods	Randomised clinical trial	
Participants	Sample size: 80 (experimental 45; control 35) Sex: not reported Age: not reported Country: not reported Setting: not specified. Population: ICU patients receiving enteral feeding and free of known GI bleeding. Inclusion criteria: enteral feeding. Exclusion criteria: known GI bleeding.	
Interventions	Experimental: "standard practices of SUP in the ICU using PPI (not specified) Control: no intervention	
Outcomes	• Any GI bleeding (defined as a drop of more than 2 gm/dL of haemoglobin along with overt bleeding)	
Notes	Information on number of events in each group were not specified in the conference abstract. Number of events used in our analyses has been calculated based on the information given in the abstract. E-mail was sent to Dr. Kam 25.07.18. E-mail re-sent 27.08.18. No reply was received.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	High risk	Un-blinded.
Blinding of outcome assessment (detection bias)	High risk	Un-blinded.
Incomplete outcome data (attrition bias)	Unclear risk	Not enough information in abstract to assess incomplete outcome data.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Unclear risk	It was unclear how the trial was funded.
Kantorova (H2RA) 2004		
Methods	Randomised clinical trial	
Participants	Sample size: 108 (experimental 71; control 37 (75 in total) Sex (male:female): experimental 62%; control: 67% Age (mean): experimental 47; control 46 Country: Czech Republic Population: mixed ICU Inclusion criteria: polytraumatized patients and patients with absolved major intraabdominal or intrathoracic surgery admitted to any ICUs. All patients 18 years or older who were projected to require mechanical ventilation for at least 48 hours or had coagulopathy and had a nasogastric tube in place. Exclusion criteria: exclusion criteria were 1) expected stay in ICU 48 hours or Less, 2) esophagogastric surgery including vagotomy in patients history, 3) evidence of gastrointestinal bleeding at the time of admission to the ICU and during the previous year, 4) pneumonia, 5) treatment with PPI, 1--12 blockers, antacids, or sucralfate during the previous 72 hours, 6) documented peptic ulcer disease during the last year, 7) use of systemic anticoagulants, high-dose oral corticosteroids or thrombolytic agents during the previous week, 8) renal insufficiency requiring haemodialysis, 9) thrombocytopenia 000/mL, 10) patients with life expectancy <3 months, 11) patient was not able or willing to give informed consent.	
Interventions	Experimental: famotidine 40 mg twice a day at 12h intervals by slow iv Control: placebo Duration: not specified Co-intervention: not specified Multi-arm trial: The trial also included a Sucralfate group (n=69) which we have excluded. We have divided the control group in two according to the two intervention groups.	
Outcomes	• Mortality • Any GI bleeding (defined as hematemesis, melena, positive nasogastric aspirate, or hematochezia) • Clinically important bleeding (defined as overt bleeding plus at least one of the following: 1) drop of systolic blood pressure 220 mm Hg or increase in the pulse rate of 220 beats/min within 24 hours after upper gastrointestinal bleeding, or 2) decrease in the haemoglobin concentration 22g/dL) • Hospital-acquired pneumonia • Serious adverse events	
Notes	E-mail was sent to Dr. Svoboda 25.07.18, although not delivered. Alternative e-mail addresses could not be found.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated random numbers
Allocation concealment	Low risk	Opaque sealed numbered envelopes

(selection bias)		
Blinding of participants and personnel (performance bias)	Low risk	A drug-matched placebo was used
Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	High risk	36 (11%) patients (not specified by group) were excluded from the analyses due to 1 patient died (within 2 hours after randomisation), 18 underwent mech vent for under 48h, and 16 were not assessed due to missing important data (reason not clearly stated).
Selective reporting (reporting bias)	Unclear risk	There was no pre-published protocol.
Other bias	Low risk	The study was supported by public grants.
Kantorova (PPI) 2004		
Methods	Sample size: 110 (experimental 72; control 38 (75 in total) Sex (male:female): experimental male 67%; control 67% Age (mean): experimental 44; control 46 Country: Czech Republic Population: mixed ICU Inclusion criteria: polytraumatized patients and patients with absolved major intraabdominal or intrathoracic surgery admitted to any ICUs. All patients 18 years or older who were projected to require mechanical ventilation for at least 48 hours or had coagulopathy and had a nasogastric tube in place. Exclusion criteria: exclusion criteria were 1) expected stay in ICU 48 hours or less, 2) esophagogastric surgery including vagotomy in patients history, 3) evidence of gastrointestinal bleeding at the time of admission to the ICU and during the previous year, 4) pneumonia, 5) treatment with PPI, H2 blockers, antacids, or sucralfate during the previous 72 hours, 6) documented peptic ulcer disease during the last year, 7) use of systemic anticoagulants, high-dose oral corticosteroids or thrombolytic agents during the previous week, 8) renal insufficiency requiring haemodialysis, 9) thrombocytopenia <30 000/mL, 10) patients with life expectancy <3 months, 11) patient was not able or willing to give informed consent.	
Participants	Experimental: PPI (omeprazole): 72 (40 mg iv once daily); H2 antagonists (famotidine): 71 (40mg twice a day at 12h intervals by slow iv) Control: placebo: 75 in total (in the analysis, the placebo group is divided according to the two intervention groups) (iv saline) Duration: not specified Co-intervention: not specified	
Interventions	Experimental: omeprazole 40 mg iv once daily Control: placebo Duration: not specified Co-intervention: not specified	
Outcomes	• Mortality • Any GI bleeding (defined as hematemesis, melena, positive nasogastric aspirate, or hematochezia) • Clinically important bleeding (defined as overt bleeding plus at least one of the following: 1) drop of systolic blood pressure 220 mm Hg or increase in the pulse rate of 220 beats/min within 24 hours after upper gastrointestinal bleeding, or 2) decrease in the haemoglobin concentration 22g/dL) • Hospital-acquired pneumonia • Serious adverse events	
Notes	E-mail was sent to Dr. Svoboda 25.07.18, although not delivered. Alternative e-mail addresses could not be found.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated random numbers.
Allocation concealment (selection bias)	Low risk	Opaque sealed numbered envelopes.
Blinding of participants and personnel (performance bias)	Low risk	A drug-matched placebo was used.
Blinding of outcome assessment (detection bias)	Unclear risk	Not described.
Incomplete outcome data (attrition bias)	High risk	36 (11%) patients (not specified by group) were excluded from the analyses due to 1 patient died (within 2 hours after randomizatrion), 18 underwent mech vent for under 48h, and 16 were not assessed due to missing important data (reason not clearly stated).
Selective reporting (reporting bias)	Unclear risk	There was no pre-published protocol.
Other bias	Low risk	The study was supported by public grants.
Karlstadt 1990		
Methods	Randomised clinical trial	
Participants	Sample size: 87 (experimental 54; control 33)	

	Sex (male:female): experimental 31:23; control 16:17 Age (mean): experimental 56,5; control 61,9 Country: USA Population: Mixed ICU Inclusion criteria: patients admitted to ICUs were eligible for entry into the trial if they had at least one of the following conditions generally regarded as risk factors for bleeding: (1) major thoracic or abdominal surgery; (2) major multiple trauma; (3) hypotension, defined as a decrease in blood pressure greater than 30/20 mm Hg (systolic/diastolic); (4) hypovolemic shock, defined as a syndrome of inadequate tissue perfusion characterized by systolic blood pressure of less than 90 mm Hg (or a decrease of 30 mm Hg in previously hypertensive patients); (5) sepsis, defined by the presence of peritonitis, confirmed bacteraemia, or the complex of fever, elevated white blood cell count, and hypotension with a bacteriologically determined source of infection; and (6) acute respiratory failure, defined as the need for assisted ventilation for more than 24 hours. Exclusion criteria: 1) active upper gastrointestinal bleeding, a history of peptic ulcer, or upper gastrointestinal surgery; (2) severe chronic hepatic failure, defined by the presence of portal systemic encephalopathy or ascites secondary to chronic liver disease; (3) renal failure, defined by elevated serum creatinine, indicating a creatinine clearance of less than 30 mg/min; (4) treatment with other drugs, such as antacids, other H,-receptor antagonists, and sucralfate, that would interfere with evaluation of the investigative treatment effects; (5) pregnancy or lactation; and (6) age less than 16 years. Patients with known hypersecretory disorders (e.g., peptic ulcer, burns) or who are considered likely to bleed from non-stress-related conditions (e.g., varices, uremic gastritis) were not included to focus the trial on bleeding from more typical stress-related mucosal damage.
Interventions	Experimental: patients received an initial 300-mg dose of cimetidine infused over 15 to 20 minutes, followed by a continuous infusion by IVAC or IMED (San Diego, CA) pump of cimetidine at a rate of 50 mg/hours. Control: placebo Duration: not specified Co-intervention: not specified
Outcomes	• Mortality • Clinically important bleeding (defined as (1) hematemesis or the presence of more than 10 ml of bright red blood in a single aspirate; (2) melena or haematochezia (unless upper gastrointestinal endoscopy clearly indicated that the melena did not arise from an upper gastrointestinal site); (3) the presence of "coffee grounds," positive for haemoglobin by Gastroccult (SmithKline Diagnostics, Sunnyvale, CA), in the nasogastric aspirate on each of 3 consecutive 6hourly observations (over 12 hours) and a I-gm decrease in haemoglobin over 24 hours; or (4) Gastroccult-positive "coffee grounds" in aspirate that did not clear with lavage) • Hospital-acquired pneumonia • Serious adverse events
Notes	Contact information was not identified; thus, e-mail was not sent.

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	Low risk	The trial was described as "double-blind" using placebo.
Blinding of outcome assessment (detection bias)	Unclear risk	Not described.
Incomplete outcome data (attrition bias)	Low risk	2 patients were withdrawn from the cimetidine group, and reasons stated.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	High risk	Funding by industry, Smith Kline & French Laboratories. First author works at Smith Kline & French Laboratories.

Koelz 1990

Methods	Randomised clinical trial
Participants	Sample size: 67 (experimental 33; control 34) Sex (male:female): experimental 19:10; control 19:8 Age (mean): experimental 37,1; control 38,1 Country: Switzerland Setting: mixed ICU. Inclusion criteria: patients of either sex aged 16 years and more admitted to the intensive care unit because of sepsis and/or polytrauma. Exclusion criteria: patients already receiving treatment with H2-receptor antagonists, those with evidence of upper gastrointestinal bleeding at the beginning of the study, those with basal skull fracture, and those with facial injury that precluded upper gastrointestinal endoscopy or could represent an unacceptable risk to the patient.
Interventions	Experimental group: ranitidine 50 mg iv every 8 h, or 25 mg in patients with a serum creatinine concentration exceeding 360 µmol/l. Control: placebo Duration: 7 days Co-intervention: 10 ml antacid every 4 h.

Outcomes	• Clinically important bleeding (defined as endoscopically visible lesions thought to be responsible for the need for transfusion of at least two units of blood or for surgery because of bleeding)	
Notes	3/56 patients died. Mortality not specified by group, thus these data are not included in our analysis. E-mail was sent to Dr. Koelz and Dr. Halter 25.07.18, although not delivered. Alternative e-mail addresses could not be found.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	Low risk	Drug-matched placebo was used.
Blinding of outcome assessment (detection bias)	Unclear risk	Not described.
Incomplete outcome data (attrition bias)	High risk	16% of the included patients dropped out. 21% in the control group and 12% in the Experimental group. Two patients dropped out and not included in the analysis due to being dead. However, reasons for dropouts are given in the paper.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Unclear risk	It was unclear how the trial was funded.
Krag and Marker 2018		
Methods	Randomised clinical trial	
Participants	Sample size: 3298 (experimental 1645; control 1653) Sex (male no): experimental 1039 (63%); control 1067 (65%) Age: experimental 67; control 67 Country: Denmark (patients recruited multinational) Setting: mixed ICU Inclusion criteria: all adult (18 years or older) patients who are acutely admitted to the ICU with one or more of the following risk factors for gastrointestinal bleeding: · Shock (continuous infusion with vasopressors or inotropes, systolic blood pressure below 90 mmHg, mean arterial blood pressure below 70 mmHg or plasma lactate level 4 mmol/l or above) · Acute or chronic intermittent or continuous renal replacement therapy (RRT) · Invasive mechanical ventilation which is expected to last more than 24 hours · Coagulopathy (platelets below 50 × 10^9/l, or international normalized ratio (INR) above 1.5, or prothrombin time (PT) above 20 s) documented within the last 24 hours · Ongoing treatment with anticoagulant drugs (prophylactic doses excluded) · History of coagulopathy (platelets below 50 × 10^9/l or INR above 1.5 or PT above 20 s within the 6 months prior to hospital admission) · History of chronic liver disease (portal hypertension, cirrhosis proven by biopsy, computed tomography (CT) scan or ultrasound or history of variceal bleeding or hepatic encephalopathy) Exclusion criteria: · Contraindications to proton pump inhibitors: any history of intolerance to proton pump inhibitors or additives, or treatment with atazanavir (HIV medication) · Ongoing treatment with proton pump inhibitors and/or histamine-2-receptor antagonists on a daily basis. Ongoing is defined as treatment not being discontinued at ICU admission · Gastrointestinal bleeding of any origin (both upper and lower gastrointestinal bleeding) during current hospital admission, documented in the patient charts · Diagnosed with peptic ulcer confirmed by endoscopy or other method during current hospital admission · Organ transplant during current hospital admission · Withdrawal from active therapy or brain death documented in the patient charts · Fertile woman with positive urine human chorionic gonadotropin (hCG) or plasma-hCG · Consent according to national regulations not obtainable: patients where the clinician or investigator is unable to obtain necessary consent before inclusion of the patient according to the national regulations	
Interventions	Experimental: pantoprazole 40 mg x 1 iv Control: placebo x 1 iv Duration: until GI bleeding, ICU discharge, (death)	
Outcomes	• Mortality • Any GI bleeding (overt GI bleeding defined as hematemesis, coffee ground emesis, melena, haematochezia or bloody nasogastric aspirate) • Clinically important bleeding (defined as overt gastrointestinal bleeding and at least one of the following four features within 24 hours of gastrointestinal bleeding, in the absence of other causes, in the ICU: a spontaneous decrease in systolic blood pressure, mean arterial pressure, or diastolic blood pressure of 20 mm Hg or more; initiation of treatment with a vasopressor or a 20% increase in vasopressor dose; a decrease in haemoglobin of at least 2 g per decilitre [1.24 mmol per liter]; or transfusion of two or more units of packed red cells) • Hospital-acquired pneumonia • Myocardial ischemia • Cl. Difficile	
Notes	Sanam Safi and Kiran Kumar Katakam, who were not involved in any aspects of the trial, extracted data and	

	evaluated risk of bias of the SUP-ICU trial.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomisation was performed with a centralised, computer-generated allocation sequence. Patients were randomly allocated in a 1:1 ratio, with the use of permuted blocks of varying sizes to pantoprazole or placebo.
Allocation concealment (selection bias)	Low risk	Randomisation was performed with a centralised, computer-generated allocation sequence.
Blinding of participants and personnel (performance bias)	Low risk	A drug-matched placebo was used.
Blinding of outcome assessment (detection bias)	Low risk	Patients, personnel, investigators and statisticians were blinded to allocation.
Incomplete outcome data (attrition bias)	Low risk	16/3298 patients were lost to follow-up (for 90-day mortality). Reasons for exclusion and group is specified.
Selective reporting (reporting bias)	Low risk	Consistency with published protocol.
Other bias	Low risk	The trial was funded by public grants. By Innovation Fund Denmark.
Larson 1989		
Methods	Randomised clinical trial	
Participants	Sample size: 31 (experimental 13; control 18) Sex (male:female): not specified Age: not specified Country: USA Setting: surgical ICU. Population: Patients with severe head injury and multiple trauma Inclusion criteria: not specified Exclusion criteria: ot specified	
Interventions	Experimental: continuous infusion of iv ranitidine (6.25 mg/h, 150 mg/day) Control: placebo	
Outcomes	• Any GI bleeding (definition not specified)	
Notes	7 patients were lost to follow-up of which one died, 1 removed pH electrode and 5 developed GI bleeding. For the outcome GI bleeding, two patients were lost to follow-up. These patients were not specified to randomised group. These data can therefore not be included in BW/WB analysis. Contact information was not identified; thus, e-mail was not sent.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	Low risk	The trial was described as double-blind using placebo.
Blinding of outcome assessment (detection bias)	Unclear risk	The trial was described as double-blind, but it was unclear who was blinded and how blinding was maintained.
Incomplete outcome data (attrition bias)	High risk	For the outcome GI bleeding two patients were lost to follow-up. These two patients were not specified by intervention group.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Unclear risk	It was unclear how the trial was funded.
Lin 2016		
Methods	Randomised clinical trial	
Participants	Sample size: 120 (experimental 60; control 60) Sex (male:female): experimental 38/22; control 37/23 Age: experimental 67,7; control 64,8 Country: Taiwan Setting: mixed ICU. Population: Mechanically ventilated patients Inclusion criteria: patients who had received mechanical ventilation for >48 hours, had undergone nasogastric (NG) tube intubation, and were prepared to be weaned from the ventilator were included. Exclusion criteria: patients who were pregnant, <18 years old, allergic to lansoprazole, having active UGI bleeding, or receiving PPIs or H2RAs within 1 week were excluded.	
Interventions	Experimental: lanzoprazole OD 30 mg once daily Control: no intervention Duration: 14 days	
Outcomes	• Mortality • Any GI bleeding (defined as (1) a coffee ground substance from the NG aspirate >60 mL; (2) fresh blood from the NG tube; or (3) passage of tarry stool) • Clinically important bleeding (defined as GI with haemoglobin level decrease ≥ 2 gm/dL or in need of a	

	blood transfusion of >2 units) • Hospital-acquired pneumonia	
Notes	E-mail was sent to Dr. Lin and Dr. Lee 26.07.18. E-mail re-sent 27.08.18.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that patients were randomly allocated into two groups using block randomisation, but the method of sequence generation was nit described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	High risk	The trial was described as non-double blinded.
Blinding of outcome assessment (detection bias)	Unclear risk	Not described.
Incomplete outcome data (attrition bias)	Low risk	No patients were lost to follow up.
Selective reporting (reporting bias)	Low risk	All outcomes stated in the protocol were assessed (ClinicalTrials.gov ID: NCT00708149).
Other bias	Low risk	The trial was supported by public grants.
Liu (H2RA) 2013		
Methods	Randomised clinical trial	
Participants	Sample size: 80 (experimental 54; control 26 (54 in total)) Sex (male:female): experimental 34:20; control 35:18 Age: experimental <40: 13, 40–60: 33, >60: 8; control: <40: 10, 40–60: 40, >60: 3 Country: China Setting: surgical ICU. Population: Neurosurgical patients. Inclusion criteria: older than 18 years, had CT-proven ICH within 72 hours of ictus requiring neurosurgery, had a nasogastric tube in place and a baseline intragastric pH lower than 4 on 2 consecutive determinations, and if they or their legally authorized representative gave informed consent Exclusion criteria: patients with arteriovenous malformation or aneurysmal hemorrhage, those who had a history of peptic ulcers, those who were likely to swallow blood (for example, those with severe facial trauma), those who underwent antiplatelet and anticoagulation therapy, those with renal insufficiency requiring hemodialysis, those with thrombocytopenia less than 30,000/ml, and those who died within 72 hours after the ictus.	
Interventions	Experimental: cimetidine 300 mg iv every 6 h Control: placebo (placebo solution every 12 hours) Duration: 7 days Multi-arm trial: the trial compared three arms: omeprazole, cimetidine and placebo. We have included both arms in our study.	
Outcomes	• Mortality • Any GI bleeding (as defined by hematemesis, aspiration of coffee ground material from the nasogastric tube, or melena, which was proven by positive results of gastric occult blood testing or faecal occult blood testing, with or without hemodynamic instability resulting from gross bleeding that needed transfusion) • Hospital-acquired pneumonia	
Notes	E-mail was sent to Dr. Bing Li 26.07.18. E-mail re-sent 27.08.18.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated random numbers.
Allocation concealment (selection bias)	Unclear risk	Not specified.
Blinding of participants and personnel (performance bias)	Unclear risk	iv PPI and iv H2RA not given at the same administration intervals.
Blinding of outcome assessment (detection bias)	Unclear risk	Not described.
Incomplete outcome data (attrition bias)	High risk	19 (10.3%) patients were lost to follow up. The 19 patients were not specified by group, data from this trial could therefore not be included in the sensitivity analyses on losses to follow-up. "19 were excluded from data analysis because 11 were lost to follow-up within 30 days of ictus, 5 were not assessable due to missing important data, and 3 did not meet the enrolment criteria."
Selective reporting (reporting bias)	High risk	All outcomes stated in the protocol were assessed (ChiCTR-TRC-12001871). However, the protocol was registered retrospectively.
Other bias	Unclear risk	It was unclear how the trial was funded.
Liu (PPI) 2013		

Methods	Randomised clinical trial	
Participants	Sample size: 85 (experimental 58; control 27 (54 in total)) Sex (male:female): experimental 31:27; control 35:18 Age: experimental <40: 10, 40–60: 44, >60: 4; control: <40: 10, 40–60: 40, >60: 3 Country: China Setting: surgical ICU. Population: Neurosurgical patients. Inclusion criteria: older than 18 years, had CT-proven ICH within 72 hours of ictus requiring neurosurgery, had a nasogastric tube in place and a baseline intragastric pH lower than 4 on 2 consecutive determinations, and if they or their legally authorized representative gave informed consent Exclusion criteria: patients with arteriovenous malformation or aneurysmal hemorrhage, those who had a history of peptic ulcers, those who were likely to swallow blood (for example, those with severe facial trauma), those who underwent antiplatelet and anticoagulation therapy, those with renal insufficiency requiring hemodialysis, those with thrombocytopenia less than 30,000/ml, and those who died within 72 hours after the ictus.	
Interventions	Experimental: omeprazole 40 mg iv every 12 hours Control: placebo Duration: 7 days Multi-arm trial: the trial compared three arms: omeprazole, cimetidine and placebo. We have included both arms in our study.	
Outcomes	• Mortality • Any GI bleeding (as defined by hematemesis, aspiration of coffee ground material from the nasogastric tube, or melena, which was proven by positive results of gastric occult blood testing or fecal occult blood testing, with or without hemodynamic instability resulting from gross bleeding that needed transfusion) • Hospital-acquired pneumonia	
Notes	E-mail was sent to Dr. Bing Li 26.07.18. E-mail re-sent 27.08.18.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated random numbers.
Allocation concealment (selection bias)	Unclear risk	Not specified.
Blinding of participants and personnel (performance bias)	Unclear risk	iv PPI and iv H2RA not given at the same administration intervals.
Blinding of outcome assessment (detection bias)	Unclear risk	Not described.
Incomplete outcome data (attrition bias)	Unclear risk	19 (10.3%) patients were lost to follow up. The 19 patients were not specified by group, data from this trial could therefore not be included in the BW WB analyses. "19 were excluded from data analysis because 11 were lost to follow-up within 30 days of ictus, 5 were not assessable due to missing important data, and 3 did not meet the enrolment criteria."
Selective reporting (reporting bias)	High risk	All outcomes stated in the protocol were assessed (ChiCTR-TRC-12001871). However, the protocol was registered retrospectively.
Other bias	Unclear risk	It was unclear how the trial was funded.
Luk 1982		
Methods	Randomised clinical trial	
Participants	Sample size: 123 (experimental 62; control 61) Sex: not specified Age: not specified Country: US Setting: not specified Inclusion criteria: not specified Exclusion criteria: not specified	
Interventions	Experimental: cimetidine 300 mg iv for 6 hours Control: placebo iv for 6 hours Multi-arm trial: two additional arms were used (antacid and no drug). We excluded these two arms from our analysis	
Outcomes	• Any GI bleeding (definition not specified)	
Notes	25% of the patients died, but these deaths were not specified to randomisation group. Therefore, we could not use these data in our analysis. Contact information was not identified; thus, e-mail was not sent.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the methods of sequence generation was not described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	Low risk	The trial was described as double-blind using placebo.
Blinding of outcome	Unclear risk	The trial was described as double-blind, but it was unclear who was blinded and how

assessment (detection bias)		blinding was maintained.
Incomplete outcome data (attrition bias)	Low risk	No patient was lost to follow-up.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Unclear risk	It was not stated how the trial was funded.
Macdougall 1977		
Methods	Randomised clinical trial	
Participants	Sample size: 62, (experimental 26; control 36) Sex (male:female): not specified Age: not specified Country: UK Setting: medical ICU. Population: Patients with fulminant hepatic failure Inclusion criteria: not specified Exclusion criteria: not specified	
Interventions	Experimental: metiamide (n=10) or cimetidine (n=16) 150 mg iv Control: control not specified Duration: dose was repeated as necessary to maintain an intragastric pH above 5 measured two-hourly. Until recovery or death	
Outcomes	• Mortality • Any GI bleeding (definition not specified)	
Notes	We have combined the two experimental groups into one H2RA group. Contact information was not identified; thus, e-mail was not sent.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	High risk	Un-blinded.
Blinding of outcome assessment (detection bias)	High risk	Un-blinded.
Incomplete outcome data (attrition bias)	Low risk	No patients were lost to follow up.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Unclear risk	Smith, Kline and French Laboratories supplied metiamide and cimetidine. It was not stated whether the company was involved in other aspects of the trial. Contact information was not identified; thus, e-mail was not sent.
Martin 1993		
Methods	Randomised clinical trial	
Participants	Sample size: 131 (experimental 65; control 66) Sex (male:female): experimental 41:24; control 48:18 Age (mean): experimental 59; control 60 Country: USA Setting: Mixed ICU Inclusion criteria: eligible patients were males or nonlactating, nonpregnant females, ≥ 16 years of age, with a nasogastric tube in place, who were admitted to the ICU for a minimum anticipated treatment period of 36 hours and who had at least one of the following risk factors for upper GI hemorrhage: a) major surgery; b) multiple trauma to head, chest, abdomen, solid organs, or limbs; c) hypotension; d) hypovolemic shock; e) sepsis, including patients with peritonitis, confirmed bacteremia, or the complex of fever, increased WBC count, and hypotension with a bacteriologically determined source of infection; f) acute respiratory failure; g) jaundice with a plasma total bilirubin concentration of >513 umoVL (>30 mg/dL); or h) burns involving 230% of the body surface area. Patients were allowed to receive enteral feedings through a tube that traversed the pylorus into the small bowel. Exclusion criteria: a) if >24 hours had elapsed since they had become eligible for enrollment into the study; b) if patients had been intubated for >24 hours; c) if ICU admission esophageal, gastric, or duodenal surgery; or d) if had a history of gastrectomy or upper GI lesions that were likely to bleed. Patients were also excluded if they had received H2RA 12 hours of admission to the study or treatment 24 hours before admission to the study with omeprazole, anticoagulants (except low-dose heparin), aspirin, nonsteroidal anti-inflammatory agents, or treatment with an investigational drug within 30 days before entry. Additionally, during the screening phase, patients were excluded from the study if either of the two gastric aspirates demonstrated bright red blood, coffee ground material, or a strongly positive test for occult blood.	
Interventions	Experimental: cimetidine loading dose 300 mg and then 50 mg/hour Control: placebo Duration: 7 days	

	Co-intervention: not specified	
Outcomes	• Mortality • Any GI bleeding (defined as a) hematemesis or bright red blood that did not clear after nasogastric tube adjustment and a 5 to 10 min lavage; or b) persistent coffee ground material (eight consecutive hours) that were Gastroccult positive, not clearing with a 100 mL lavage, and/or accompanied by a 5% decrease in haematocrit) • Hospital-acquired pneumonia	
Notes	E-mail was sent to Dr. Rockhold 26.07.18 and reply was received.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	Low risk	Drug-matched placebo.
Blinding of outcome assessment (detection bias)	Low risk	Independent monitoring board.
Incomplete outcome data (attrition bias)	Low risk	No patients were lost to follow up.
Selective reporting (reporting bias)	High risk	Protocol was not pre-registered or published (clarification received by e-mail).
Other bias	High risk	The trial was supported in part by SmithKline Beecham Pharmaceuticals; however, the trial report did not state whether the company was involved in other aspects of the trial. The trial was stopped early due to the result of an interim analysis requested by the company (early stopping bias).
Metz 1993		
Methods	Randomised clinical trial	
Participants	Sample size: 167 (experimental 86; control 81) Sex (male:female): experimental 67:19; control 56:25 Age: experimental 32.5; control 35.4 Country: USA Setting: surgical ICU. Population: Patients with severe head injury Inclusion criteria: patients with severe head injury, defined as a Glasgow Coma Score of ≤ 10, were considered for study inclusion if they were at least 18 years of age, had a nasogastric tube in place, an expected intensive care unit (ICU) stay of at least 72 hours, and could be enrolled within 24 hours of injury Exclusion criteria: any patient with active gastrointestinal tract bleeding at baseline, severe burns (>20% of body surface area), renal insufficiency, documented peptic ulcer disease within 6 months, a baseline platelet count of thrombocytes/uL, or who received antacids within 4 hours or a histamine-2-receptor antagonist within 24 hours of study entry	
Interventions	Experimental: ranitidine 6.25 mg/hour continuous infusion Control: placebo Duration: 5 days Co-intervention: treatment with histamine-2-receptor antagonists (other than the study drug), antacids, prostaglandins, sucralfate, somatostatin analogues, vasopressin, nitroglycerin, propranolol, digitalis, and salicylates was prohibited by the study protocol. All other medications could be prescribed at the discretion of the attending physician.	
Outcomes	• Any GI bleeding (defined as a positive answer to any of the following questions: a) Was the gastric drainage occult blood positive and were "coffee grounds" present for the previous 8 hrs? b) Was there a minimum of 50 mL of bright red blood aspirated per nasogastric tube? c) Did the patient experience hematemesis in the last 8 hrs? d) Was there endoscopic or surgical confirmation of an upper gastrointestinal source of bleeding?) • Hospital-acquired pneumonia	
Notes	Contact information was not identified; thus e-mail was not sent.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated randomisation scheme.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	Low risk	A drug-matched placebo was used.
Blinding of outcome assessment (detection bias)	Unclear risk	Not described.
Incomplete outcome data	Low risk	No patients were lost to follow up.

(attrition bias)		
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Unclear risk	The trial was funded in part by Glaxo Pharmaceuticals, and the trial report did not state whether the company was involved in other aspects of the trial.
Nielsen 1989		
Methods	Randomised clinical trial	
Participants	Sample size: 25 (experimental 12; control 13) Sex (male:female): experimental 6:6; control 3:10 Age: experimental 56; control 59 Country: Denmark Setting: mixed ICU. Population: Patients with septicaemia or intra-abdominal sepsis. Inclusion criteria: diagnosis of septicaemia or intra-abdominal sepsis, with pyrexia >38.5^{0}C persisting for more than 48 hours despite comprehensive medical and/or early surgical treatment. Exclusion criteria: corticosteroids, nonsteroidal anti-inflammatory drugs, insulin, theophyllamine, antiviral or other known immunomodulating drugs.	
Interventions	Experimental: ranitidine 50 mg every 6 h Control: no prophylaxis. Duration: 8 days Co-intervention: all received prescribed drugs such as antibiotics, opiates and circulatory stimulants, with respiratory assistance, re-operation etc., as appropriate.	
Outcomes	• Mortality	
Notes	E-mail was sent to Dr. Nielsen and Dr. Kehlet 26.07.18. E-mail re-sent 27.08.18. Reply was received from Dr. Kehlet and Dr. Nielsen 27.08.18 (no clarifications were received).	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	The trial was described as randomised, bit the method of sequence generation was not described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	High risk	Un-blinded.
Blinding of outcome assessment (detection bias)	Unclear risk	Not described.
Incomplete outcome data (attrition bias)	Low risk	No patients were lost to follow up.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Unclear risk	It was unclear how the trial was funded.
Peura 1985		
Methods	Randomised clinical trial	
Participants	Sample size: 39 (experimental 21; control 18) Sex (male:female): 29 men and 11 women in total (not specified to randomisation group). Age: 21-84 (not specified to randomised group) Country: USA Setting: medical ICU Inclusion criteria: all patients admitted to the medical intensive care unit with an illness of sufficient severity to expect a minimum of 5 days' care in the unit were considered for entry to this study. Exclusion criteria: age less than 18 years, or the presence of acute myocardial infarction or pregnancy, prior gastric surgery, and contraindications to upper gastrointestinal endoscopy. Also excluded were patients with clinical evidence of active or recent gastrointestinal bleeding, such as hematemesis, melena, Hemoccult, positive stools, or nasogastric aspirate.	
Interventions	Experimental: cimetidine 300mg iv every 6h Control: placebo Duration: 14 days Co-intervention: medical management of patients was as standard as their underlying conditions allowed. Concomitant use of ulcerogenic drugs, such as salicylates and nonsteroidal anti-inflammatory agents, was not permitted nor were antacids used to neutralize intragastric pH. Otherwise, supportive management and therapy were permitted according to the needs of the patient.	
Outcomes	• Mortality • Clinically important bleeding (definition not specified)	
Notes	E-mail was sent to Dr. Peura 26.07.18 and reply was received (no clarifications was received).	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	The trial was described as being randomised, but the method of random sequence generation was not described.
Allocation concealment	Unclear risk	Not described.

(selection bias)		
Blinding of participants and personnel (performance bias)	Low risk	A drug-matched placebo was used.
Blinding of outcome assessment (detection bias)	Low risk	All endoscopic examinations were done by a single investigator and witnessed by a second investigator who observed the procedure though a lecturescope. During the endoscopy, findings were discussed and agreed on by both investigators before an entry was made on the report form. Both investigators were uninformed as to the patient's treatment group
Incomplete outcome data (attrition bias)	Low risk	No patients were lost to follow up.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Unclear risk	It was unclear how the trial was funded.
Powell (H2RA) 1993		
Methods	Randomised clinical trial	
Participants	Sample size: 16 (experimental 11; control 5) Sex (male:female): experimental 8:3; control 5:0 Age (mean): experimental 59.5; control 53.3 Country: UK Setting: surgical ICU Inclusion criteria: patients scheduled for coronary artery bypass graft surgery Exclusion criteria: active peptic ulcer disease, a previous definitive acid-lowering operation or current treatment with an H2 antagonist or other gastric antisecretory agent. Also excluded were patients with a history of severe allergy, those with concomitant renal or liver disease or receiving treatment with warfarin or phenytoin, and those who had received any non-licensed drug within the preceding 30 days.	
Interventions	Experimental: ranitidine 50 mg iv every 8 h Control: placebo Co-intervention: during the period on the ICU patients were given papaveretum for analgesia as required, midazolam or propofol for sedation, antibiotics, SNP or hydralazine to control hypertension, and frusemide to maintain adequate urine output. Duration: 6 days Multi-arm trial: we have divided the trial into two trials: 1) omeprazole infusion + omeprazole bolus = 20 patients vs. placebo 5 patients; 2) ranitidine 11 patients vs placebo 5 patients.	
Outcomes	• Mortality • Any GI bleeding (definition not specified)	
Notes	Contact information was not identified; thus, e-mail was not sent.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	High risk	Placebo controlled; however, obvious difference between bolus and infusion groups.
Blinding of outcome assessment (detection bias)	Unclear risk	It is stated that personnel collecting the aspirates did not know which treatment the patients received. It was not stated whether other personnel were blinded.
Incomplete outcome data (attrition bias)	Low risk	No patients were lost to follow up.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Unclear risk	Astra Clinical Research Unit supplied the drugs. It was not stated whether the company was involved in other aspects of the trial.
Powell (PPI) 1993		
Methods	Randomised clinical trial	
Participants	Sample size: 25 (experimental 20; control 5) Sex (male:female): experimental 17:3; control 5:0 Age (mean): experimental 57.7 (bolus) and 55.6 (infusion); control 53.3 Country: UK Setting: surgical ICU Inclusion criteria: patients scheduled for coronary artery bypass graft surgery Exclusion criteria: active peptic ulcer disease, a previous definitive acid-lowering operation or current treatment with an H2 antagonist or other gastric antisecretory agent. Also excluded were patients with a history of severe allergy, those with concomitant renal or liver disease or receiving treatment with warfarin or phenytoin, and those who had received any non-licensed drug within the preceding 30 days.	
Interventions	Experimental: omeprazole 80 mg iv loading dose then 40 mg every 8 hours by iv bolus or infusion Control: placebo Co-intervention: during the period on the ICU patients were given papaveretum for analgesia as required, midazolam or propofol for sedation, antibiotics, SNP or hydralazine to control hypertension, and frusemide to	

	maintain adequate urine output. Duration: 6 days Multi-arm trial: we have divided the trial into two trials: 1) omeprazole infusion + omeprazole bolus = 20 patients vs. placebo 5 patients; 2) ranitidine 11 patients vs placebo 5 patients.	
Outcomes	• Mortality • Any GI bleeding (definition not specified)	
Notes	We have divided the trial into two trials: 1) omeprazole infusion + omeprazole bolus = 20 patients vs. placebo 5 patients; 2) ranitidine 11 patients vs placebo 5 patients Contact information was not identified, thus e-mail was not sent.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	High risk	Placebo controlled; however, obvious difference between bolus and infusion groups.
Blinding of outcome assessment (detection bias)	Unclear risk	It is stated that personnel collecting the aspirates did not know which treatment the patients received. It was not stated whether other personnel were blinded.
Incomplete outcome data (attrition bias)	Low risk	No patients were lost to follow up.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Unclear risk	Astra Clinical Research Unit supplied the drugs. It was not stated whether the company was involved in other aspects of the trial.
Rigaud 1988		
Methods	Randomised clinical trial	
Participants	Sample size: 12 (experimental 6; control 6) Sex (male:female): experimental 4:2; control 4:2 Age: experimental 65; control 70 Country: France Setting: medical ICU. Inclusion criteria: ICU patients admitted for acute respiratory failure due to chronic obstructive pulmonary disease. Exclusion criteria: history of peptic ulcer, gastrointestinal bleeding, gastric surgery, renal insufficiency, or if taking steroids or other anti-inflammatory drugs	
Interventions	Experimental: ranitidine 0.25µg/kg/h Control: placebo Duration: 3 days Co-intervention: all patients were given respiratory assistance during the study and received antibiotics (ampicillin) intravenously Multi-arm trial: 4 arms. The patients were randomly assigned to receive, iv either ranitidine (ranitidine group) or placebo (placebo group) at constant rates in a double-blind manner. Four therapeutic regimens: placebo only, continuous enteral nutrition (CEN) only, ranitidine, ranitidine and CEN were studied.	
Outcomes	• No relevant outcomes reported.	
Notes	Contact information was not identified; thus, e-mail was not sent.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	Low risk	Drug-matched placebo.
Blinding of outcome assessment (detection bias)	Unclear risk	Not described.
Incomplete outcome data (attrition bias)	Unclear risk	No patients were lost to follow up.
Selective reporting (reporting bias)	High risk	No protocol could be found, and no relevant outcomes for our review were reported.
Other bias	Unclear risk	Glaxo Laboratories and Lilly Laboratories sponsored PPI and antibiotics. It was not stated whether the company was involved in other aspects of the trial.
Rohde 1980		
Methods	Clinical randomised trial	

Participants	Sample size: 28 (experimental 14, control 14) Sex (male:female): not specified Age: not specified Country: Germany Setting: surgical ICU. Population: Patients with polytrauma Inclusion criteria: not specified Exclusion criteria: not specified	
Interventions	Experimental: cimetidine Control: placebo Co-intervention: not specified Duration: not specified	
Outcomes	• Mortality • Any GI bleeding (definition not specified)	
Notes	Contact information was not identified; thus, e-mail was not sent.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	High risk	The trial was described as "single-blinded", but it was unclear who was blinded and how blinding was maintained.
Blinding of outcome assessment (detection bias)	Unclear risk	The trial was described as "single-blinded", but it was unclear who was blinded and how blinding was maintained.
Incomplete outcome data (attrition bias)	High risk	The trial was stopped early. Not described whether the interim analysis was pre-defined.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Unclear risk	It was unclear how the trial was funded.
Ruiz-Santana 1991		
Methods	Randomised clinical trial	
Participants	Sample size: 49 (experimental 19; control 30) Sex (male:female): experimental 14:5; control 19:11 Age: experimental 39; control 39 Country: Spain Setting: mixed ICU Inclusion criteria: ICU patients with an illness of sufficient severity to expect a minimum of 6 days of mechanical ventilation. Patients had to be in metabolic stress, hemodynamically stable, with normal hepatic and renal function, and on total parenteral nutrition. Exclusion criteria: patients with spinal cord injury.	
Interventions	Experimental: parenteral nutrition + ranitidine 50 mg iv every 6 h Control: parenteral nutrition Duration: until weaning, GI bleeding, or death Co-intervention: total parenteral nutrition Multi-arm trial: the trial had a third arm (sucralfate) which we excluded from our study.	
Outcomes	• Mortality • Any GI bleeding (defined as hematemesis, bloody aspirate, melena, "coffee grounds" material)	
Notes	E-mail was sent to Dr. Ruiz-Santana 26.07.18 and reply was received.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Simple randomisation (clarification received by e-mail).
Allocation concealment (selection bias)	Low risk	Sealed envelopes (clarification received by e-mail).
Blinding of participants and personnel (performance bias)	High risk	Not blinded as the control group received no intervention.
Blinding of outcome assessment (detection bias)	High risk	Endoscopic evaluations were done by a single investigator uninformed as to the treatment group - except for a few emergency cases.
Incomplete outcome data (attrition bias)	High risk	24/97 (25%) patients are excluded from the analysis. The following reasons were stated: 10 due to mech ventilation < than 6 days, 8 deaths, 5 acute upper GI haemorrhage, 1 early tolerance to enteral feeding. As these patients have received the intervention it is not correct to exclude them from the analysis.
Selective reporting (reporting bias)	High risk	Protocol was not pre-registered or published (clarification received by e-mail).
Other bias	Low risk	The trial was supported by hospital funds (clarification received by e-mail).

Selvanderan 2016	
Methods	Randomised clinical trial
Participants	Sample size: 216 (experimental 107; control 109) Sex (male): experimental 64%; control 67% Age: experimental 52; control 52 Country: Australia Setting: mixed ICU. Population: Mechanically ventilated patients Inclusion criteria: patients who were anticipated to be invasively mechanically ventilated for greater than 24 hours and receive enteral nutrition within 48 hours of admission were eligible for inclusion. Exclusion criteria: exclusion criteria included are as follows: 1) use of acid-suppressive therapy prior to admission, 2) admission with gastrointestinal bleeding, 3) history of proven peptic ulcer disease, 4) administration of greater than 100 mg daily of prednisolone (or equivalent of other corticosteroid), 5) surgery on the upper gastrointestinal tract or cardiac surgery during the current hospital admission, 6) pregnancy, 7) Jehovah's witnesses, 8) patients who could not receive their first dose of study medication within 36 hours of initiation of mechanical ventilation, 9) admission for the sole purpose of providing palliative care, and 10) patients readmitted to the ICU.
Interventions	Experimental: pantoprazole (40mg in 10ml of 0.9% saline iv) Control: placebo (10ml of 0.9% saline iv) Duration: until extubation or max 14 days Co-intervention: not specified.
Outcomes	• Mortality • Any GI bleeding (defined as overt GI bleeding: hematemesis, bloody gastric aspirate, melena, or haematochezia) • Clinically important bleeding (defined as an episode of overt bleeding (hematemesis, bloody gastric aspirate, melena, or haematochezia), accompanied by at least one of the following: 1) a reduction in mean arterial blood pressure of more than or equal to 20 mm Hg within 24 hours in the absence of another cause, 2) a reduction in haemoglobin of more than or equal to 20 g/L within 24 hours, or 3) a need for endoscopy or surgery to achieve haemostasis) • Hospital-acquired pneumonia • Cl. difficile
Notes	E-mail was sent to Dr. Deane 26.07.18. E-mail re-sent 27.08.18.

Risk of bias assessment

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer generated randomisation.
Allocation concealment (selection bias)	Low risk	Hospital pharmacy performed the randomisation.
Blinding of participants and personnel (performance bias)	Low risk	A drug matched placebo was used.
Blinding of outcome assessment (detection bias)	Low risk	All outcomes were assessed while the investigators remained blinded to treatment allocation.
Incomplete outcome data (attrition bias)	Low risk	One patient in each group was excluded from analysis due to withdrawn consent for ongoing participation and use of data after randomisation.
Selective reporting (reporting bias)	Low risk	All outcomes stated in the protocol (ACTRN12613000807752) were assessed.
Other bias	Low risk	The trial was supported by public grants.

Spapen 1995	
Methods	Randomised clinical trial
Participants	Sample size: 30 (experimental 20; control 10) Sex (male:female): not specified Age: not specified Country: Belgium Setting: mixed ICU Inclusion criteria: haemodynamic stability. Exclusion criteria: patients treated with corticosteroids and Hr blocking agents within 4 weeks before study, or suffering from any known intercurrent endocrinologic disease were excluded from the study. Patients who continued to fight the ventilator, even after the use of heavy sedation 0.1 mg fentanyl/ hour and/ or > 3 mg midazolam/ hour) were equally not included. Patients were withdrawn from the study when significant circulatory failure lasting 2 6 hours developed or death occurred.
Interventions	Experimental: cimetidine 1200mg or ranitidine 200 mg iv during study period Control: placebo Duration: 24h Co-intervention: not specified Multi-arm trial: three arms (cimetidine vs ranitidine vs placebo) were compared. We pooled the two H2RA group into one group.
Outcomes	• Mortality
Notes	E-mail was sent to Dr. Spapen 26.07.18. E-mail re-sent 27.08.18.

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Not described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	High risk	Not blinded.
Blinding of outcome assessment (detection bias)	High risk	Not blinded.
Incomplete outcome data (attrition bias)	High risk	The authors excluded 10/30 patients, 6 due to mortality and 4 due to severe cardiocirculatory dysfunction. As the exclusion were reported by group, we have re-included the patients being excluded due to mortality.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Unclear risk	It was unclear how the trial was funded.
Vlatten 1998		
Methods	Randomised clinical trial	
Participants	Sample size: 60 (experimental 30; control 30) Sex (male:female): not specified Age: not specified Country: Germany Setting: mixed ICU. Population: Intensive care patients with septic shock, polytrauma or skull-brain trauma Inclusion criteria: not specified Exclusion criteria: not specified	
Interventions	Experimental: 2 x 40 mg omeprazole Control: no prophylaxis Duration: not specified Multi-arm trial: the trial compared three interventions (omeprazole vs pirenzepine vs no prophylaxis). We did not use the results from the pirenzepine group. We assume 30 patients were randomised to each group as this was not specified	
Outcomes	• Any GI bleeding (definition not specified)	
Notes	E-mail was sent to Dr. Vlatten and Dr. Georgieff 26.07.18. E-mail re-sent 27.08.18.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	High risk	Un-blinded.
Blinding of outcome assessment (detection bias)	High risk	Un-blinded.
Incomplete outcome data (attrition bias)	Unclear risk	Not described.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Unclear risk	It was unclear how the trial was funded
Zinner 1981		
Methods	Randomised clinical trial	
Participants	Sample size: 226 (experimental 113; control 113) Sex (male:female): experimental 63:37; control 63:37 Age: experimental 56.7; control 55.5 Country: US Setting: surgical ICU patients Inclusion criteria: patient admitted for at least 48 h. Exclusion criteria: patients with upper gastrointestinal tract bleeding, those with recent active peptic ulcer disease or those who had undergone an operation on the oesophagus or the stomach.	
Interventions	Experimental: cimetidine 300 mg iv every 6h during the entire stay in the ICU. Control: no prophylaxis Duration: during entire stay in the ICU Co-intervention: if a nasogastric tube was not in place, a dose of 20 mL of Maalox Therapeutic Concentrate was given orally every two hours. Multi-arm trial: the trial compared three interventions (cimetidine vs no prophylaxis vs antacid). We excluded the results from the antacid group.	
Outcomes	• Mortality • Any GI bleeding (definition not specified)	

	• Clinically important bleeding (GI bleeding requirering transfusion)	
Notes	40 withdrawn (evenly distributed according to report), 14 antacid, 13 cimetidine, 13 no treatment. Contact information was not identified; thus, e-mail was not sent.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Table of random numbers was used.
Allocation concealment (selection bias)	Unclear risk	Not described.
Blinding of participants and personnel (performance bias)	High risk	Un-blinded.
Blinding of outcome assessment (detection bias)	High risk	Un-blinded.
Incomplete outcome data (attrition bias)	High risk	No patients were excluded from analysis. Three hundred patients met the criteria. Forty additional patients were entered into the randomised study but were removed from the protocol. Thirty-one of these were excluded because of protocol errors or because of the request of the physician. These were evenly distributed between the treatment groups. There were three treatment groups, with 100 patients in each group.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found.
Other bias	Unclear risk	It was unclear how the trial was funded.

* In total contact details could not be identified for authors on 16 trial reports, 23 emails were sent, and 9 replies were received

Table S5. Excluded studies and ongoing trials

EXCLUDED STUDIES

1. Adt M, Schutz A, Baumert JH, Reimann HJ. Histamine response to various stimuli during CABG-surgery: a study in patients with and without prophylactically administered H1- and H2-receptor antagonists. Inflammation Research 1995; 44 Suppl 1: S80-1.
2. Agnew NM, Kendall JB, Akrofi M, et al. Gastroesophageal reflux and tracheal aspiration in the thoracotomy position: should ranitidine premedication be routine? Anesthesia & Analgesia 2002; 95(6): 1645-9, table of contents.
3. Apte NM, Karnad DR, Medhekar TP, et al. GASTRIC COLONIZATION AND PNEUMONIA IN INTUBATED CRITICALLY ILL PATIENTS RECEIVING STRESS-ULCER PROPHYLAXIS - A RANDOMIZED, CONTROLLED TRIAL. Critical Care Medicine 1992; 20(5): 590-593.
4. Baccino E, Boles JM, Guillou M, et al Attempt at preventive treatment of esophagitis caused by intubation during intensive care. 1987. 11, 24-28.
5. Baccino E, Boles JM, Le Guillou M, et al. [Attempt at preventive treatment of esophagitis caused by intubation during intensive care]. Gastroenterologie Clinique et Biologique 1987; 11(1): 24-8.
6. Bailey RJ, MacDonald BRD, Williams R A controlled trial of H2 -receptor antagonists in prophylaxis of bleeding from gastrointestinal erosions in fulminant hepatic failure. 1976. 17, 389.
7. Bailey RJ, Macdougall BR, Williams R. Proceedings: A controlled trial of H2-receptor antagonists in prophylaxis of bleeding from gastrointestinal erosions in fulminant hepatic failure. Gut 1976; 17(5): 389.
8. Bailey RJ, MacDougall BRD, Williams R. A controlled trial of H-2-receptor antagonists in prophylaxis of bleeding from gastrointestinal erosions in fulminant hepatic failure. Gut 1976; 17(5): 389.
9. Bailey RJ, Macdougall BRD, Williams R. CONTROLLED TRIAL OF H-2-RECEPTOR ANTAGONISTS IN PROPHYLAXIS OF BLEEDING FROM GASTROINTESTINAL EROSIONS IN FULMINANT HEPATIC-FAILURE. Gut 1976; 17(5): 389-389.
10. Bailey RJ, MacDougall BRD, Williams R A contolled trial of H2 -receptor antagonists in prophylaxix of bleeding from gastrointestinal erosions in fulminant hepatic failure. 1976. 17, 389.
11. Baltas B, Nagy A, Baradnay G, Pepo J. [Experience with prophylactic cimetidine therapy in high risk patients]. 1980; 121(50): 3043-7.
12. Barnes PJ, Havill JH. Preoperative cimetidine--effects on gastric fluid. 1980; 8(4): 464-8.
13. Barri YM, Ramos EL, Balagtas RS, Peterson JC, Karlix JL Cimetidine or rantidine in prophylactic doses do not affect renal function or cyclosporine levels in renal transplant patients (rtp). 1993. 4, 925.
14. Ben-Menachem T, Fogel R, Patel RV, et al. Prophylaxis for stress-related gastric hemorrhage in the medical intensive care unit. A randomized, controlled, single-blind study. Annals of Internal Medicine 1994; 121(8): 568-75.
15. Ben-Menachem T, Patel RV, Bresalier RS, Fogel R. Efficacy and safety of prophylaxis for stress-related gastrointestinal hemorrhage in the medical intensive care unit: A controlled clinical trial. Clinical Research 1993; 41(2): 200A.
16. Bhanot RD. Nosocomial pneumonia in mechanically ventilated patients receiving ranitidine, omeprazole or sucralfate as stress ulcer prophylaxis. American Thoracic Society International Conference, ATS 2010 New Orleans 181 (1 Meeting Abstracts) 2010.
17. Bhatt DL, Cryer BL, Contant CF, et al. Clopidogrel with or without omeprazole in coronary artery disease. New England Journal of Medicine 2010; 363(20): 1909-1917.
18. Cai J, Wu Q, Fan L, et al. [Impact of different proton pump inhibitors on the antiplatelet activity of clopidogrel in combination with aspirin for patients undergoing coronary stent implantation]. Zhongguo Ying Yong Sheng Li Xue Za Zhi 2010; 26(3): 266-9.
19. Calvet X, Baigorri F, Duarte M, et al. Effect of ranitidine on gastric intramucosal pH in critically ill patients. Intensive Care Medicine 1998; 24(1): 12-7.
20. Chan KH, Lai EC, Tuen H, Ngan J Prospective double blind randomized study on the use of ranitidine in preventing acute gastroduodenal disease after neurosurgery. 1991. 6, 90.
21. Cheadle WG, Vitale GC, Mackie CR, Cuschieri A Prophylactic postoperative nasogastric decompression. A prospective study of its requirement and the influence of cimetidine in 200 patients. 1985. 202, 361-366.
22. Chen HM, Li XB, Ge ZZ, et al. Rabeprazole combined with hydrotalcite is effective for patients with bile reflux gastritis after cholecystectomy. Canadian Journal of Gastroenterology 2010; 24(3): 197-201.
23. Chen ZQ. Influence of Ranitidine on gastrointestinal hemorrhage and thrombosis caused by dual antiplatelet therapy after PCI in patients with coronary heart disease. Cardiology 2010; 117: 48-49.
24. Chernov MS, Hale HW, Jr., Wood M. Prevention of stress ulcers. American Journal of Surgery 1971; 122(5): 674-7.
25. Chia YY, Chiang HL, Liu K, Ko NH. Randomized, double-blind study comparing postoperative effects of treatment timing with histamine H-2-receptor antagonist cimetidine. Acta Anaesthesiologica Scandinavica 2005; 49(6): 865-869.
26. Chia YY, Chiang HL, Liu K, Ko NH Randomized, double-blind study comparing postoperative effects of treatment timing with histamine H-receptor antagonist cimetidine. 2005. 49, 865-869 DOI: 10.1111/j.1399-6576.2005.00696.x.
27. Cloud ML, Offen W. Continuous infusions of nizatidine are safe and effective in the treatment of intensive care unit patients at risk for stress gastritis. The Nizatidine Intensive Care Unit Study Group. Scandinavian Journal of Gastroenterology - Supplement 1994; 206: 29-34.
28. Cloud ML, Offen W. Continuous infusions of nizatidine are safe and effective in the treatment of intensive care unit patients at risk for stress gastritis. Scandinavian Journal of Gastroenterology, Supplement 1994; 29(206): 29-34.
29. Cloud ML, Offen W. CONTINUOUS INFUSIONS OF NIZATIDINE ARE SAFE AND EFFECTIVE IN THE TREATMENT OF INTENSIVE-CARE UNIT PATIENTS AT RISK FOR STRESS GASTRITIS. Scandinavian Journal of Gastroenterology 1994; 29: 29-34.
30. Darlong V, Jayalakhsmi TS, Kaul HL, Tandon R. Stress ulcer prophylaxis in patients on ventilator. Tropical Gastroenterology 2003; 24(3): 124-8.
31. Deane A, Selvanderan S, Summers M, et al. Pantoprazole or placebo for stress ulcer prophylaxis (Popup) study. Anaesthesia and Intensive Care 2016; 44 (2): 303-304.
32. Dearden J, Grant D, van Someren N. Proton pump inhibitors are only effective in the short term for preventing PEG site excoriation. Gastroenterology 2004; 126(4, Suppl. 2): A646.
33. Delalande JP, Perramant M, Perramant-Creach Y, Gerard L, Egreteau JP. [Value of cimetidine in the prevention of Mendelson's syndrome in abdominal surgery]. Annales de l'Anesthesiologie Francaise 1981; 22(1): 48-52.
34. Delalande JP, Perramant M, Perramant-Creach Y, Gérard L, Egreteau JP Value of cimetidine in the prevention of Mendelson's syndrome in abdominal surgery. 1981. 22, 48-52.
35. DiMango E, Walker P, Keating C, et al. Effect of esomeprazole versus placebo on pulmonary exacerbations in cystic fibrosis. BMC Pulmonary Medicine 2014; 14 (1) (no pagination)(21).
36. DiMango E, Walker P, Keating C, et al. Effect of esomeprazole versus placebo on pulmonary exacerbations in cystic fibrosis.

BMC Pulmonary Medicine 2014; 14: 7.
37. Duffy BL, Woodhouse PC, Schramm MD, Scanlan CM Ranitidine prophylaxis before anaesthesia in early pregnancy. 1985. 13, 29-32.
38. Essardas Daryanani H, Santolaria Fernandez FJ, Gonzalez Reimers CE, et al. Acute pancreatitis: comparative therapeutic study among cimetidine, oxyphenonium and simple nasogastric aspiration. [Spanish] Pancreatitis aguda: estudio terapeutico comparativo entre cimetidina, oxifenonio y aspiracion nasogastrica simple. Medicina Clinica 1983; 81(1): 4-6.
39. Fang XW, Chang S, Zhao JH, Qian XY. Prevention and treatment of stress-induced gastrointestinal bleeding after severe brain injury. [Chinese]. World Chinese Journal of Digestology 2014; 22(3): 404-408.
40. Fang XW, Chang S, Zhao JH, Qian XY Prevention and treatment of stress-induced gastrointestinal bleeding after severe brain injury. 2014. 22, 404-408 DOI: 10.11569/wcjd.v22.i3.404.
41. Fischer M, Lorenz W, Reimann HJ Cimetidine prophylaxis of acute gastroduodenal lesions in patients at risk. <COLLECTIVE> Cimetidine: proceedings of an international symposium on histamine Hsub 2-receptor antagonists. 1978. 443, 280-291.
42. Gao QP, Sun Y, Sun YX, Wang LF, Fu L. Early use of omeprazole benefits patients with acute myocardial infarction. Journal of Thrombosis & Thrombolysis 2009; 28(3): 282-7.
43. Geus WP, Haas JAM, Smith SJ, et al Antisecretory effects of intravenous ranitidine and omeprazole in postoperative intensive care unit patients during prolonged dosing. 1997. 9, A 2.
44. Goel C, Anand LK, Gombar KK Comparative evaluation of single dose intravenous pantoprazole and ranitidine on gastric pH and volume: a double blind study. 2006. 22, 145-149.
45. Griffin PJ, Salaman JR. Ranitidine or aluminium hydroxide as prophylaxis against gastrointestinal complications after renal transplantation. Journal of the Royal College of Surgeons of Edinburgh 1986; 31(2): 88-90.
46. Halloran LG, Zfass AM, Gayle WE, Wheeler CB, Miller JD Prevention of acute gastrointestinal complications after severe head injury: a controlled trial of cimetidine prophylaxis. 1980. 139, 44-48.
47. Halloran Lg GWEZAMWCBMJD. Cimetidine prevention of gastrointestinal bleeding in severe head injury: a controlled clinical trial. Surgical forum 1978; 29: 428-30.
48. Hanisch EW, Encke A, Naujoks F, Windolf J. A randomized, double blind trial for stress ulcer prophylaxis shows no evidence of increased pneumonia. American Journal of Surgery 1998; 176(5): 453-457.
49. Hastings PR, Skillman JJ, Bushnell LS, Silen W. Antacid titration in the prevention of acute gastrointestinal bleeding: a controlled, randomized trial in 100 critically ill patients. New England Journal of Medicine 1978; 298(19): 1041-5.
50. Hastings PR, Skillman JJ, Bushnell LS, Silen W. ANTI ACID TITRATION IN THE PREVENTION OF ACUTE GASTRO INTESTINAL BLEEDING A CONTROLLED RANDOMIZED TRIAL IN 100 CRITICALLY ILL PATIENTS. New England Journal of Medicine 1978; 298(19): 1041-1045.
51. He Gw ZYMZJ. Naloxone combined with pantoprazole for prevention of upper gastrointestinal bleeding in patients with respiratory failure. World Chinese Journal of Digestology 2017; 25(1): 102.
52. Heiselman DE, Chapman J, Malik M, Riegnor E. Hemodynamic status during famotidine infusion. DICP 1990; 24(12): 1163-5.
53. Holzapfel L DLCA, et al. Prophylaxis against stress induced bleeding and nosocomial pneumonia in patients on mechanical ventilation (abstract). Intensive Care Med 1990; 16(Suppl 1)(S4).
54. Hongo M, Hoshihara Y. Efficacy and safety of lansoprazole (AG-1749) 15 mg and 30 mg in Japanese patients with non-erosive reflux disease (NERD): A phase III multicenter, double-blind, placebo-controlled trial. [Japanese]. Japanese Pharmacology and Therapeutics 2008; 36(7): 655-671.
55. Hudzik B, Szkodzinski J, Danikiewicz A, et al. Effect of omeprazole on the concentration of interleukin-6 and transforming growth factor-beta1 in patients receiving dual antiplatelet therapy after percutaneous coronary intervention. European Cytokine Network 2010; 21(4): 257-63.
56. Hummer-Sigiel M, Jacquier A, Girard A. The effect of ranitidine is compared of that of a placebo in 22 patients who are treated in intensive care for severe cerebral injuries. [French] Ranitidine Pour La Prophylaxie De L'ulcere De Stress Chez Les Traumatises Craniens Graves. Annales Medicales de Nancy et de l'Est 1986; 25(2): 101-103.
57. Hummer-Sigiel M, Jacquier A, Girard A The effect of ranitidine is compared of that of a placebo in 22 patients who are treated in intensive care for severe cerebral injuries. <ORIGINAL> RANITIDINE POUR LA PROPHYLAXIE DE L'ULCERE DE STRESS CHEZ LES TRAUMATISES CRANIENS GRAVES. 1986. 25, 101-103.
58. Idjadi F, Robbins R, Stahl WM, Essiet G. Prospective study of gastric secretion in stressed patients with intracranial injury. J Trauma 1971; 11(8): 681-8.
59. Jones LK. Stress-induced gastroduodenal lesions and total parenteral nutrition in critically ill patients: frequency, complications, and the value of prophylactic treatment. A prospective, randomized study. Jpen 1992; Journal of parenteral and enteral nutrition. 16(2): 182-183.
60. Jones RH, Rudge CJ, Bewick M, Parsons V, Weston MJ. Cimetidine: prophylaxis against upper gastrointestinal haemorrhage after renal transplantation. British Medical Journal 1978; 1(6110): 398-400.
61. Kam J, Modi C, Doraiswamy V, et al. Role of gastrointestinal ulcer prophylaxis in critically ill patients ACG IBD award. American Journal of Gastroenterology 2011; 106: S420.
62. Kantorova I, Svoboda P, Scheer P, et al. Stress ulcer prophylaxis in critically III patients: A randomized controlled trial. Hepato-Gastroenterology 2004; 51(57): 757-761.
63. Karlstadt RG, Dambrosio CA, Palmer RH, et al. NO INCREASED RISK OF NOSOCOMIAL PNEUMONIA WITH CIMETIDINE - EXPERIENCE IN A MULTICENTER, PLACEBO-CONTROLLED ICU STUDY. Clinical Research 1990; 38(2): A273-A273.
64. Katoh J, Tsuchiya K, Osawa H, et al. Cimetidine reduces impairment of cellular immunity after cardiac operations with cardiopulmonary bypass. J Thorac Cardiovasc Surg 1998; 116(2): 312-8.
65. Katoh J, Tsuchiya K, Sato W, Nakajima M, Iida Y. Cimetidine and immunoreactivity. Lancet 1996; 348(9024): 404-5.
66. Kaushal S MVSACSCGC. A comparative study of the effects of famotidine and sucralfate in prevention of upper gastro-intestinal bleeding in patients of head injury. Indian Journal of Pharmacology 2000; 32(3): 246-9.
67. Kessing R. Intensive Patients: PPI Administration without increased Risk of Bacteremia. Zeitschrift fuer Gastroenterologie 2017; 55(10): 984.
68. Kohler TR, Dellinger EP, Simonowitz DA, Blair AD, Edwards WA. Cimetidine pharmacokinetics in trauma patients. Surgical Forum 1979; 30: 12-4.
69. Kokhaei P, Barough MS, Hassan ZM. Cimetidine effects on the immunosuppression induced by burn injury. International Immunopharmacology 2014; 22(1): 273-6.
70. Laheij RJ, Rossum LG, Jansen JB, Verheugt FW Proton-pump inhibitor therapy for acetylsalicylic acid associated upper gastrointestinal symptoms: a randomized placebo-controlled trial. 2003. 18, 109-115.
71. Laheij RJ, van Rossum LG, Jansen JB, Verheught FW. PROTON-PUMP INHIBITOR THERAPY FOR ACETYLSALICYLIC ACID ASSOCIATED UPPER GASTROINTESTINAL SYMPTOMS, A RANDOMISED CLINICAL TRIAL. Digestive Disease Week Abstracts and Itinerary Planner 2003; 2003: T1328.
72. Larson GM, Livingston DH, Smith JS, et al. Impact of multiple risk factors and ranitidine prophylaxis on the development of stress-related upper gastrointestinal bleeding: A prospective, multicenter, double-blind, randomized trial. Critical Care Medicine 1993; 21(12): 1844-1849.

73. Lavercombe P, Bowler S, Smith S The effect of stress ulcer prophylaxis on the incidence of pulmonary infection in patients receiving mechanical ventilation. 1989. 19, 668.
74. Lehot JJ, Deleat-Besson R, Bastien O, et al. Should we inhibit gastric acid secretion before cardiac surgery? Anesthesia & Analgesia 1990; 70(2): 185-90.
75. Liu BL, Li B, Zhang X, et al. A randomized controlled study comparing omeprazole and cimetidine for the prophylaxis of stress-related upper gastrointestinal bleeding in patients with intracerebral hemorrhage. Journal of Neurosurgery 2013; 118(1): 115-20.
76. Liu BL, Li B, Zhang X, et al. A randomized controlled study comparing omeprazole and cimetidine for the prophylaxis of stress-related upper gastrointestinal bleeding in patients with intracerebral hemorrhage Clinical article. Neurosurgical Focus 2013; 34(1): 115-120.
77. Liu F, Zhang X, Yang X, et al. Influence of different proton pump inhibitors on the antiplatelet action of clopidogrel associated with aspirin. Circulation 2012; 125 (19): e854.
78. Lorenz W, Fischer M, Rohde H, et al. Histamine and stress ulcer: New components in organizing a sequential trial on cimetidine prophylaxis in seriously ill patients and definition of a special group at risk (severe polytrauma). Klinische Wochenschrift 1980; 58(13): 653-665.
79. Lu YH Cimetidine for preventing stress ulcer associated with cerebral hemorrhage resulting from high pressure. 1998. 10, 446.
80. Macdougall BR, Bailey RJ, Williams R. H2-receptor antagonists and antacids in the prevention of acute gastrointestinal haemorrhage in fulminant hepatic failure. Two controlled trials. Lancet 1977; 1(8012): 617-9.
81. MacDougall BR, Williams R. H2-receptor antagonist in the prevention of acute upper gastrointestinal hemorrhage in fulminant hepatic failure: a controlled trial. Gastroenterology 1978; 74(2 Pt 2): 464-5.
82. Macdougall BRD, Bailey RJ, Williams R. H2-RECEPTOR ANTAGONISTS AND ANTACIDS IN PREVENTION OF ACUTE GASTROINTESTINAL HEMORRHAGE IN FULMINANT HEPATIC-FAILURE - 2 CONTROLLED TRIALS. Lancet 1977; 1(8012): 617-619.
83. MacDougall BRD, Williams R. H-2-receptor antagonist in the prevention of acute upper gastrointestinal hemorrhage in fulminant hepatic failure. A controlled trial. Gastroenterology 1978; 74(2 II): 464-465.
84. Macdougall BRD, Williams R. H2-RECEPTOR ANTAGONIST IN PREVENTION OF ACUTE UPPER GASTROINTESTINAL HEMORRHAGE IN FULMINANT HEPATIC-FAILURE - CONTROLLED TRIAL. Gastroenterology 1978; 74(2): 464-465.
85. Macdougall BRD, Williams R. H-2 RECEPTOR ANTAGONIST IN THE PREVENTION OF ACUTE UPPER GASTRO INTESTINAL HEMORRHAGE IN FULMINANT HEPATIC FAILURE A CONTROLLED TRIAL. Gastroenterology 1978; 74(2 PART 2): 464-465.
86. Marti MC, Suter P, Dubouloz M. [Cimetidine prevention of stress ulcers]. Schweizerische Medizinische Wochenschrift. Journal Suisse de Medecine 1979; 109(16): 615-6.
87. Martin JN. Prophylaxis for stress-related gastric hemorrhage. Ann Intern Med 1995; 122(8): 633.
88. Maynard N ASMRSMBD. Influence of intravenous ranitidine on gastric intramucosal pH in critically ill patients. Crit Care Med 1994; 22(A79).
89. Mayumi H, Toshima Y, Tokunaga K. Pretreatment with H2 blocker famotidine to ameliorate protamine-induced hypotension in open-heart surgery. J Cardiovasc Surg (Torino) 1992; 33(6): 738-45.
90. McCarthy G, Mirakhur RK, Elliott P, Wright J. Effect of H2-receptor antagonist pretreatment on vecuronium- and atracurium-induced neuromuscular block. 1991; 66(6): 713-5.
91. Metz CA. Ranitidine prophylaxis for stress-related upper gastrointestinal bleeding in patients with severe head injury: Clinical and pharmacoeconomic outcome. Digestive Surgery 1996; 13(4-5): 405-409.
92. Metz CA, Livingston DH, Smith JS, et al. IMPACT OF MULTIPLE RISK-FACTORS AND RANITIDINE PROPHYLAXIS ON THE DEVELOPMENT OF STRESS-RELATED UPPER GASTROINTESTINAL-BLEEDING - A PROSPECTIVE, MULTICENTER, DOUBLE-BLIND, RANDOMIZED TRIAL. Critical Care Medicine 1993; 21(12): 1844-1849.
93. Misra UK, Kalita J, Pandey S, Mandal SK, Srivastava M. A randomized placebo controlled trial of ranitidine versus sucralfate in patients with spontaneous intracerebral hemorrhage for prevention of gastric hemorrhage. Journal of the neurological sciences 2005; 239(1): 5-10.
94. Mizia-Stec K, Haberka M, Mizia M, et al. Effects of pantoprazole on dual antiplatelet therapy in stable angina pectoris patients after percutaneous coronary intervention. Pharmacological Reports: PR 2012; 64(2): 360-8.
95. Moesgaard F, Nielsen HJ, Bardram L. Postoperative infectious complications following emergency colorectal surgery; A randomized, prospective and double blind trial comparing ranitidine vs. placebo. Gastroenterology 1994; 106(4 SUPPL.): A737.
96. Nagasue N, Yukaya H, Ogawa Y, Sasaki Y, Hirose S. Prophylaxis of upper gastrointestinal bleeding with cimetidine in patients undergoing partial hepatectomy. Annales Chirurgiae et Gynaecologiae 1984; 73(1): 6-10.
97. Nielsen HJ, Hammer JH, Gronvall S, et al. The effect of ranitidine on immune function, tumor response and survival in patients with liver metastases from colorectal cancer. Gi Cancer 1996; 1(3): 183-190.
98. Nielsen HJ, Hammer JH, Moesgaard F, Heron I, Kehlet H. RANITIDINE IMPROVES POSTOPERATIVE SUPPRESSION OF ANTIBODY-RESPONSE TO PREOPERATIVE VACCINATION. Surgery 1992; 111(1): 69-73.
99. Nielsen HJ, Moesgaard F, Kehlet H. Ranitidine for prevention of postoperative suppression of delayed hypersensitivity. American Journal of Surgery 1989; 157(3): 291-4.
100. Nielsen HJ, Mynster T, Jensen S, Hammer J, Nielsen H. Effect of ranitidine on soluble interleukin 2 receptors and CD8 molecules in surgical patients. British Journal of Surgery 1994; 81(12): 1747-51.
101. Nielsen HJ, Nielsen H, Jensen S, Moesgaard F. Ranitidine improves postoperative monocyte and neutrophil function. Archives of Surgery 1994; 129(3): 309-15.
102. Nielsen HJ, Nielsen H, Moesgaard F, et al. The effect of ranitidine on cellular immunity in patients with multiple myeloma. Cancer Immunol Immunother 1990; 32(3): 201-5.
103. Nielsen HJ, Nielsen HI, Jensen S, Moesgaard FA. [The effect of ranitidine on postoperative monocyte and neutrophil granulocyte function]. Ugeskrift for Laeger 1995; 157(44): 6119-24.
104. Nielsen HJ, Nielsen HI, Jensen S, Moesgaard FA. The effect of ranitidine on postoperative monocyte and neutrophil granulocyte function. [Danish] Effekten Af Ranitidin Paa Den Postoperative Monocyt- Og Neutrofile Granulocytfunktion. Ugeskrift for Laeger 1995; 157(44): 6119-6124.
105. Nielsen HJ, Nielsen HI, Jensen S, Moesgaard FA The effect of ranitidine on postoperative monocyte and neutrophil granulocyte function. 1995. 157, 6119-6124.
106. Nielsen HJ, Pedersen BK, Moesgaard F, Haahr PM, Kehlet H. Effect of ranitidine on postoperative suppression of natural killer cell activity and delayed hypersensitivity. Acta Chirurgica Scandinavica 1989; 155(8): 377-82.
107. Nikcevic G, Raspopovic S, Pejic M, et al. Prognostic implications of acute gastrointestinal bleeding in acute coronary syndrome: Intrahospital follow up. European Heart Journal 2011; 32: 100.
108. Or?owski T, Smogorzewski M The influence of the H-2 receptor antagonist ranitidine on cadaver kidney allograft function. 1986. 18, 205-210.
109. Peloso LJ, Faria PN, Bossolani MV, De Oliveira HB, Filho SRF. The serum concentration of tacrolimus after ingesting omeprazole: A pilot study. Transplantation 2014; 98(6): e63-e64.
110. Pescina J. Use of cimetidine in the prevention of stress ulcers in intensive care units. Clinical therapeutics 1981; 3(6): 453-5.
111. Philbin DM, Moss J, Akins CW, et al. The use of H1 and H2 histamine antagonists with morphine anesthesia: a double-blind

study. Anesthesiology 1981; 55(3): 292-6.
112. Pinilla JC. Does antacid prophylaxis prevent upper, gastrointestinal bleeding in critically ill patients? Crit Care Med 1985; 13(8): 646-50.
113. Popovic J, Cameron JS. Effects of ranitidine on renal function in transplant recipients. Nephrology Dialysis Transplantation 1990; 5(11): 980-1.
114. Powell H, Morgan M, Li SK, Baron JH. Inhibition of gastric acid secretion in the intensive care unit after coronary artery bypass graft. A pilot control study of intravenous omeprazole by bolus and infusion, ranitidine and placebo. Theoretical Surgery 1993; 8(3): 125-130.
115. Ren YH, Zhao M, Chen YD, et al. Omeprazole affects clopidogrel efficacy but not ischemic events in patients with acute coronary syndrome undergoing elective percutaneous coronary intervention. 2011; 124(6): 856-61.
116. Reusser P, Gyr K, Scheidegger D, et al. Prospective endoscopic study of stress erosions and ulcers in critically ill neurosurgical patients: current incidence and effect of acid-reducing prophylaxis. Critical care medicine 1990; 18(3): 270-4.
117. Rixen D, Livingston DH, Loder P, Denny TN. Ranitidine improves lymphocyte function after severe head injury: results of a randomized, double-blind study. Critical Care Medicine 1996; 24(11): 1787-92.
118. Rixen D, Livingston DH, Loder P, Denny TN. Ranitidine improves lymphocyte function after sever head injury: Results of a randomized, double-blind study. Critical Care Medicine 1996; 24(11): 1787-1792.
119. Rohde H, Lorenz W, Fischer M. [A randomized clinical study of stress ulcer prophylaxis with cimeytidine in severe multiple injuries]. Zeitschrift fur Gastroenterologie 1980; 18(6): 328-9.
120. Rohde H, Lorenz W, Fischer M A randomized clinical study of stress ulcer prophylaxis with cimeytidine in severe multiple injuries. 1980. 18, 328-329.
121. Rohde H, Lorenz W, Fischer M Sequential trial on cimetidine prophylaxis in severe polytrauma. 1980. 18, 328-329.
122. Ruizsantana S, Ortiz E, Gonzalez B, et al. STRESS-INDUCED GASTRODUODENAL LESIONS AND TOTAL PARENTERAL-NUTRITION IN CRITICALLY ILL PATIENTS - FREQUENCY, COMPLICATIONS, AND THE VALUE OF PROPHYLACTIC TREATMENT - A PROSPECTIVE, RANDOMIZED STUDY. Critical Care Medicine 1991; 19(7): 887-891.
123. Sadawarte SM, Vaidyanathan P, Sareen R Evaluating efficacy of single dose intravenous esomeprazole and pantoprazole on gastric pH and volume: a double blind study. 2009. 25, 217-220.
124. Scheiman J, Agewall S, Svedberg LE, Naucler E, Nagy P. Prevention of peptic ulcers with once-daily esomeprazole 20 mg and 40 mg in low-dose acetylsalicylic acid users at gastrointestinal risk: Outcome analysis by cardiovascular risk (OBERON). European Heart Journal 2011; 32: 323.
125. Schiessel R, Starlinger M, Wolf A Failure of cimetidine to prevent gastroduodenal ulceration and bleeding after renal transplantation. 1981. 90, 456-458.
126. Selvanderan SP, Summers M, Finnis M, et al. Administration of stress ulcer prophylaxis may cause harm in critically ill patients: A randomized double blind exploratory study. Gastroenterology 2016; 1): S882.
127. Selvanderan SP, Summers M, Finnis M, et al. Administration of Stress Ulcer Prophylaxis May Cause Harm in Critically Ill Patients: A Randomized Double Blind Exploratory Study. Gastroenterology 2016; 150(4, Suppl. 1): S882.
128. Selvanderan SP, Summers MJ, Plummer MP, et al. Withholding stress ulcer prophylaxis to mechanically ventilated enterally-fed critically ill patients appears safe: A randomised double-blind placebo controlled pilot study. Intensive Care Medicine Experimental 2015; 3 (no pagination)(A41).
129. Selvanderan SP, Summers MJ, Plummer MP, et al Withholding stress ulcer prophylaxis to mechanically ventilated enterally-fed critically ill patients appears safe: A randomised double-blind placebo controlled pilot study. 2015. 3, DOI: 10.1186/2197-425X-3-S1-A41.
130. Silvestri N CMMUdPPBFMG. Cimetidine to prevent stress ulcers. Lancet 1980; 8173: 885.
131. Soria F JPHRMVH. [Incidence of acute gastroduodenal lesions in patients with intracranial lesions. Prophylactic value of treatment with ranitidine]. Rev Fac Cienc Med Cordoba 1985; 1(32-5).
132. Stremple Jf MMDMJJMHD. Post-traumatic gastric bleeding. Archs surg 1972; 105: 177-185.
133. Summers MJ, Selvanderan SP, Plummer MP, et al. Comparison of macroscopic abnormalities in patients receiving routine pantoprazole when compared to placebo. Intensive Care Medicine Experimental 2015; 3 (no pagination)(A980).
134. Wajima Z, Shitara T, Inoue T, Yoshikawa T, Ogawa R The effect of previous administration of nizatidine on the neuromuscular effects of vecuronium and the effect of nizatidine on gastric secretion. 2000. 28, 46-48.
135. Walter S, Thorup Andersen J, Christensen U. Effect of cimetidine on upper gastrointestinal bleeding after renal transplantation: A prospective study. British Medical Journal 1984; 289(6453): 1175-1176.
136. Wei P, Zhang YG, Ling L, et al. Effects of the short-term application of pantoprazole combined with aspirin and clopidogrel in the treatment of acute STEMI. Experimental and Therapeutic Medicine 2016; 12(5): 2861-2864.
137. Whellan DJ, Goldstein JL, Cryer BL, et al. PA32540 (a coordinated-delivery tablet of enteric-coated aspirin 325 mg and immediate-release omeprazole 40 mg) versus enteric-coated aspirin 325 mg alone in subjects at risk for aspirin-associated gastric ulcers: Results of two 6-month, phase 3 studies. American Heart Journal 2014; 168(4): 495-502.e4.
138. Whellan DJ, Goldstein JL, Scheiman JM, et al. Treatment continuation and cardiovascular safety of antiplatelet therapy with pa32540, a tablet with enteric-coated aspirin and immediate-release omeprazole: Results of two 6-month, phase 3 studies. Circulation. Conference: American Heart Association 2012; 126(21 SUPPL. 1).
139. Wu H, Jing Q, Wang J, Guo X. Pantoprazole for the prevention of gastrointestinal bleeding in high-risk patients with acute coronary syndromes. Journal of Critical Care 2011; 26(4): 434.e1-434.e6.
140. Yang HX, Luo D, Zou YY. [Effect of ranitidine on the gastric acid, plasma endothelin, and calcitonin gene-related peptide in patients undergoing the brain operation]. Zhong Nan da Xue Xue Bao. Yi Xue Ban = Journal of Central South University. Medical Sciences 2007; 32(2): 295-8.
141. Yang HX, Luo D, Zou YY. Effect of ranitidine on the gastric acid, plasma endothelin, and calcitonin gene-related peptide in patients undergoing the brain operation. [Chinese]. Journal of Central South University (Medical Sciences) 2007; 32(2): 295-298.
142. Yang HX, Luo D, Zou YY Effect of ranitidine on the gastric acid, plasma endothelin, and calcitonin gene-related peptide in patients undergoing the brain operation. 2007. 32, 295-298.
143. Yao HY, Liu XJ, Cheng K, et al. Early enteral nutrition combined with omeprazole for prevention of stress ulcer in patients with cerebral apoplexy. [Chinese]. World Chinese Journal of Digestology 2017; 25(11): 1016-1020.
144. Yoshihisa A, Takiguchi M, Kanno Y, et al. Associations of Acid Suppressive Therapy With Cardiac Mortality in Heart Failure Patients. Journal of the American Heart Association 2017; 6(5): 9.
145. Yu L, Wang Q, Sha W. Prevention effect of PPIs on induced gastrointestinal injury in CHD patients. Journal of Gastroenterology and Hepatology 2013; 28: 13.
146. Zach GA, Gyr KE, Von Alvensleben E, et al. A DOUBLE-BLIND RANDOMIZED CONTROLLED STUDY TO INVESTIGATE THE EFFICACY OF CIMETIDINE GIVEN IN ADDITION TO CONVENTIONAL THERAPY IN THE PREVENTION OF STRESS ULCERATION AND HEMORRHAGE IN PATIENTS WITH ACUTE SPINAL INJURY. Digestion 1984; 29(4): 214-222.
147. Zhang JR, Wang DQ, Du J, et al. Efficacy of Clopidogrel and Clinical Outcome When Clopidogrel Is Coadministered With Atorvastatin and Lansoprazole: A Prospective, Randomized, Controlled Trial. Medicine 2015; 94(50): e2262.
148. Zumtobel V, Teichmann RK, Inthorn D. [Prevention and therapy of gastroduodenal stress hemorrhage in intensive care patients using the histamine H2-receptor antagonist cimetidine]. Chirurgisches Forum fur Experimentelle und Klinische Forschung 1979: 247-50.
149. Zumtobel V, Teichmann RK, Inthorn D On the prophylaxis and therapy of gastro-duodenal stress haemorrhaging using the

Histamin-H2 receptor antagonist Cimetidin in intensive care patients. 1979. Suppl, 247-250.
150. Zumtobel V, Teichmann RK, Inthorn D Prevention and therapy of gastroduodenal stress bleeding with cimetidine]. [German. 1980. 18, 330-333.
151. Zäch GA, Gyr KE, Alvensleben E, et al A double-blind randomized, controlled study to investigate the efficacy of cimetidine given in addition to conventional therapy in the prevention of stress ulceration and haemorrhage in patients with acute spinal injury. 1984. 29, 214-222.

ONGOING STUDIES

NCT 03374800 – Re-evaluationg the inhibition of stress

NCT03452865 – PPI-SEPSIS

Table S6. Results of analyses on primary/secondary outcomes, subgroups analyses and sensitivity analyses

Results of primary and secondary outcomes							
Outcome or subgroup	**Stu-dies**	**Partici-pants**	**Statistical method**	**95% CI effect estimated in conventional meta-analysis**	**Bayes factor**	**TSA adjusted CI (α-spending adjusted CI)**	**Boundary identifier**
All-cause mortality	28	5656	FEM	1.01 [0.93, 1.10]	941833	0.93, 1.10	Alpha 3.3% beta 10% RRR 20% CEP 26.7% Diversity 0%
						0.84, 1.20	Confidence limit closest to null effect: RRR 7% (RRR 7% alpha 3.3% beta 10% CEP 26.7% Diversity 0%)
						0.89, 1.13	Diversity set to 20% in REM (Diversity 20% alpha 3.3% beta 10% RRR 20% CEP 26.7)
						0.91, 1.13	Incl. trials with no events[b] (Alpha 3.3% beta 10% RRR 20% CEP 26.7% Diversity 0%)
All-cause mortality - Overall low risk of bias	3	3587	FEM	1.03 [0.94, 1.14]	239649	0.94, 1.14	Alpha 3.3% beta 10% RRR 20% CEP 30% Diversity 0%
Any gastrointestinal bleeding	39	6627	REM	0.51 [0.39, 0.67]	$4x10^{-9}$	0.31, 0.84	RRR 20% alpha 3.3% beta 10% CEP 12.26 % Diversity 66.84%
						0.38, 0.69	Confidence limit closest to null effect: RRR 33% (RRR 33% alpha 3.3% beta 10% CEP 12.26% Diversity 66.84%)
						0.36, 0.86	Incl. trials with no events[b] (RRR 20% alpha 3.3% beta 10% CEP 12.26 % Diversity 66.84%)
Any gastrointestinal bleeding	39	6627	FEM	0.52 [0.45, 0.61]	$9x10^{-9}$	0.39, 0.68	RRR 20% alpha 3.3% beta 10% CEP 12.26 % Diversity 66.84%
						0.35, 0.75	Confidence limit closest to null effect: RRR 39% (RRR 39% alpha 3.3% beta 10% CEP 12.26% Diversity 66.84%)
						0.41, 0.78	Incl. trials with no events[b] (RRR 20% alpha 3.3% beta 10% CEP 12.26 % Diversity 66.84%)
Any gastrointestinal bleeding - Overall low risk of bias	3	3596	FEM	0.60 [0.47, 0.77]	0.004	0.36, 1.00	Alpha 3.3% beta 10% RRR 20% CEP 8.74% Diversity 0%
Clinically important gastrointestinal bleeding	14	4833	FEM	0.63 [0.48, 0.81]	0.017	0.35, 1.13	Alpha 3.3% beta 10% RRR 20% CEP 5.41% Diversity 0%
Clinically important gastrointestinal bleeding	14	4833	REM	0.65 [0.49, 0.85]	0.024	0.35, 1.22	Alpha 3.3% beta 10% RRR 20% CEP 5.41% Diversity 3%
Serious adverse events (highest proportion)	42	6744	FEM	0.92 [0.85, 1.00]	33	0.85, 1.00	Alpha 1.7 beta 10% RRR 20% CEP 26.20 Diversity 0%
Serious adverse events (highest proportion) – Overall low risk of bias	3	3587	FEM	1.03 [0.94, 1.14]	239649	0.94,1.14	Alpha 1.7% beta 10% RRR 20% CEO 30% Diversity 0%
Serious adverse events (cumulated)	42	6748	FEM	0.89 [0.85, 0.93]	0.2	0.85,0.93	Alpha 1.7% beta 10% RRR 20% CEP 50% Diversity 0%
Serious adverse events (cumulated) - Overall low risk of bias	3	3587	REM	1.04 [0.85, 1.26]	$7x10^{20}$	0.64, 1.68	Alpha 1.7 beta 10% RRR 20% CEP 63.13% Diversity 93.35%
Serious adverse	3	3587	FEM	0.94 [0.90, 0.99]	7288838	0.68, 1.58	Alpha 1.7 beta 10% RRR 20% CEP

events (cumulated) - Overall low risk of bias							63.13% Diversity 0%
Pneumonia	16	4951	FEM	1.07 [0.94, 1.21]	7465	0.89, 1.27 0.63, 1.80 0.90, 1.32 0.89, 1.28	Alpha 1.7 beta 10% RRR 20% CEP 14.91% Diversity 0% Confidence limit closest to null effect: RRR 6% (Alpha 1.7 beta 10% RRR 6% CEP 14.91% Diversity 0%) Diversity set to 20% in REM: Alpha 1.7 beta 10% RRR 20% CEP 14.91% Diversity 20% Incl. trials with no events[b] (Alpha 1.7 beta 10% RRR 20% CEP 14.91% Diversity 0%)
Pneumonia - Overall low risk of bias	3	3596	REM	1.01 [0.87, 1.18]	82	0.77, 1.33	Alpha 1.7% beta 10% RRR 20% CEP 15.19 Diversity 0%
Cl. Difficile	4	3698	FEM	0.78 [0.46, 1.34]	0.67	N/A	TSA not possible due to too little information (4.73)
Cl. Difficile - Overall low risk of bias	3	3596	FEM	0.84 [0.48, 1.47]	0.84	N/A	

[a]Primary outcome 96,7%, Secondary outcome: 98,3%
[b]Method: constant, value: 0.001

Results of subgroup analyses

Outcome/subgroup Test-of-interaction	**Studies**	**Participants**	**Statistical method**	**Effect estimate**
All-CAUSE MORTALITY				
Risk of bias Test-of-interaction P = 0.47	28	5656	Risk Ratio (M-H, Fixed, 95% CI)	1.01 [0.93, 1.10]
Overall low risk of bias	3	3587	Risk Ratio (M-H, Fixed, 95% CI)	1.03 [0.94, 1.14]
Overall high risk of bias	25	2069	Risk Ratio (M-H, Fixed, 95% CI)	0.96 [0.82, 1.13]
ICU department Test-of-interaction P = 0.08	28	5589	Risk Ratio (M-H, Random, 95% CI)	1.01 [0.93, 1.10]
Medical ICU	5	447	Risk Ratio (M-H, Random, 95% CI)	1.16 [0.81, 1.66]
Surgical ICU	11	817	Risk Ratio (M-H, Random, 95% CI)	0.76 [0.57, 1.00]
Mixed ICU	12	4325	Risk Ratio (M-H, Random, 95% CI)	1.04 [0.94, 1.14]
Mechanical ventilation Test-of-interaction P = 0.21	29	5656	Risk Ratio (M-H, Random, 95% CI)	1.01 [0.93, 1.09]
Invasive mechanical ventilation	12	1015	Risk Ratio (M-H, Random, 95% CI)	0.96 [0.75, 1.21]
Mixed ventilation	10	4041	Risk Ratio (M-H, Random, 95% CI)	1.05 [0.95, 1.16]
No info. on supplemental oxygen administration	7	600	Risk Ratio (M-H, Random, 95% CI)	0.86 [0.71, 1.06]
No mechanical ventilation	0	0	Risk Ratio (M-H, Random, 95% CI)	Not estimable
PPI versus H2RA Test-of-interaction P = 0.51	29	5656	Risk Ratio (M-H, Random, 95% CI)	1.01 [0.93, 1.09]
PPI	9	4104	Risk Ratio (M-H, Random, 95% CI)	1.02 [0.93, 1.13]
H2RA	20	1552	Risk Ratio (M-H, Random, 95% CI)	0.96 [0.82, 1.13]
Placebo versus no prophylaxis Test-of-interaction P = 0.51	29	5656	Risk Ratio (M-H, Random, 95% CI)	1.01 [0.93, 1.09]
Placebo	22	4966	Risk Ratio (M-H, Random, 95% CI)	1.00 [0.92, 1.10]
No prophylaxis	7	690	Risk Ratio (M-H, Random, 95% CI)	1.15 [0.77, 1.70]
According to dose of PPI: Max 40 mg daily versus >40 mg daily Test-of-interaction P = 0.42	9	4104	Risk Ratio (M-H, Fixed, 95% CI)	1.02 [0.93, 1.13]
Max 40 mg daily	7	3994	Risk Ratio (M-H, Fixed, 95% CI)	1.03 [0.94, 1.14]
>40 mg daily	2	110	Risk Ratio (M-H, Fixed, 95% CI)	0.79 [0.43, 1.48]
Publication year Test-of-interaction P = 0.36	28	5656	Risk Ratio (M-H, Random, 95% CI)	1.01 [0.93, 1.09]

Early trials (1977 to 1993)	14	1081	Risk Ratio (M-H, Random, 95% CI)	0.93 [0.78, 1.12]
Recent trials (1994 to 2018)	14	4575	Risk Ratio (M-H, Random, 95% CI)	1.03 [0.93, 1.13]
GI BLEEDING				
Risk of bias Test-of-interaction P = 0.27	39	6627	Risk Ratio (M-H, Fixed, 95% CI)	0.52 [0.45, 0.61]
Overall low risk of bias	3	3596	Risk Ratio (M-H, Fixed, 95% CI)	0.60 [0.47, 0.77]
Overall high risk of bias	36	3031	Risk Ratio (M-H, Fixed, 95% CI)	0.47 [0.38, 0.57]
ICU department Test-of-interaction P = 0.78	33	6074	Risk Ratio (M-H, Random, 95% CI)	0.51 [0.39, 0.67]
Medical ICU	6	462	Risk Ratio (M-H, Random, 95% CI)	0.54 [0.22, 1.37]
Surgical ICU	15	1233	Risk Ratio (M-H, Random, 95% CI)	0.44 [0.31, 0.62]
Mixed ICU	12	4379	Risk Ratio (M-H, Random, 95% CI)	0.52 [0.34, 0.81]
Mechanical ventilation Test-of-interaction P = 0.68	39	6627	Risk Ratio (M-H, Random, 95% CI)	0.51 [0.39, 0.67]
Invasive mechanical ventilation	17	1174	Risk Ratio (M-H, Random, 95% CI)	0.53 [0.35, 0.81]
Mixed ventilation	10	4255	Risk Ratio (M-H, Random, 95% CI)	0.55 [0.33, 0.89]
No info. on supplemental oxygen administration	12	1198	Risk Ratio (M-H, Random, 95% CI)	0.40 [0.22, 0.73]
No mechanical ventilation	0	0	Risk Ratio (M-H, Random, 95% CI)	Not estimable
PPI versus H2RA Test-of-interaction P = 0.38	39	6627	Risk Ratio (M-H, Random, 95% CI)	0.51 [0.39, 0.67]
PPI	11	4336	Risk Ratio (M-H, Random, 95% CI)	0.57 [0.46, 0.72]
H2RA	28	2291	Risk Ratio (M-H, Random, 95% CI)	0.47 [0.33, 0.69]
Placebo versus no prophylaxis Test-of-interaction P = 0.59	39	6627	Risk Ratio (M-H, Random, 95% CI)	0.51 [0.39, 0.67]
Placebo	28	5540	Risk Ratio (M-H, Random, 95% CI)	0.47 [0.36, 0.61]
No prophylaxis	11	1087	Risk Ratio (M-H, Random, 95% CI)	0.58 [0.29, 1.14]
According to dose of PPI: Max 40 mg daily versus >40 mg daily Test-of-interaction P = 0.18	11	4256	Risk Ratio (M-H, Fixed, 95% CI)	0.57 [0.45, 0.72]
Max 40 mg daily	8	4086	Risk Ratio (M-H, Fixed, 95% CI)	0.59 [0.46, 0.75]
>40 mg daily	3	170	Risk Ratio (M-H, Fixed, 95% CI)	0.35 [0.17, 0.73]
Publication year Test-of-interaction P = 0.19	39	6627	Risk Ratio (M-H, Random, 95% CI)	0.51 [0.39, 0.67]
Early trials (1977 to 1993)	21	1608	Risk Ratio (M-H, Random, 95% CI)	0.40 [0.24, 0.67]
Recent trials (1994 to 2018)	18	5019	Risk Ratio (M-H, Random, 95% CI)	0.59 [0.46, 0.76]
SERIOUS ADVERSE EVENTS – highest event proportion				
Risk of bias	42	6744	Risk Ratio (M-H, Fixed, 95% CI)	0.92 [0.85, 1.00]
Overall low risk of bias	3	3587	Risk Ratio (M-H, Fixed, 95% CI)	1.03 [0.94, 1.14]
Overall high risk of bias	39	3157	Risk Ratio (M-H, Fixed, 95% CI)	0.75 [0.65, 0.86]
SERIOUS ADVERSE EVENTS – added up				
Risk of bias	42	6748	Risk Ratio (M-H, Random, 95% CI)	0.84 [0.75, 0.93]
Overall low risk of bias	3	3587	Risk Ratio (M-H, Random, 95% CI)	1.04 [0.85, 1.26]
Overall high risk of bias	39	3161	Risk Ratio (M-H, Random, 95% CI)	0.78 [0.68, 0.90]
QUALITY OF LIFE				

MYOCARDIAL ISCHEMIA				
HOSPITAL-ACQUIRED PNEUMONIA				
Risk of bias Test-of-interaction P = 0.14	16	4951	Risk Ratio (M-H, Random, 95% CI)	1.09 [0.96, 1.23]
Overall low risk of bias	3	3596	Risk Ratio (M-H, Random, 95% CI)	1.01 [0.87, 1.18]
Overall high risk of bias	13	1355	Risk Ratio (M-H, Random, 95% CI)	1.23 [1.00, 1.51]
ICU department Test-of-interaction P = 0.06	16	4951	Risk Ratio (M-H, Random, 95% CI)	1.09 [0.96, 1.23]
Medical ICU	2	234	Risk Ratio (M-H, Random, 95% CI)	1.74 [1.11, 2.74]
Surgical ICU	8	783	Risk Ratio (M-H, Random, 95% CI)	1.16 [0.92, 1.47]
Mixed ICU	6	3934	Risk Ratio (M-H, Random, 95% CI)	1.00 [0.86, 1.17]
Mechanical ventilation Test-of-interaction P = 0.88	16	4951	Risk Ratio (M-H, Random, 95% CI)	1.09 [0.96, 1.23]
Invasive mechanical ventilation	8	827	Risk Ratio (M-H, Random, 95% CI)	1.20 [0.94, 1.52]
Mixed ventilation	8	4124	Risk Ratio (M-H, Random, 95% CI)	1.16 [0.86, 1.57]
No info. on supplemental oxygen administration	0	0	Risk Ratio (M-H, Random, 95% CI)	Not estimable
No mechanical ventilation	0	0	Risk Ratio (M-H, Random, 95% CI)	Not estimable
PPI versus H2RA Test-of-interaction P = 0.18	16	4951	Risk Ratio (M-H, Random, 95% CI)	1.09 [0.96, 1.23]
PPI	6	3911	Risk Ratio (M-H, Random, 95% CI)	1.02 [0.88, 1.19]
H2RA	10	1040	Risk Ratio (M-H, Random, 95% CI)	1.24 [0.97, 1.58]
Placebo versus no prophylaxis Test-of-interaction P = 0.16	16	4951	Risk Ratio (M-H, Random, 95% CI)	1.09 [0.96, 1.23]
Placebo	13	4597	Risk Ratio (M-H, Random, 95% CI)	1.05 [0.93, 1.20]
No prophylaxis	3	354	Risk Ratio (M-H, Random, 95% CI)	1.52 [0.92, 2.51]
CI. DIFFICILE				
Risk of bias Test-of-interaction P = 0.35	4	3698	Risk Ratio (M-H, Fixed, 95% CI)	0.78 [0.46, 1.34]
Overall low risk of bias	3	3596	Risk Ratio (M-H, Fixed, 95% CI)	0.84 [0.48, 1.47]
Overall high risk of bias	1	102	Risk Ratio (M-H, Fixed, 95% CI)	0.28 [0.03, 2.65]
ICU department Test-of-interaction P = 0.36	4	3698	Risk Ratio (M-H, Random, 95% CI)	0.78 [0.45, 1.35]
Medical ICU	1	102	Risk Ratio (M-H, Random, 95% CI)	0.28 [0.03, 2.65]
Surgical ICU	0	0	Risk Ratio (M-H, Random, 95% CI)	Not estimable
Mixed ICU	3	3596	Risk Ratio (M-H, Random, 95% CI)	0.83 [0.47, 1.47]
Mechanical ventilation Test-of-interaction P = 0.83	4	3698	Risk Ratio (M-H, Random, 95% CI)	0.78 [0.45, 1.35]
Invasive mechanical ventilation	3	407	Risk Ratio (M-H, Random, 95% CI)	0.91 [0.21, 3.85]
Mixed ventilation	1	3291	Risk Ratio (M-H, Random, 95% CI)	0.76 [0.42, 1.38]
No info. on supplemental oxygen administration	0	0	Risk Ratio (M-H, Random, 95% CI)	Not estimable
No mechanical ventilation	0	0	Risk Ratio (M-H, Random, 95% CI)	Not estimable
PPI versus H2RA Test-of-interaction not applicable	4	3698	Risk Ratio (M-H, Random, 95% CI)	0.78 [0.45, 1.35]
PPI	4	3698	Risk Ratio (M-H, Random, 95% CI)	0.78 [0.45, 1.35]
H2RA	0	0	Risk Ratio (M-H, Random, 95% CI)	Not estimable
Placebo versus no prophylaxis Test-of-interaction not applicable	4	3698	Risk Ratio (M-H, Random, 95% CI)	0.78 [0.45, 1.35]
Placebo	4	3698	Risk Ratio (M-H, Random, 95% CI)	0.78 [0.45, 1.35]

No prophylaxis	0	0	Risk Ratio (M-H, Random, 95% CI)	Not estimable

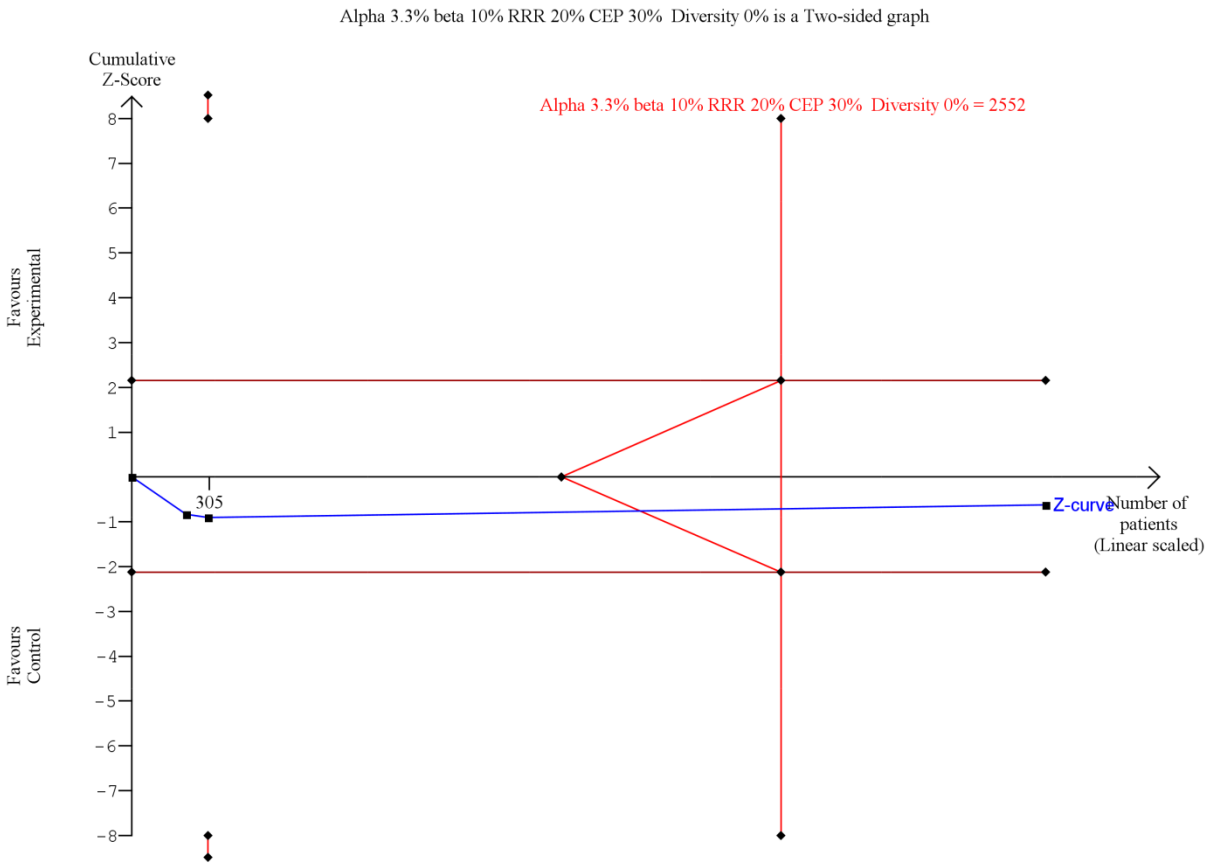
Results of sensitivity analyses

Sensitivity analysis	Studies	Participants	Statistical method	Effect estimate
All-CAUSE MORTALITY				
Best worst-case scenario on missing data	28	5725	Risk Ratio (M-H, Fixed, 95% CI)	0.96 [0.89, 1.05]
Worst best-case scenario on missing data	28	5725	Risk Ratio (M-H, Random, 95% CI)	1.04 [0.96, 1.13]
GI BLEEDING				
Best worst-case scenario on missing data	39	6720	Risk Ratio (M-H, Fixed, 95% CI)	0.47 [0.40, 0.54]
Worst best-case scenario on missing data	39	6720	Risk Ratio (M-H, Random, 95% CI)	0.64 [0.47, 0.87]
Clinically important bleeding	14	4833	Risk Ratio (M-H, Fixed, 95% CI)	0.63 [0.48, 0.81]
SERIOUS ADVERSE EVENTS				
QUALITY OF LIFE				
MYOCARDIAL ISCHEMIA				
HOSPITAL-ACQUIRED PNEUMONIA				
Best worst-case scenario on missing data	16	4953	Risk Ratio (M-H, Fixed, 95% CI)	1.06 [0.93, 1.21]
Worst best-case scenario on missing data	16	4953	Risk Ratio (M-H, Fixed, 95% CI)	1.07 [0.94, 1.22]
CL. DIFFICILE				
Best worst-case scenario on missing data	4	3722	Risk Ratio (M-H, Random, 95% CI)	0.51 [0.12, 2.13]
Worst best-case scenario on missing data	4	3722	Risk Ratio (M-H, Fixed, 95% CI)	1.06 [0.65, 1.75]

CEP: control event proportion; CI: confidence interval; FEM: fixed effects model; GI bleeding: gastrointestinal bleeding; H2RA: histamin-2 receptor antagonist ICU: intensive care unit; M-H: mantel-haenszel; PPI: proton pump inhibitor; prop: proportion; REM: random effects model; RRR: relative risk reduction; SAE: serious adverse events; TSA: trial sequential analysis;

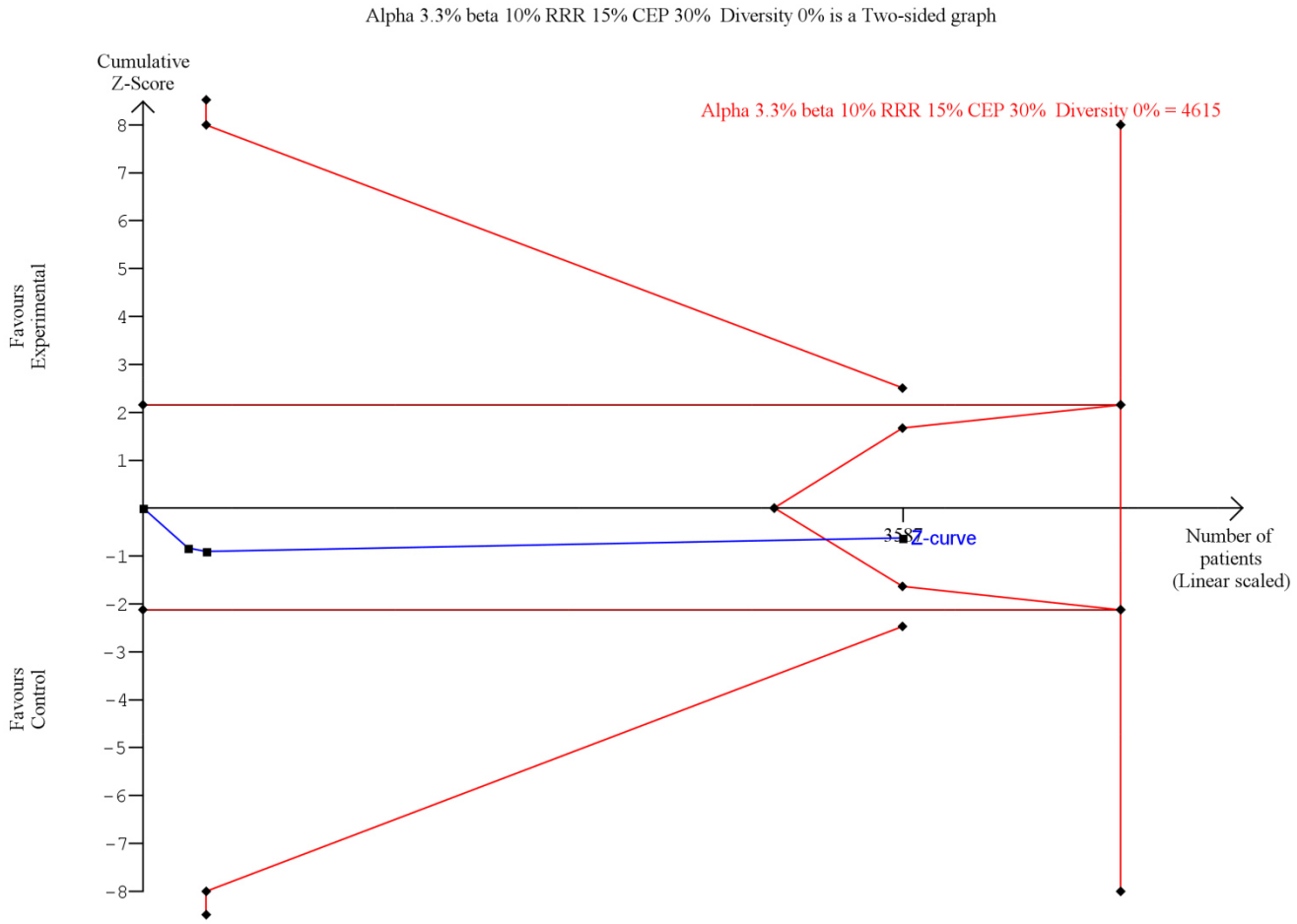
MORTALITY

Fig. S7. TSA of trials with overall low risk of bias for a 20% RRR/RRI on mortality



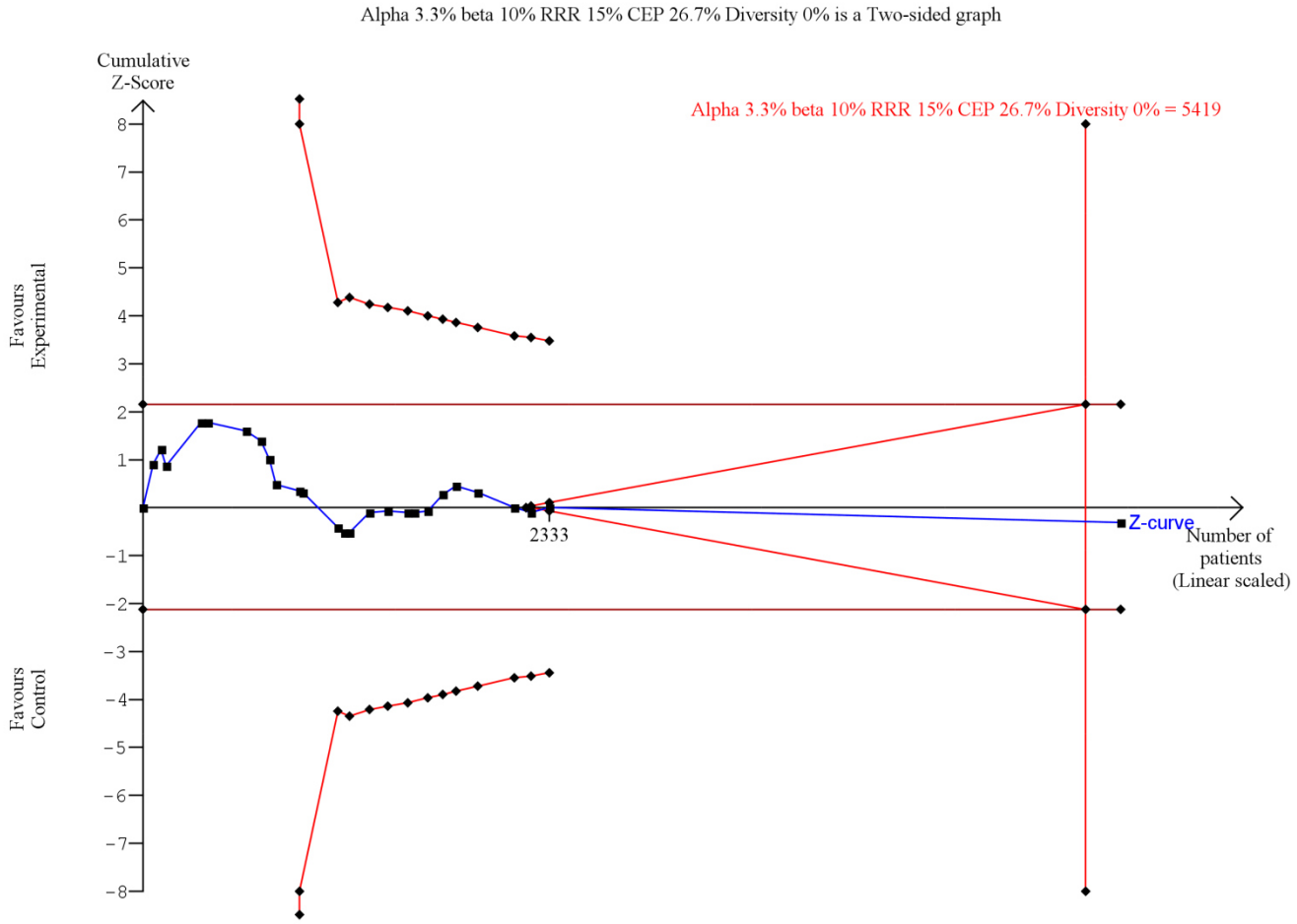
Control event proportion of 30%, diversity (D2) of 0%, alpha of 3.3%, power of 90% and RRR/RRI of 20%. We used an adjusted maximal type 1 error risk (alpha) of 0.033 due to two co-primary outcomes. The RR was 1.03 with a naive 95% CI of 0.94, 1.14 in a fixed effects model and the TSA-adjusted CI 0.94, 1.14. As the cumulative Z-curve reaches required information size we may exclude a 20% RRR/RRI and even a 15% RRR/RRI may be rejected as the CIs exclude these effects.

Fig. S8. TSA of trials with overall low risk of bias for a 15% RRR/RRI on mortality



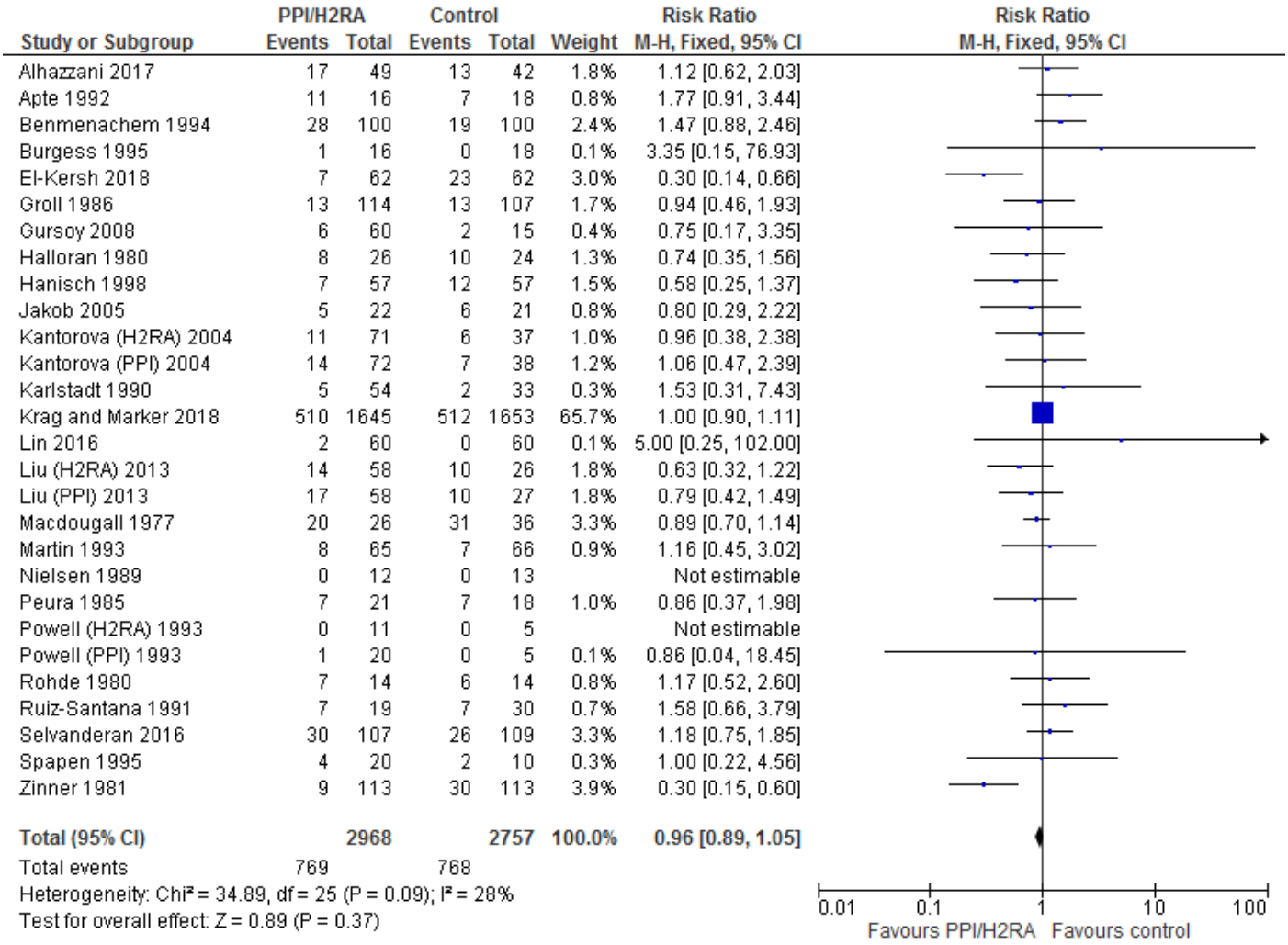
Control event proportion of 30%, diversity (D2) of 0%, alpha of 3.3%, power of 90% and RRR/RRI of 15%. We used an adjusted maximal type 1 error risk (alpha) of 0.033 due to two co-primary outcomes. The RR was 1.03 with a naive 95% CI of 0.94, 1.14 in a fixed effects model and the TSA-adjusted CI 0.91, 1.17. As the cumulative Z-curve reaches futility area for non-inferiority we can exclude a 15% RRR/RRI.

Fig. S9. TSA of all trials for a 15% RRR/RRI on mortality



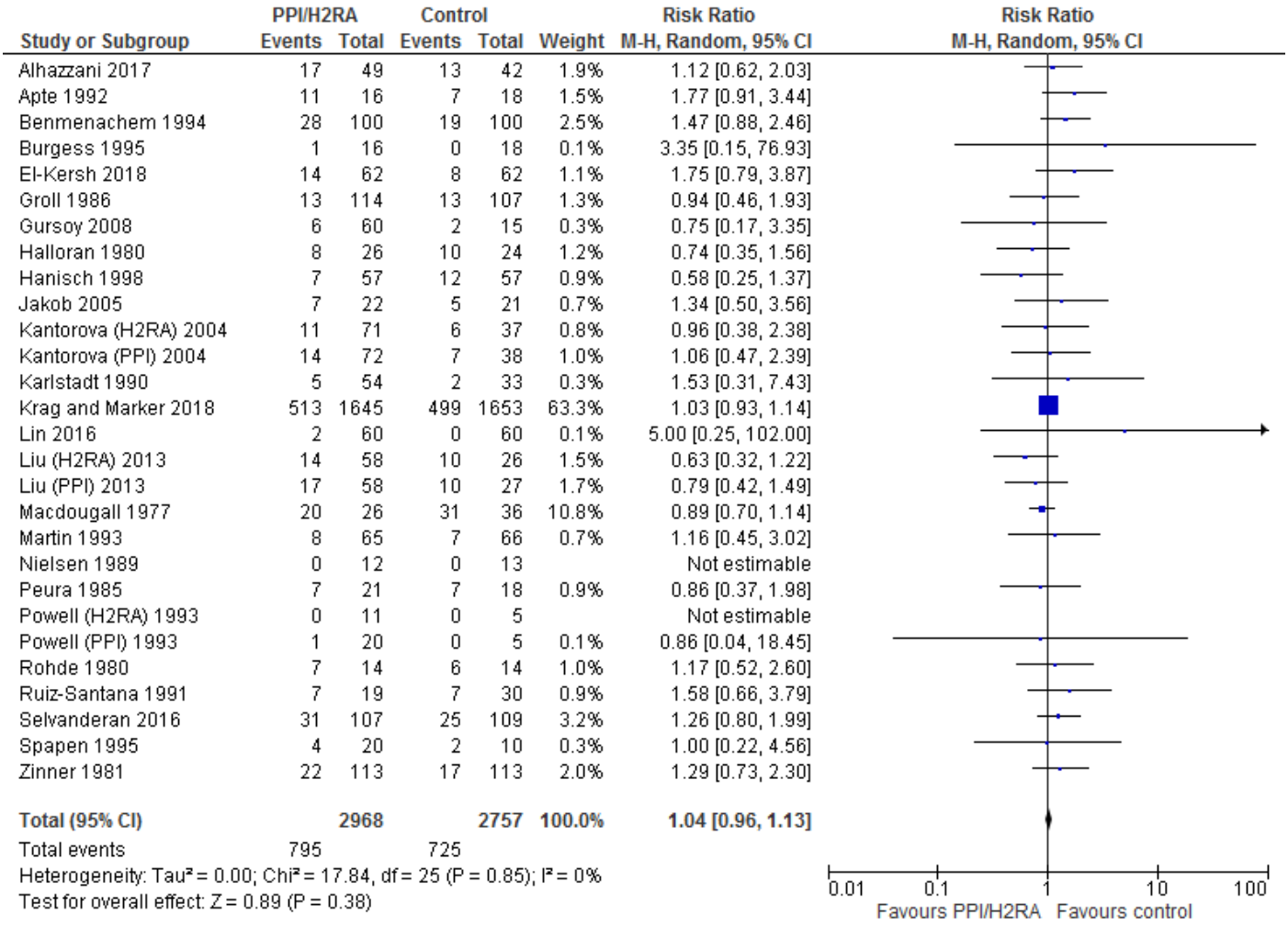
Control event proportion of 26.7%, diversity (D2) of 0%, alpha of 3.3%, power of 90% and RRR/RRI of 15%. We used an adjusted maximal type 1 error risk (alpha) of 0.033 due to two co-primary outcomes. The RR was 1.01 with a naive 95% CI of 0.93, 1.10 in a fixed effects model and the TSA-adjusted CI 0.93, 1.10. As the cumulative Z-curve reaches required information size we can exclude a 15% RRR/RRI.

Fig. S10. Forest plot of best worst-case scenario sensitivity analysis for missing data on mortality



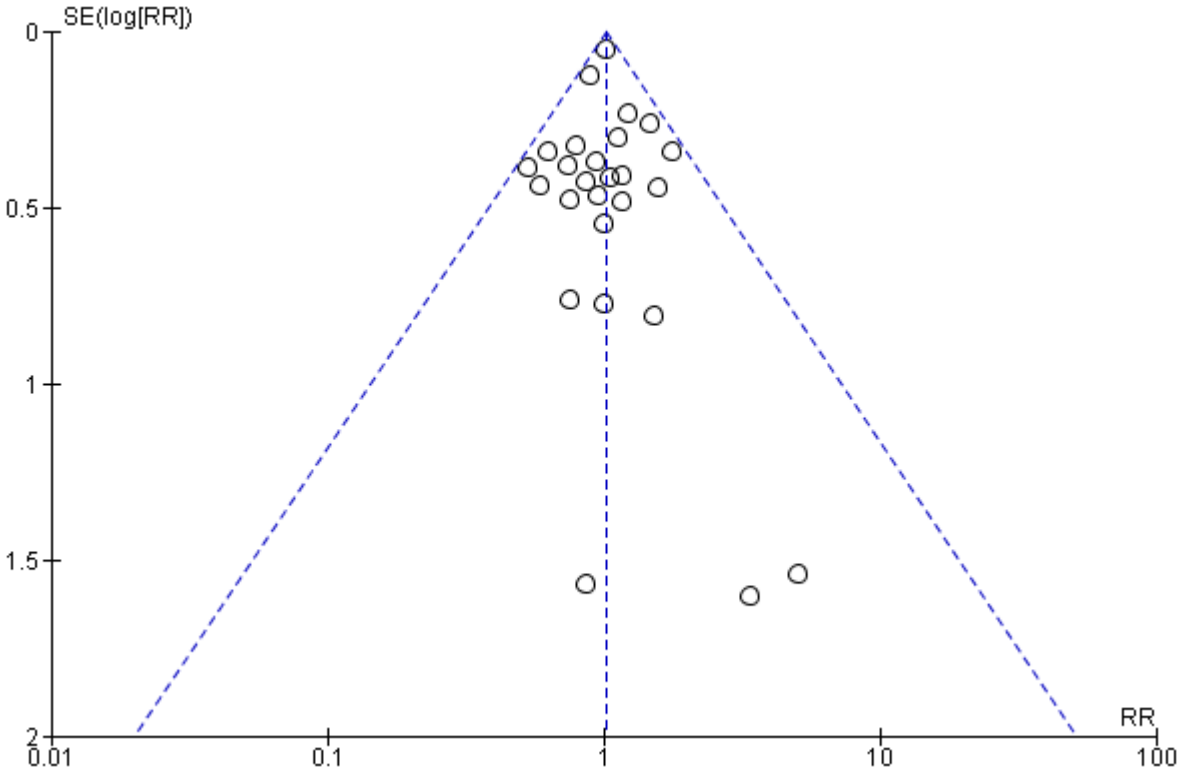
In this analysis it is assumed that all participants lost to follow-up in the PPI/H2RA group have survived and all those with missing outcomes in the control group have died. The primary meta-analysis result (RR 1.01, 95% CI 0.93, 1.10) and the result from this sensitivity analysis (RR 0.96, 95% CI 0.89, 1.05) show similar P values and CIs.

Fig. S11. Forest plot of worst best-case scenario sensitivity analysis for missing data on mortality



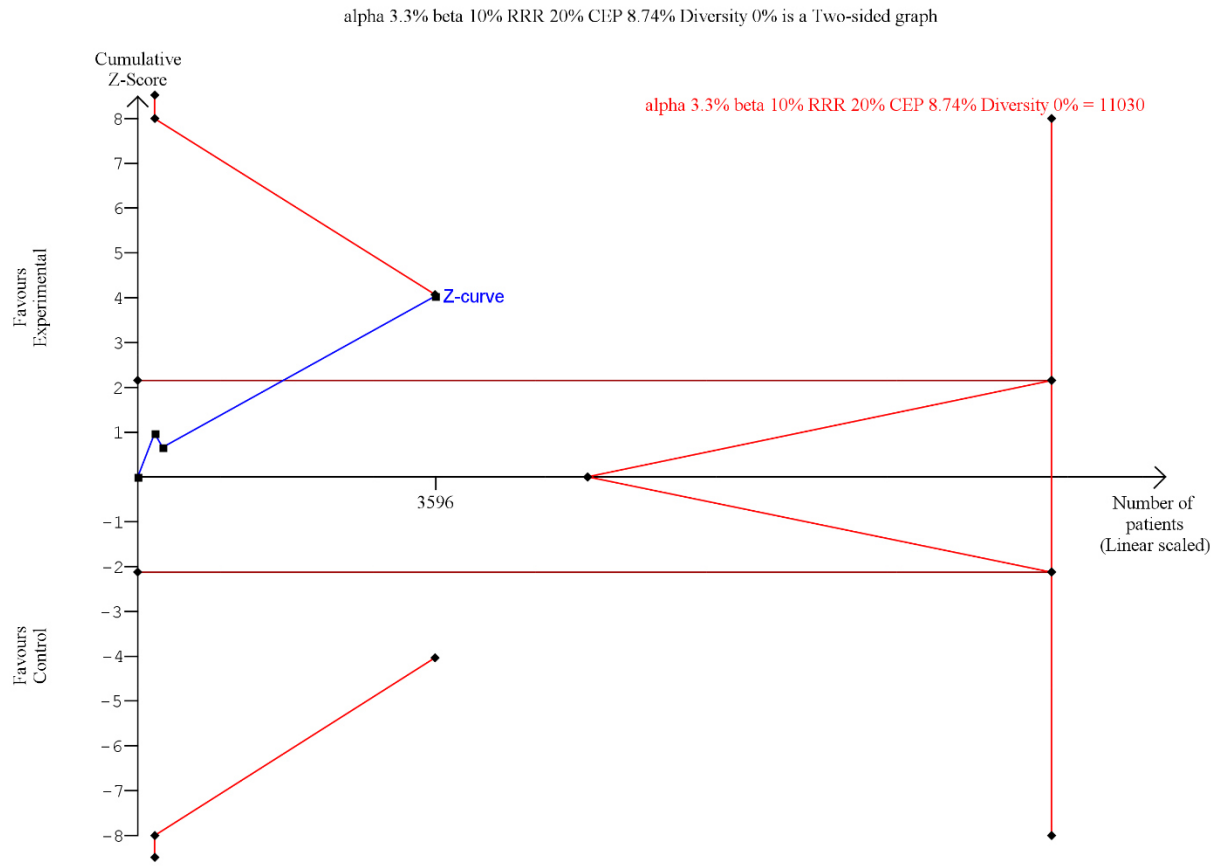
In this analysis it is assumed that all participants lost to follow-up in the PPI/H2RA group died and all those with missing outcomes in the control group have survived. The primary meta-analysis result (RR 1.01, 95% CI 0.93, 1.10) and the result from this sensitivity analysis (RR 1.04, 95% CI 0.96, 1.13) show similar P values and CIs.

Fig. S12. Funnel plot of all-cause mortality



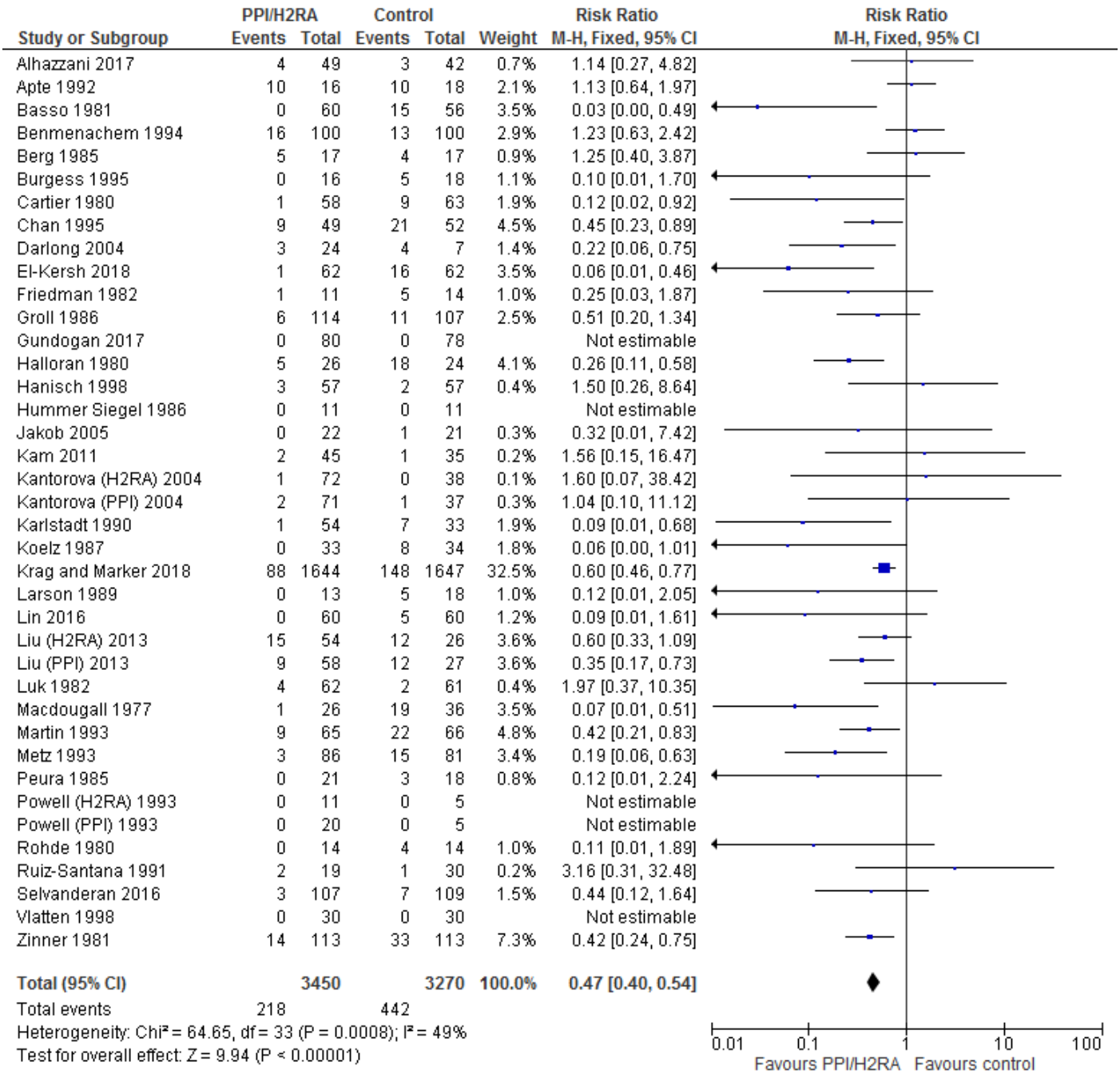
GASTROINTESTINAL BLEEDING

Fig. S13. TSA of trials with overall low risk of bias for a 20% RRR/RRI on any gastrointestinal bleeding



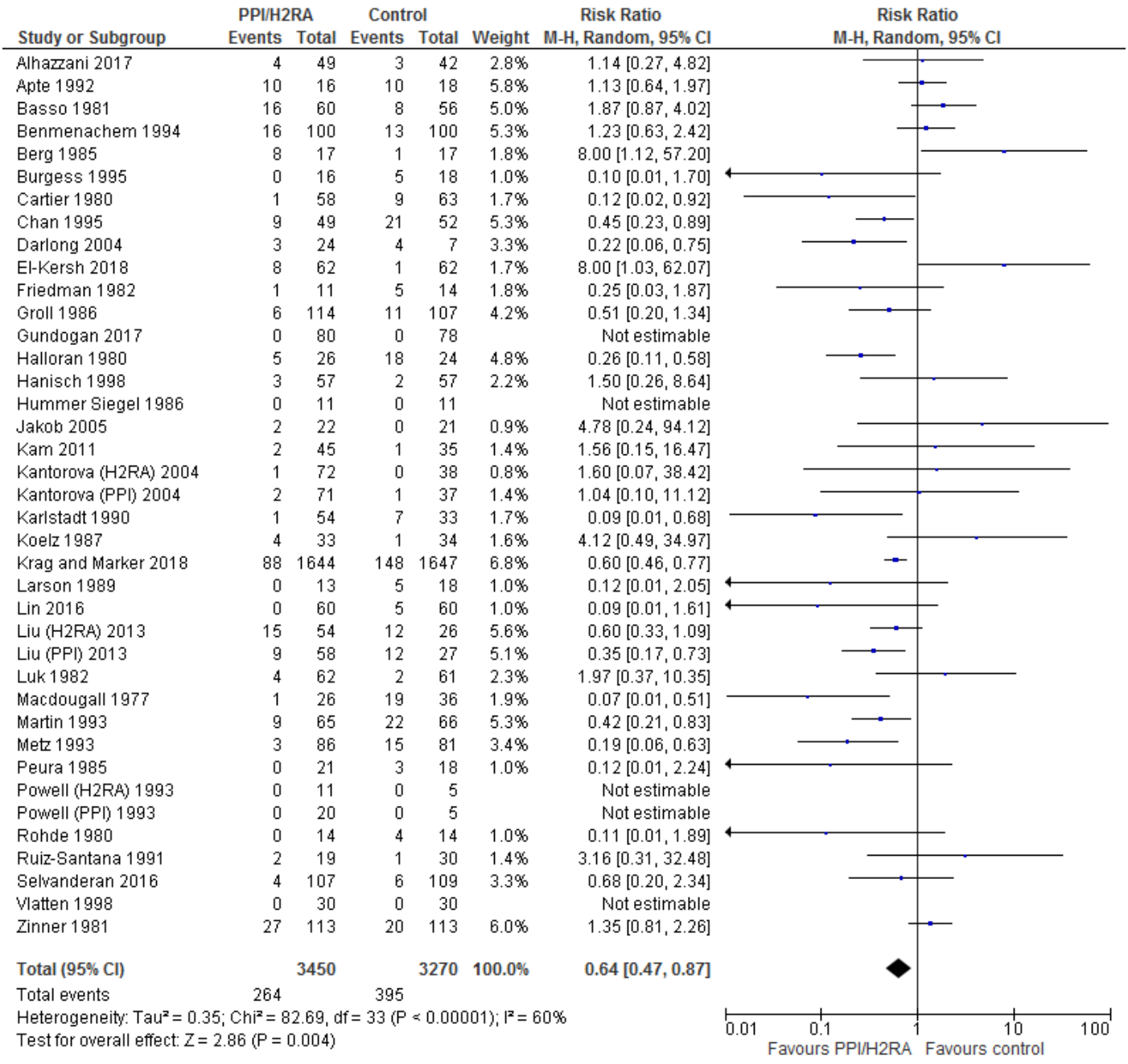
Control event proportion of 8.74%, diversity (D2) of 0%, alpha of 3.3%, power of 90% and RRR of 20%. We used an adjusted maximal type 1 error risk (alpha) of 0.033 due to two co-primary outcomes. The RR was 0.60 with a naive 95% CI of 0.47, 0.77 in a fixed effects model and the TSA-adjusted CI. 0.36, 1.00. As the cumulative Z-curve reaches trial sequential monitoring boundary for benefit we may accept at least a 20% RRR.

Fig. S14. Forest plot of best worst-case scenario sensitivity analysis for missing data on GI bleeding



In this analysis it is assumed that all participants lost to follow-up in the PPI/H2RA group have survived and all those with missing outcomes in the control group have died. The primary meta-analysis result (RR 0.52, 95% CI 0.45, 0.61) and the result from this sensitivity analysis (RR 0.47, 95% CI 0.40, 0.54) show similar P values and CIs.

Fig. S15. Forest plot of worst best-case scenario sensitivity analysis for missing data on GI bleeding



In this analysis it is assumed that all participants lost to follow-up in the PPI/H2RA group died and all those with missing outcomes in the control group have survived. The primary meta-analysis result (RR 0.52, 95% CI 0.45, 0.61) and the result from this sensitivity analysis (RR 0.64, 95% CI 0.47, 0.87) show similar P values and CIs.

Fig. S16. Funnel plot of any gastrointestinal bleeding

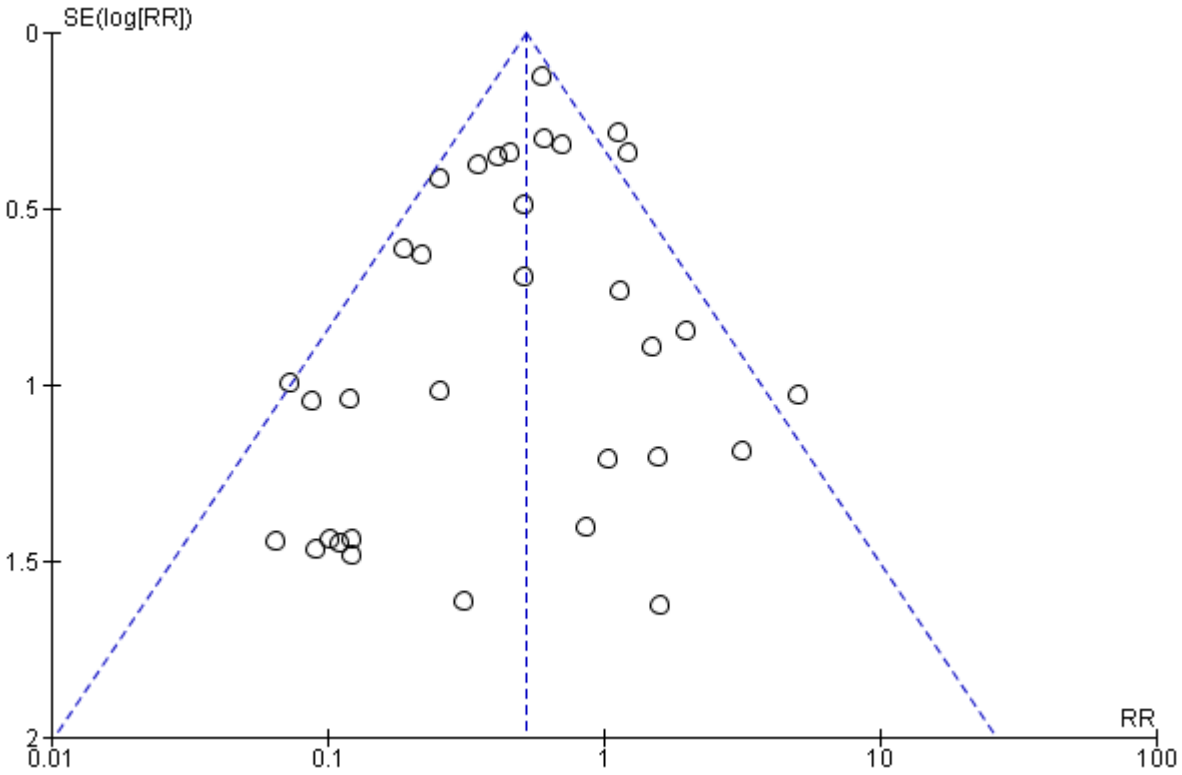
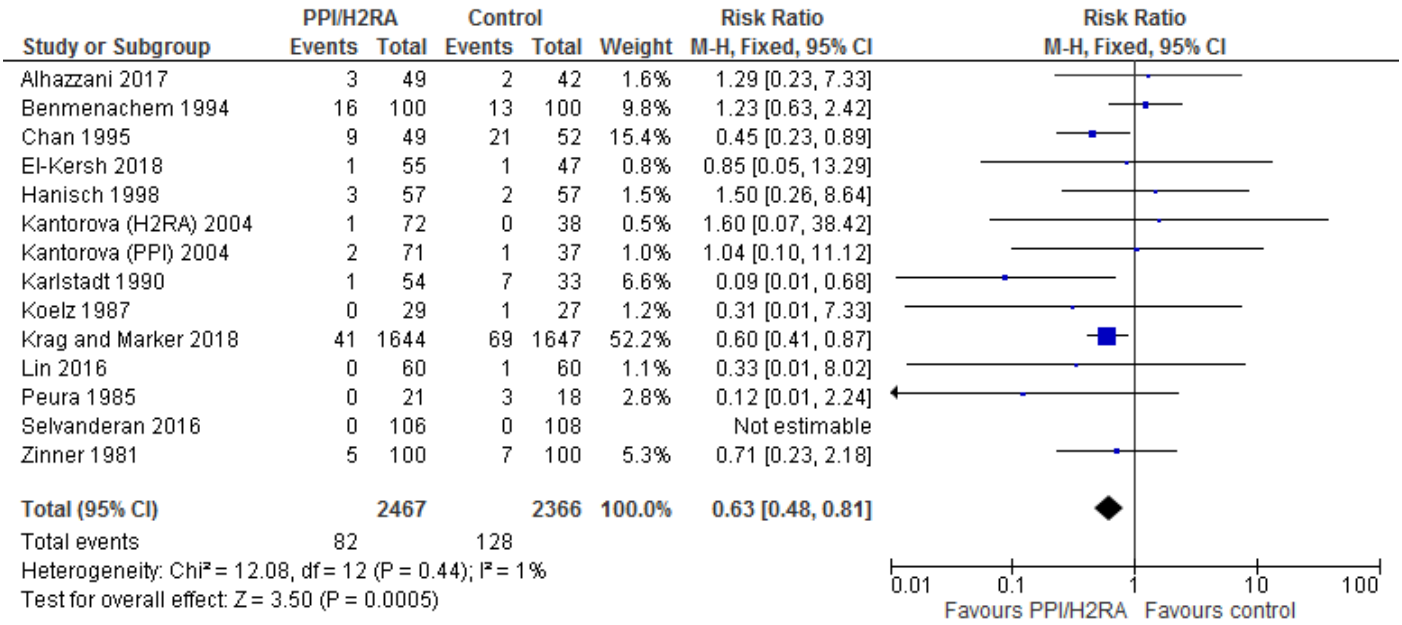
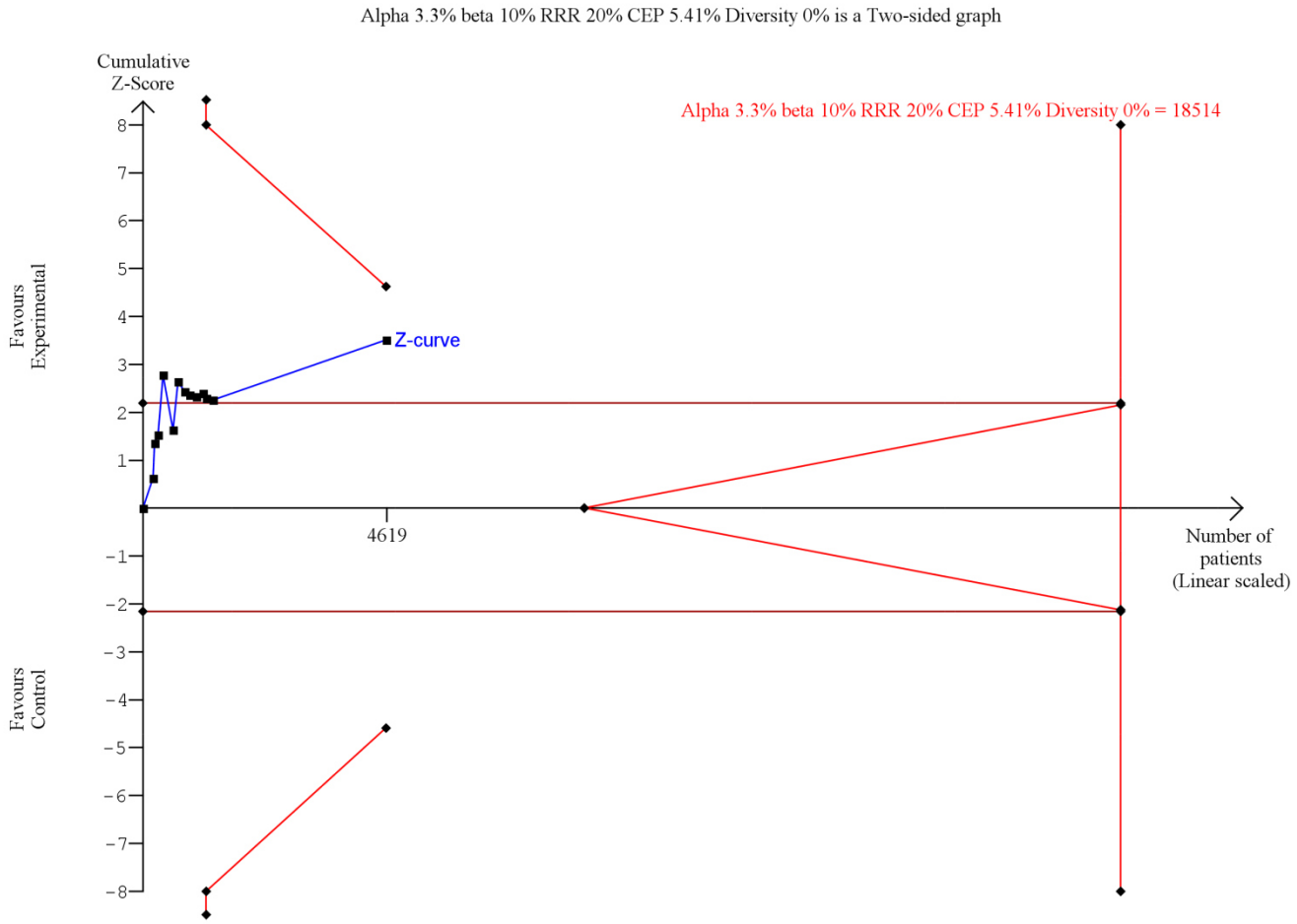


Fig. S17. Forest plot of clinically important gastrointestinal bleeding



A conventional meta-analysis showed that stress ulcer prophylaxis with PPI/H2RA versus placebo/no prophylaxis may reduce GI bleeding.

Fig. S18. TSA of all trials for a 20% RRR/RRI on clinically important gastrointestinal bleeding



Control event proportion of 5.41%, diversity (D2) of 0%, alpha of 3.3%, power of 90% and RRR of 20%. We used an adjusted maximal type 1 error risk (alpha) of 0.033 due to two co-primary outcomes. The RR was 0.63 with a naive 95% CI of 0.48, 0.81 in a fixed effects model and the TSA-adjusted CI. 0.35, 1.13. As the cumulative Z-curve does not reach the trial sequential monitoring boundary, futility area or required information size, we do not have enough information to detect or reject a 20% RRR/RRI.

SERIOUS ADVERSE EVENTS

Additional information on the host-hoc analyses of serious adverse events

A total of 42 trials reported on outcomes categorised as serious adverse events. As the reporting of serious adverse events, as a combined outcome, were not carried out according to the ICH-GCP recommendation, we estimated the number of patients with one or more serious adverse events in two ways (post-hoc analyses):

1) By choosing the one specific serious adverse event with the highest proportion reported in each trial which we will address the lowest possible proportion of patients with one or more SAEs (somehow be a best-case scenario).

2) By cumulating all reported serious adverse events, assuming that patients only experience one serious adverse event (the number of patients in each group will constitute a maximum) we will address the highest possible reported proportion of patients with one or more SAEs (somehow a worst-case scenario).

In the analysis adding all reported serious adverse events, three trials [1-3] had higher number of events than the number of included patients and it is obvious that some patients had more than one event. In a meta-analysis, number of patients having one or more event cannot exceed number of included patients; we therefore changed the number of events to the number of patients actually enrolled estimating that all patients had one or more events in these trials. In the analysis of the highest proportion of serious adverse events in each trial, we included extracted data from the trial with the highest total event proportion, as some trials had higher event proportions in the intervention group in one outcome and lower in another outcome – compared to the control group.

Definition of serious adverse events

- Mortality
- Any GI bleeding (with the exclusion of endoscopically documented asymptomatic lesions)
- Myocardial ischaemia or infarction as defined by trialists
- *Clostridium difficile* infection as defined by trialists
- Nosocomial pneumonia as defined by trialists
- Stroke as defined by trialists
- Anaphylactic reactions as defined by trialists
- Agranulocytosis as defined by trialists
- Pancytopenia as defined by trialists
- Acute hepatic failure as defined by trialists
- Stevens-Johnson syndrome and toxic epidermal necrolysis as defined by trialists
- Interstitial nephritis as defined by trialists
- Angioedema (Quincke's edema) as defined by trialists
- Other serious adverse events according to the Good Clinical Practice Guideline of the International Conference on Harmonization (ICH-GCP) definition.

Results of the analyses of serious adverse events

In the meta-analysis of the estimated highest reported proportion of serious adverse events in the three trials with low risk of bias, we found no difference between stress ulcer prophylaxis and placebo/no prophylaxis: RR 1.03 (95% CI 0.94, 1.14; P = 0.52; I^2 = 0%; TSA-adjusted CI 0.94-1.14; Bayes factor 239649). TSA showed that the required information size to detect a RRR of at least 20% had been reached (S19-21, ESM). The certainty of evidence was judged as low due to very serious indirectness (Table 1).

The corresponding summary effect estimate of all 42 trials regardless of bias reporting on serious adverse events (n = 6744 participants) was RR 0.92 (95% CI 0.85, 1.00; P = 0.05, I^2 = 44%; TSA-adjusted CI: 0.85, 1.00; Bayes factor 33), and TSA showed that the required information size to detect a RRR of at least 20% had been reached (S20, S22, ESM). Harbord's test indicated asymmetry (P = 0.019 (S23, ESM)). The certainty of evidence was very low due to risk of bias, inconsistency, very serious indirectness and suspected publication bias (Table 1).

In the meta-analysis of the estimated cumulated number of serious adverse events in the three trials with overall low risk of bias, we found no difference between stress ulcer prophylaxis and placebo/no prophylaxis: RR 1.04 (95% CI 0.85, 1.26; P = 0.72; I^2 = 53%; TSA-adjusted CI: 0.64, 1.68; Bayes factor $7x10^{20}$). TSA showed that only 29% of the required information size had been reached (S24-27, ESM). The certainty of evidence was very low due to inconsistency, very serious indirectness, and imprecision (Table 1).

The corresponding effect estimate of all 42 trials regardless of bias reporting on serious adverse events (n = 6748 participants) was RR 0.89 (95% CI 0.85, 0.93; P < 0.00001; I^2 = 59%; TSA-adjusted CI 0.85, 0.93; Bayes factor 0.2), and TSA showed that the required information size to detect a RRR of at least 20% had been reached (S26, S28, ESM). Harbord's test did not indicate asymmetry (P = 0.06 (S29, ESM)). The certainty of evidence was very low due to risk of bias, inconsistency and very serious indirectness (Table 1).

SERIOUS ADVERSE EVENTS (HIGHEST PROPORTION OF REPORTED SERIOUS ADVERSE EVENTS IN EACH TRIAL)

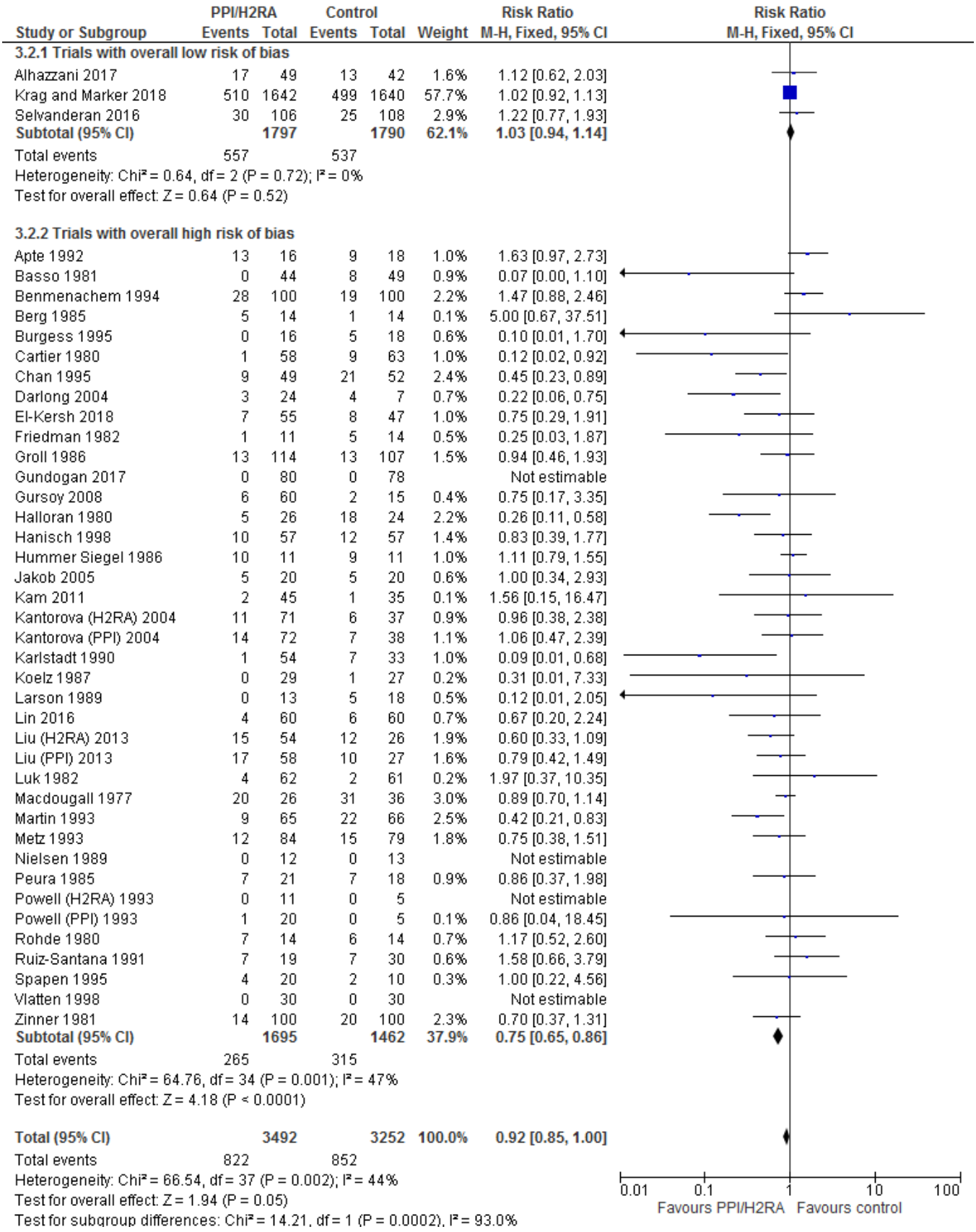
Table S19. Types of serios adverse events, from each trial, included in analysis (highest proportion)

	All-cause mortality				Prop. of pts with CLIN IMP GI bleeding				Prop. of pts with ANY GI bleeding				Prop. of pts with hospital acquired pneumonia			
Trial id	E: No. of events	E: No. analysed	C: No. of events	C: No. analysed	E: No. of events	E: No. analysed	C: No. of events	C: No. analysed	E: No. of events	E: No. analysed	C: No. of events	C: No. analysed	E: No. of events	E: No. analysed	C: No. of events	C: No. analysed
Alhazzani 2017 [4]	17	49	13	42												
Apte 1992 [1]													13	16	9	18
Basso 1981 [5]									0	44	8	49				
Benmenachem 1994 [6]	28	100	19	100												
Berg 1985 [7]									5	14	1	14				
Burgess 1995 [8]									0	16	5	18				
Cartier 1980 [9]									1	58	9	63				
Chan 1995 [10]					9	49	21	52								
Darlong 2003 [11]									3	24	4	7				
El-Kersh 2018 [12]	7	55	8	47												
Friedman 1982 [13]																
Groll 1986 [14]	13	114	13	107												
Gundogan 2017 [15]									0	80	0	78				
Gursoy 2008 [16]	6	60	2	15												
Halloran 1980 [2]									5	26	18	24				
Hanisch 1998 [17]													10	57	12	57
Hummer Sigiel 1986 [18]													10	11	9	11
Jakob 2005 [19]	5	20	5	20												
Kam 2011 [20]									2	45	1	35				
Kantotova H2RA 2004 [21]	11	71	6	37												
Kantorova PPI 2004 [21]	14	72	7	38												

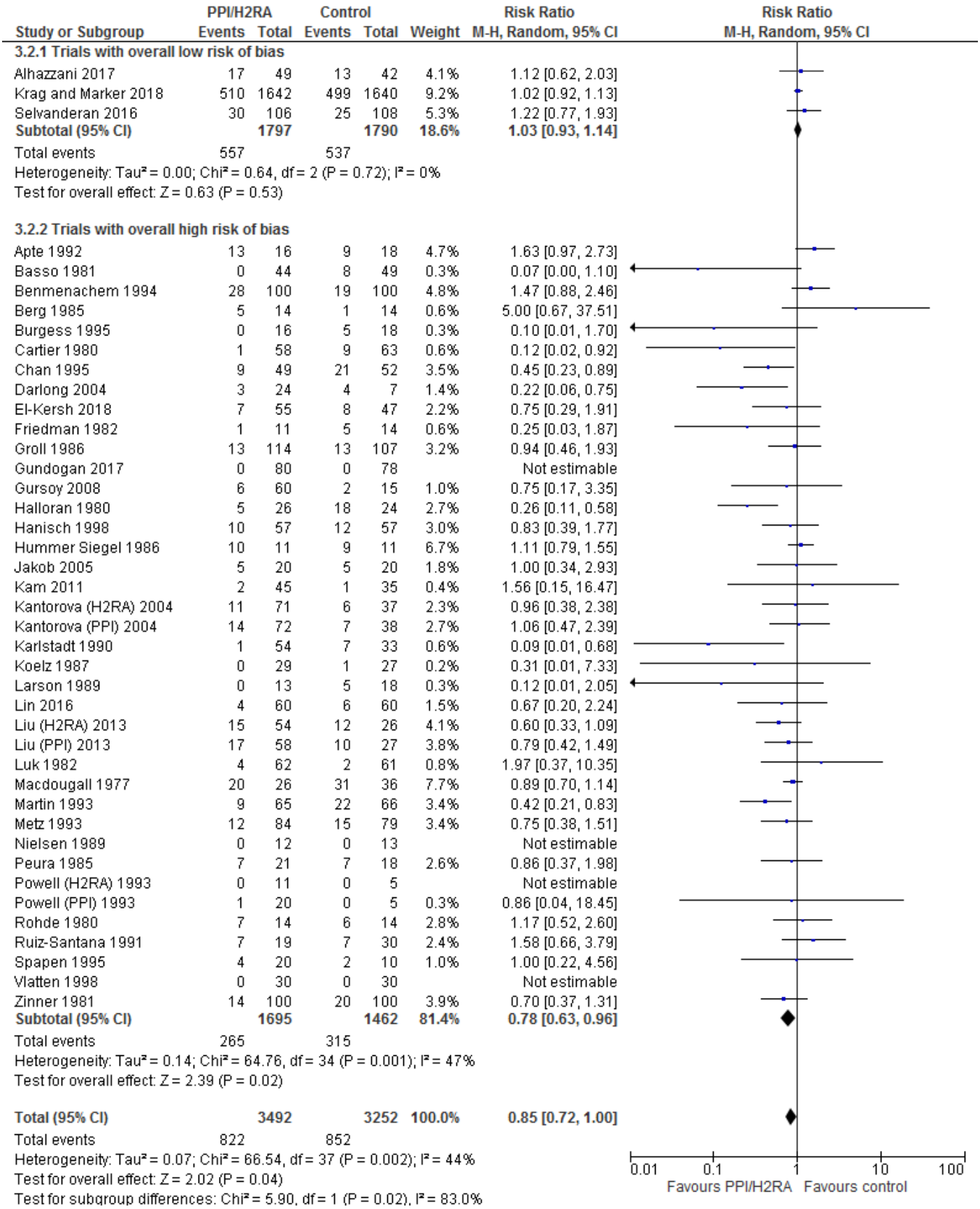
Karlstadt 1990 [22]					1	54	7	33								
Koelz 1987 [23]					0	29	1	27								
Krag & Marker 2018 [24]	510	1642	499	1640												
Larson 1989 [25]									0	13	5	18				
Lin 2016 [26]													4	60	6	60
Liu H2RA 2013 [27]									15	54	12	26				
Liu PPI 2013 [27]	17	58	10	27												
Luk 1982 [28]									4	62	2	61				
MacDougall 1977 [3]	20	26	31	36												
Martin 1993 [29]									9	65	22	66				
Metz 1993 [30]													12	84	15	79
Nielsen 1989 [31]	0	12	0	13												
Peura 1985 [32]	7	21	7	18												
Powell H2RA 1993 [33]	0	11	0	5												
Powell PPI 1993 [33]	1	20	0	5												
Rohde 1980 [34]	7	14	6	14												
Ruiz-Santana 1991 [35]	7	19	7	30												
Selvanderan 2016 [36]	30	106	25	108												
Spapen 1995 [37]	4	20	2	10												
Vlatten 1998 [38]									0	30	0	30				
Zinner 1981 [39]									14	100	20	100				

CLIN IMP GI bleeding: clinically important gastrointestinal bleeding; C: control; E: experimental; No.: number; Prop.: proportion; pts: participants

Fig. S20. Forest plot (fixed and random effects model) of serious adverse event (highest proportion)

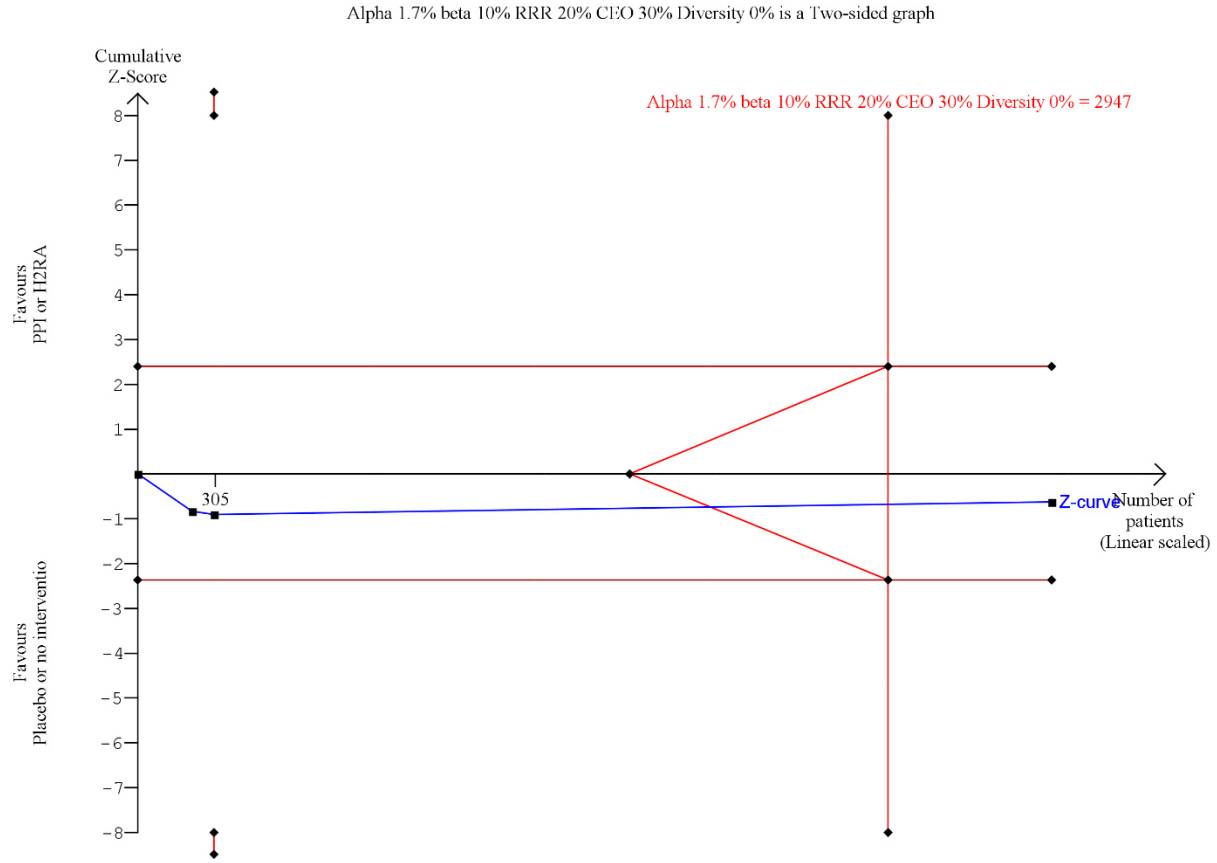


Fixed effects model. Meta-analysis showed that stress ulcer prophylaxis with PPI or H2RA versus placebo/no prophylaxis does not reduce serious adverse events.



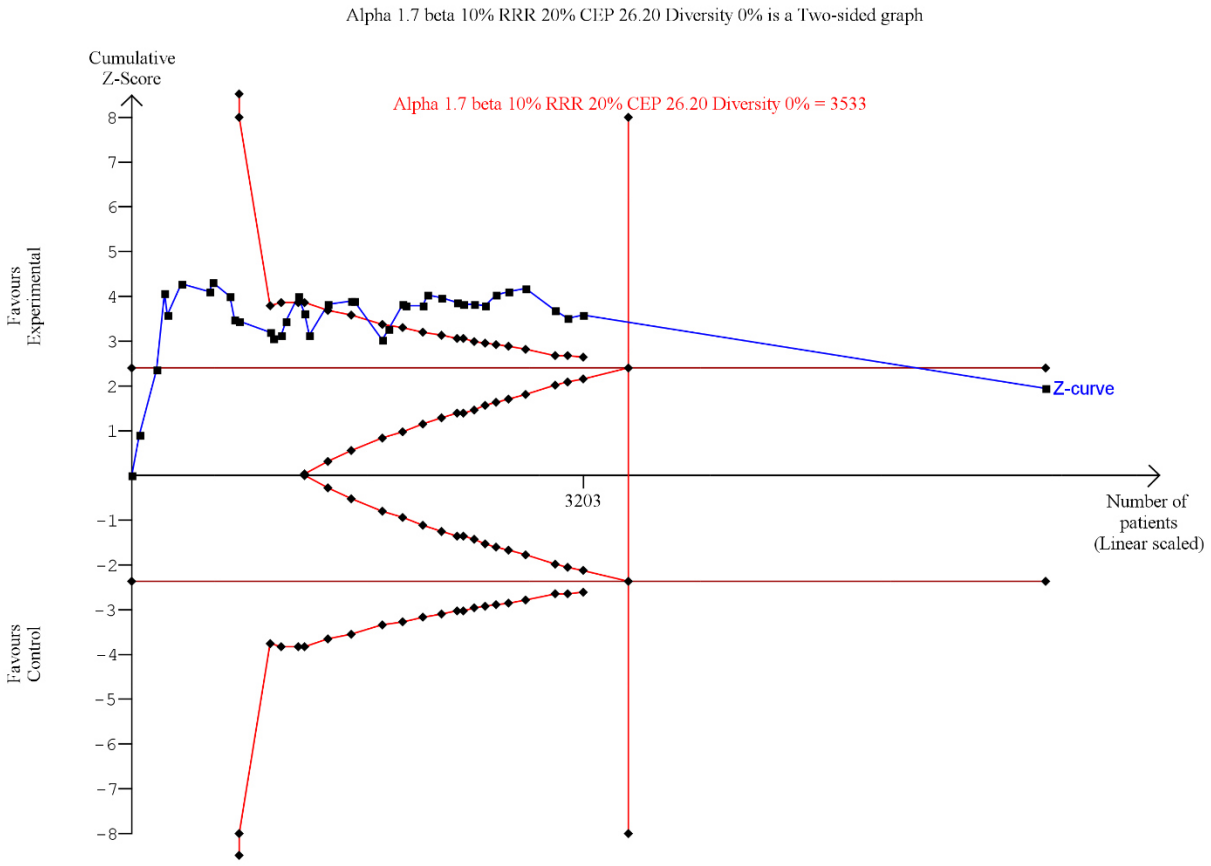
Random effects model. Meta-analysis showed that stress ulcer prophylaxis with PPI or H2RA versus placebo/no prophylaxis does not reduce serious adverse events.

Fig. S21. TSA of trials with overall low risk of bias for a 20% RRR/RRI on serious adverse events (highest proportion)



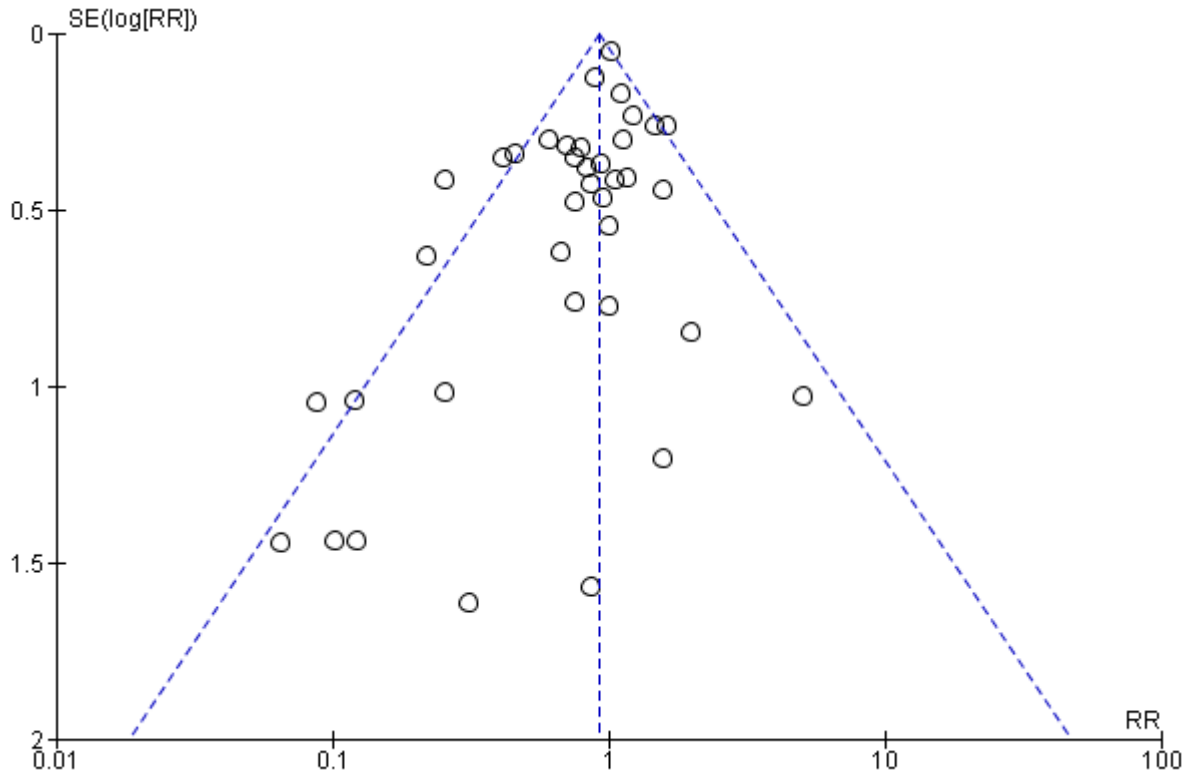
Control event proportion of 30%, diversity (D2) of 0%, alpha of 1.7%, power of 90% and RRR of 20%. We used an adjusted maximal type 1 error risk (alpha) of 0.017 due to five co-secondary outcomes. The RR was 1.03 with a naive 95% CI of 0.94, 1.14 in a fixed effects model and the TSA-adjusted CI. 0.94, 1.14. As the cumulative Z-curve reaches required information size for non-inferiority, we can exclude a 20% RRR/RRI and possible also a 15% RRR/RRI.

Fig. S22. TSA of all trials for a 20% RRR/RRI on serious adverse events (highest proportion)



Control event proportion of 26.20%, diversity (D2) of 0%, alpha of 1.7%, power of 90% and RRR of 20%. We used an adjusted maximal type 1 error risk (alpha) of 0.017 due to five co-secondary outcomes. The RR was 0.92 with a naive 95% CI of 0.85, 1.00 in a fixed effects model and the TSA-adjusted CI. 0.85, 1.00. As the cumulative Z-curve reaches trial sequential monitoring boundary for benefit and required information size, we accept a 20% RRR.

Fig. S23. Funnel plot of serious adverse events (highest proportion)



Funnel plot show signs of asymmetry.

SERIOUS ADVERSE EVENTS (CUMULATED SERIOUS ADVERSE EVENTS IN EACH TRIAL)

Table S24. Types of serious adverse events, from each trial, included in analysis (cumulated)

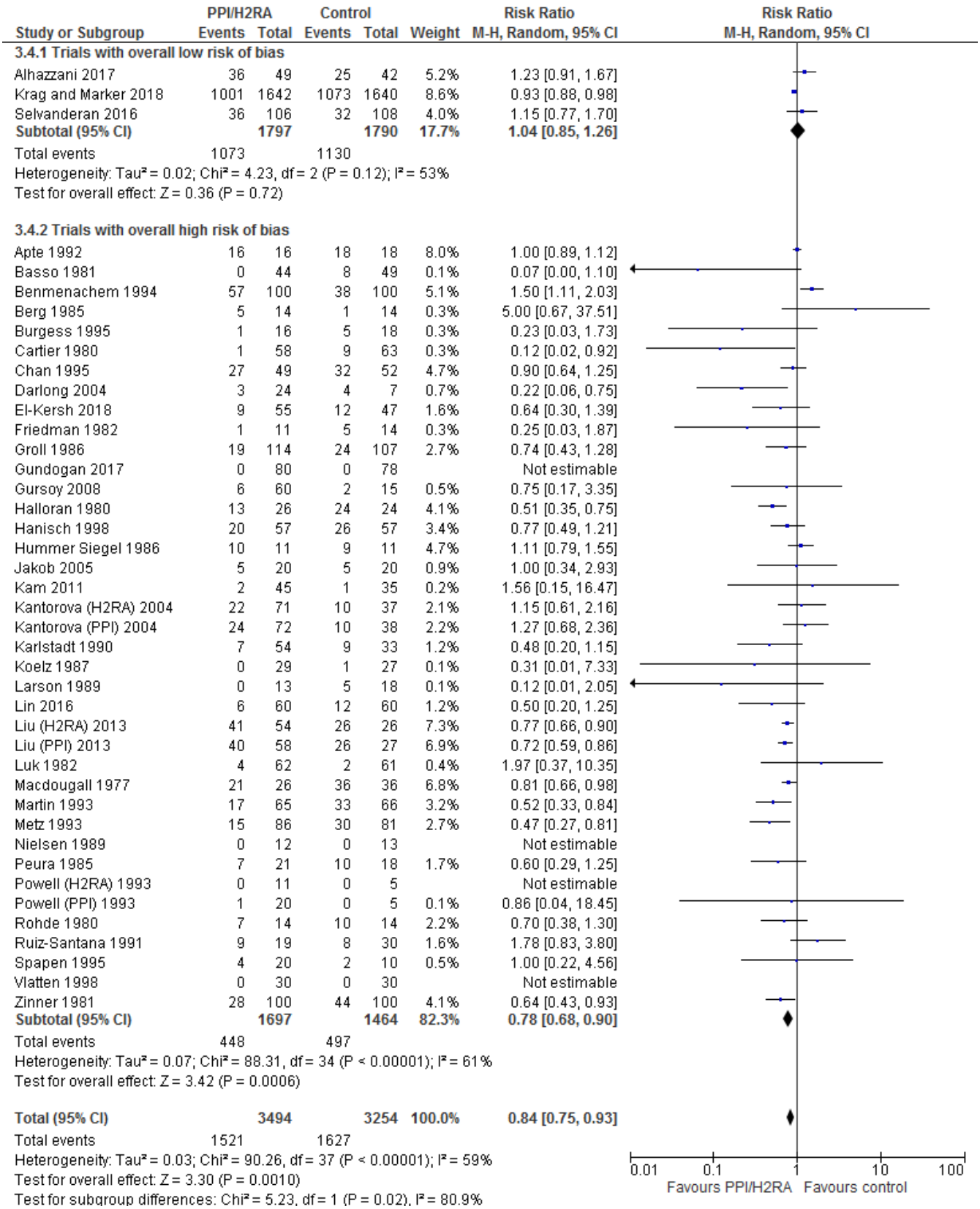
	All-cause mortality				Prop. of pts with CLIN IMP GI bleeding				Prop. of pts with ANY GI bleeding				Prop. of pts with hospital acquired pneumonia				Prop. of patients with *CL. difficile*				Prop. of pts with myocardial ischemia				SAE total events			
Trial id	E: Eve.	E: Ana.	C: Eve.	C: Ana.	E: Eve.	E: Ana.	C: Eve.	C: Ana.	E: Eve.	E: Ana.	C: Eve.	C: Ana.	E: Eve.	E: Ana.	C: Eve.	C: Ana.	E: Eve.	E: Ana.	C: Eve.	C: Ana.	E: Eve.	E: Ana.	C: Eve.	C: Ana.	E: Eve.	E: Ana.	C: Eve.	C: Ana.
Alhazzani 2017 [4]	17	49	13	42	3	49	2	42	4	49	3	42	10	49	6	42	2	49	1	42					36	49	25	42
Apte 1992 [5]	11	16	7	18					10	16	10	18	13	16	9	18									34	16	26	18
Basso 1981 [5]									0	44	8	49													0	44	8	49
Ben-menachem 1994 [6]	28	100	19	100	16	100	13	100					13	100	6	100									57	100	38	100
Berg 1985 [7]									5	14	1	14													5	14	1	14
Burgess 1995 [8]	1	16	0	18					0	16	5	18													1	16	5	18
Cartier 1980 [9]									1	58	9	63													1	58	9	63
Chan 1995 [10]					9	49	21	52					18	49	11	52									27	49	32	52
Darlong 2003 [11]									3	24	4	7													3	24	4	7
Domingues 1985 [40]																										15		15
El-Kersh 2018 [12]	7	55	8	47	1	55	1	47									1	55	3	47					9	55	12	47
Friedman 1982 [13]									1	11	5	14													1	11	5	14
Groll 1986 [14]	13	114	13	107					6	114	11	107													19	114	24	107
Gundogan 2017 [15]									0	80	0	78													0	80	0	78
Gursoy 2008 [16]	6	60	2	15																					6	60	2	15
Halloran 1980 [2]	8	26	10	24					5	26	18	24													13	26	28	24

Hanisch 1998 [17]	7	57	12	57	3	57	2	57					10	57	12	57									20	57	26	57
Hummer Sigiel 1986 [18]									0	11	0	11	10	11	9	11									10	11	9	11
Jakob 2005 [19]	5	20	5	20					0	20	0	20													5	20	5	20
Kam 2011 [20]									2	45	1	35													2	45	1	35
Kantotova H2RA 2004 [21]	11	71	6	37	2	71	1	37	2	71	1	37	7	71	2	37									22	71	10	37
Kantorova PPI 2004 [21]	14	72	7	38	1	72	0	38	1	72	0	38	8	72	3	38									24	72	10	38
Karlstadt 1990 [22]	5	54	2	33	1	54	7	33					1	54	0	33									7	54	9	33
Koelz 1987 [23]					0	29	1	27																	0	29	1	27
Krag 2018 [24]	510	1642	499	1640	41	1644	69	1647	88	1644	148	1647	266	1644	266	1647	19	1644	25	1647	77	1644	66	1647	1001	1642	1073	1640
Larson 1989 [25]									0	13	5	18													0	13	5	18
Lin 2016 [26]	2	60	0	60	0	60	1	60	0	60	5	60	4	60	6	60									6	60	12	60
Liu H2RA 2013 [27]	14	58	10	26					15	54	12	26	12	54	4	26									41	54	26	26
Liu PPI 2013 [27]	17	58	10	27					9	58	12	27	14	58	4	27									40	58	26	27
Luk 1982 [28]									4	62	2	61													4	62	2	61
MacDougall 1977 [3]	20	26	31	36					1	26	19	36													21	26	50	36
Martin 1993 [29]	8	65	7	66					9	65	22	66	0	65	4	66									17	65	33	66
Metz 1993 [30]									3	86	15	81	12	84	15	79									15	86	30	81
Nielsen 1989 [31]	0	12	0	13																					0	12	0	13
Peura 1985 [32]	7	21	7	18	0	21	3	18																	7	21	10	18
Powell H2RA 1993 [33]	0	11	0	5					0	11	0	5													0	11	0	5
Powell PPI 1993 [33]	1	20	0	5					0	20	0	5													1	20	0	5
Rigaud 1988 [41]																										6		6
Rohde 1980 [34]	7	14	6	14					0	14	4	14													7	14	10	14

Ruiz-Santana 1991 [35]	7	19	7	30					2	19	1	30													9	19	8	30
Selvanderan 2016 [36]	30	106	25	108	0	106	0	108	3	106	6	108	2	106	1	108	1	106	0	108					36	106	32	108
Spapen 1995 [37]	4	20	2	10																					4	20	2	10
Vlatten 1998 [38]									0	30	0	30													0	30	0	30
Zinner 1981 [39]	9	100	17	100	5	100	7	100	14	100	20	100													28	100	44	100

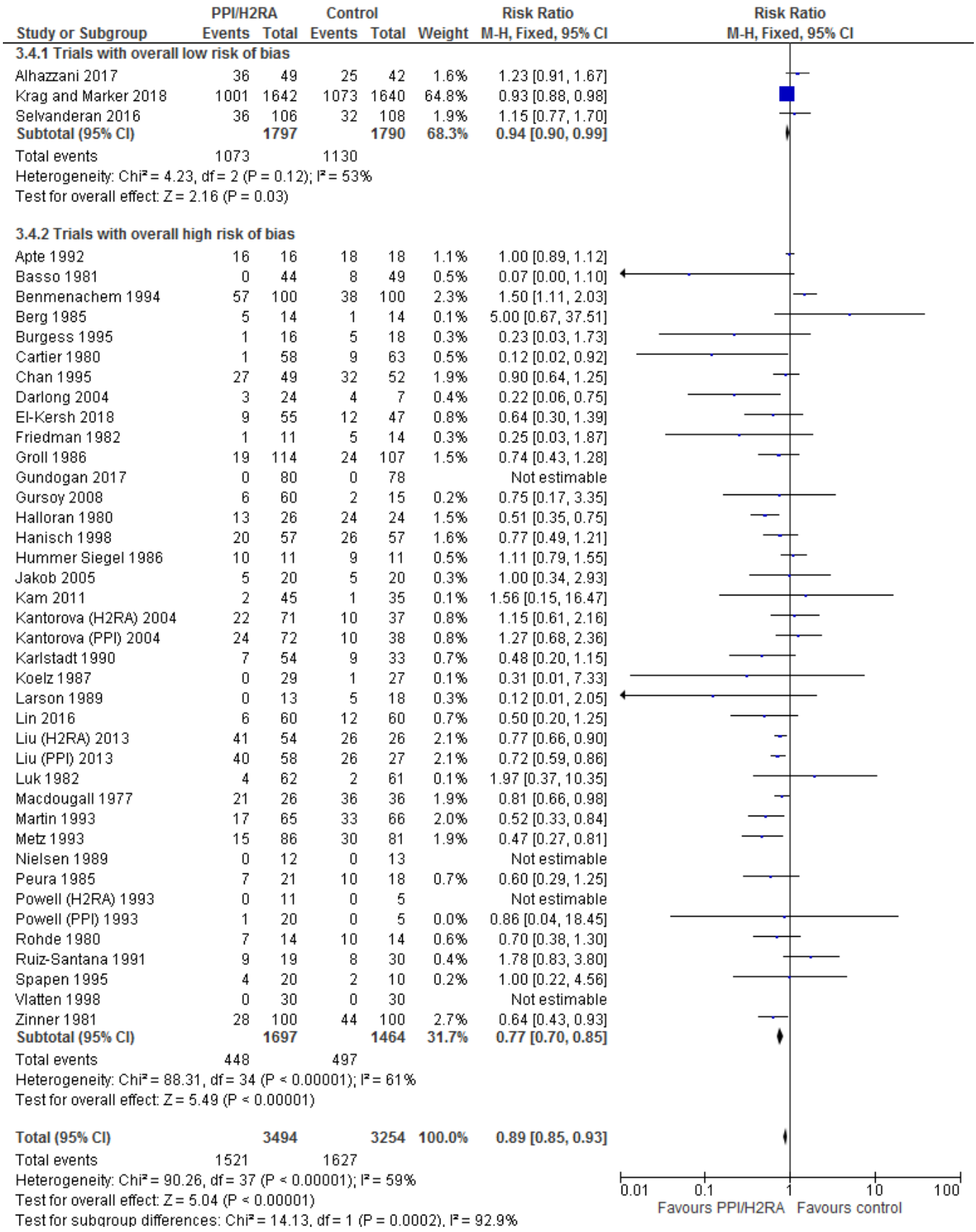
ana: analysed; C: control; CLIN IMP GI bleeding: clinically important gastrointestinal bleeding; CL. difficile: clostridium difficile; eve: events; E: experimental; No.: number; Prop.: proportion; pts: participants; SAE: serious adverse events

Fig. S25. Forest plot (random effects model) on serious adverse event (cumulated)



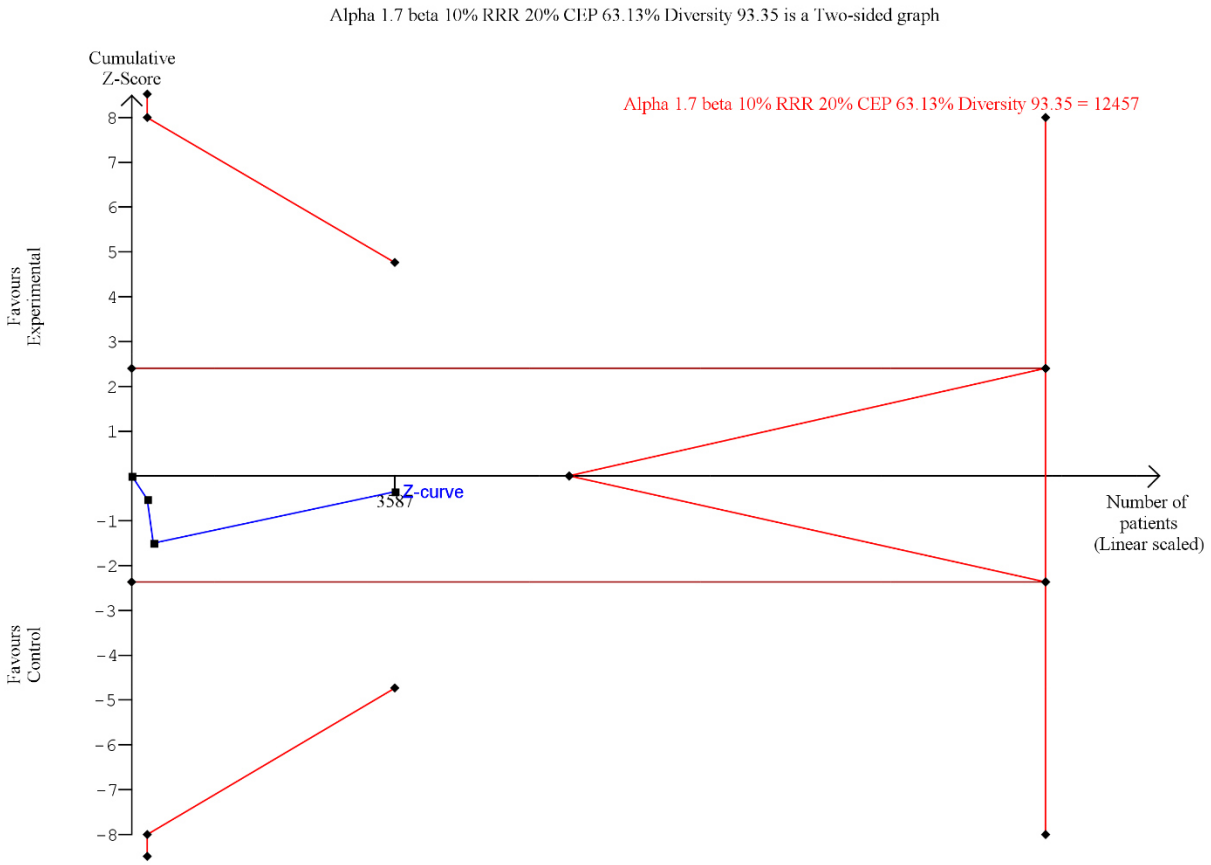
Random effects model. Meta-analysis showed that stress ulcer prophylaxis with PPI/H2RA versus placebo/no prophylaxis does not reduce serious adverse events.

Fig. S26. Forest plot (fixed effect model) on serious adverse events (cumulated)



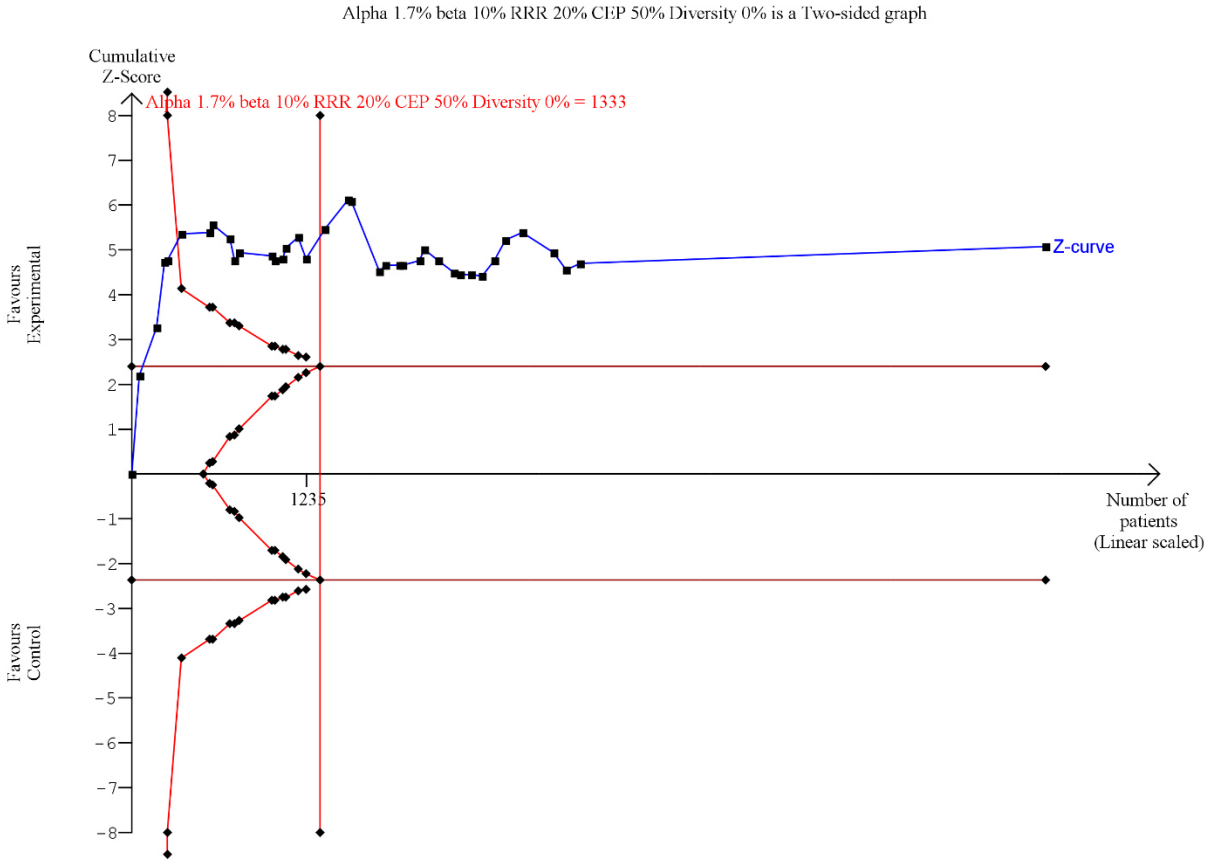
Fixed effects model. Meta-analysis showed that stress ulcer prophylaxis with PPI/H2RA versus placebo/no prophylaxis reduce serious adverse events.

Fig. S27. TSA of trials with overall low risk of bias for a 20% RRR/RRI on serious adverse events (cumulated)



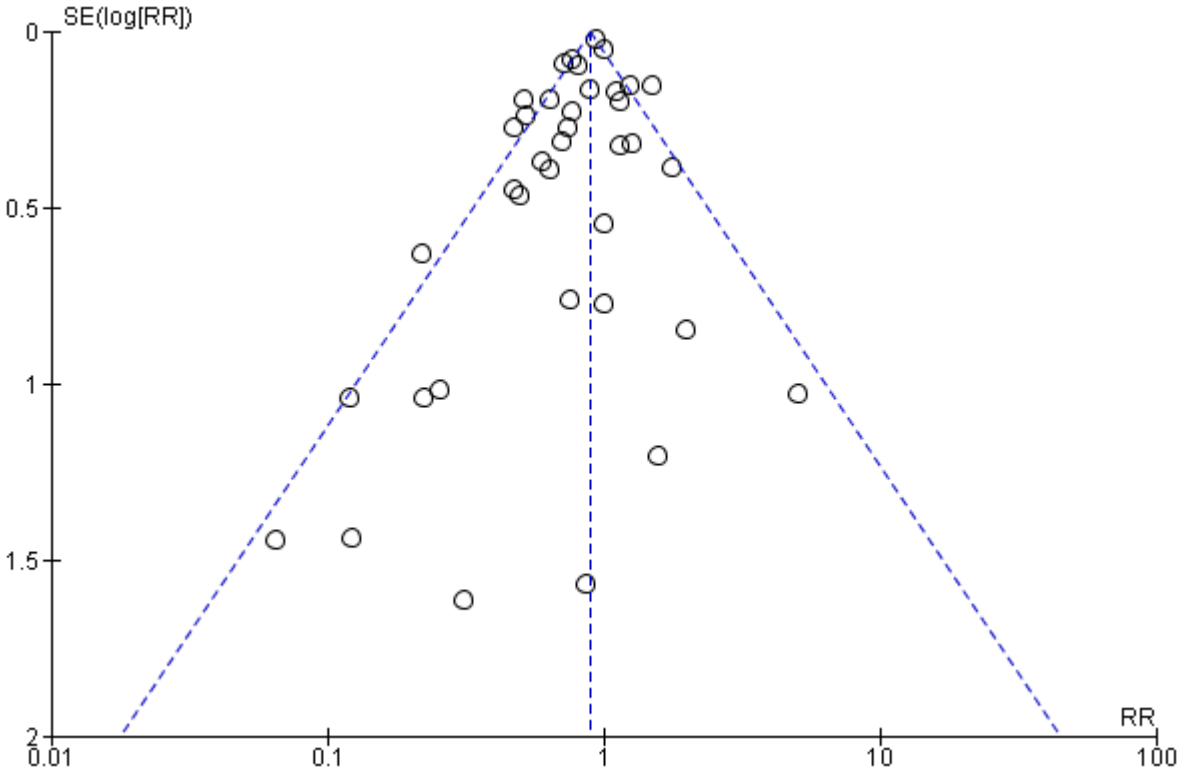
Control event proportion of 63.13%, diversity (D2) of 0%, alpha of 3.3%, power of 90% and RRR of 20%. We used an adjusted maximal type 1 error risk (alpha) of 0.033 due to two co-primary outcomes. The RR was 1.04 with a naive 95% CI of 0.85, 1.26 in a random effects model and the TSA-adjusted CI. 0.64, 1.68. As the cumulative Z-curve does not reach the trial sequential monitoring boundary, futility area or required information size, we do not have enough information to detect or reject a 20% RRR/RRI.

Fig. S28. TSA of all trials for a 20% RRR/RRI on serious adverse events (cumulated)



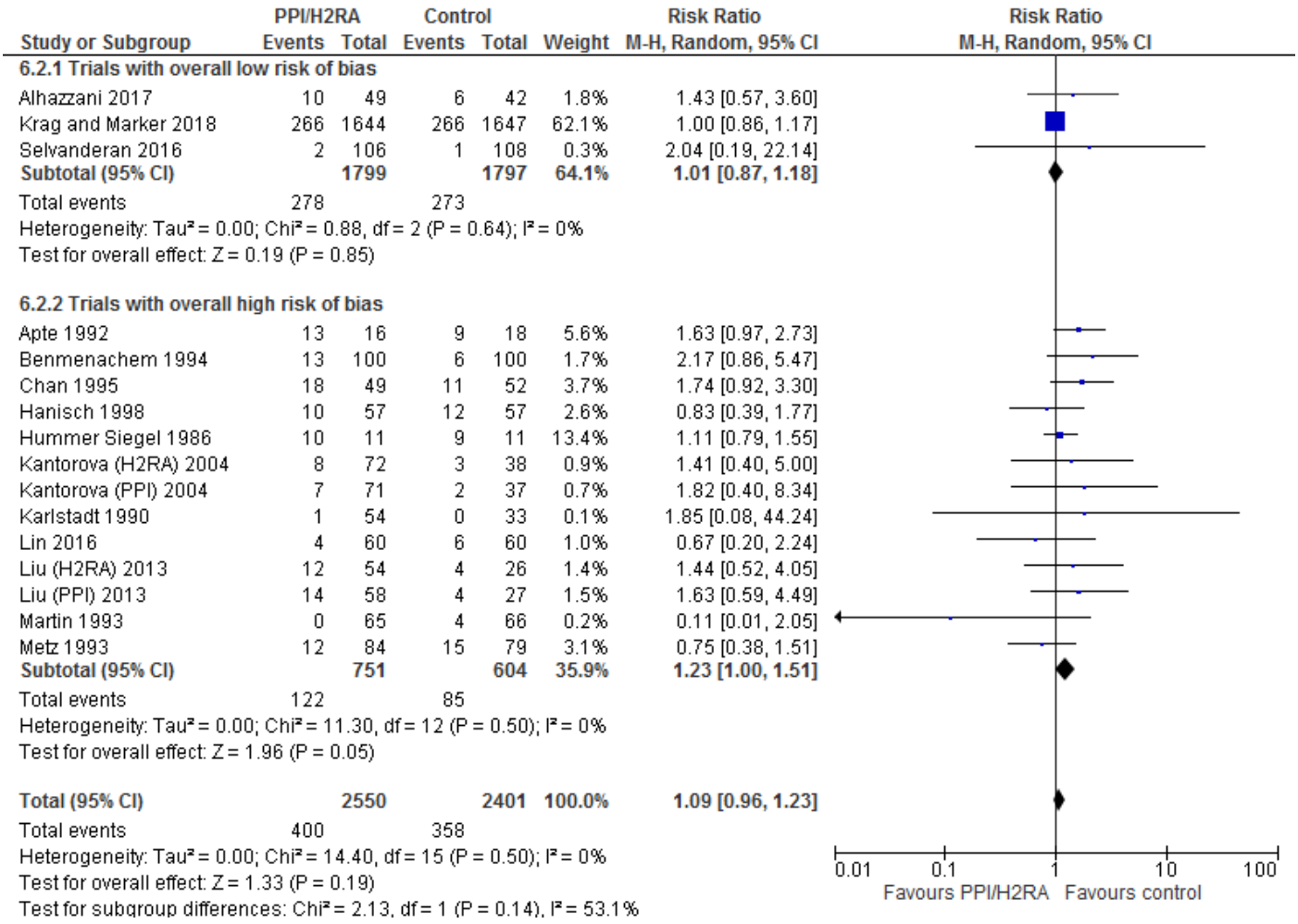
Control event proportion of 50%, diversity (D2) of 0%, alpha of 1.7%, power of 90% and RRR of 20%. We used an adjusted maximal type 1 error risk (alpha) of 0.017 due to five co-secondary outcomes. The RR was 0.89 with a naive 95% CI of 0.85, 0.93 in a fixed effects model and the TSA-adjusted CI. 0.85, 0.93. As the cumulative Z-curve reaches trial sequential monitoring boundary for benefit and required information size, we may accept a 20% RRR.

Fig. S29. Funnel plot on serious adverse events (cumulated)



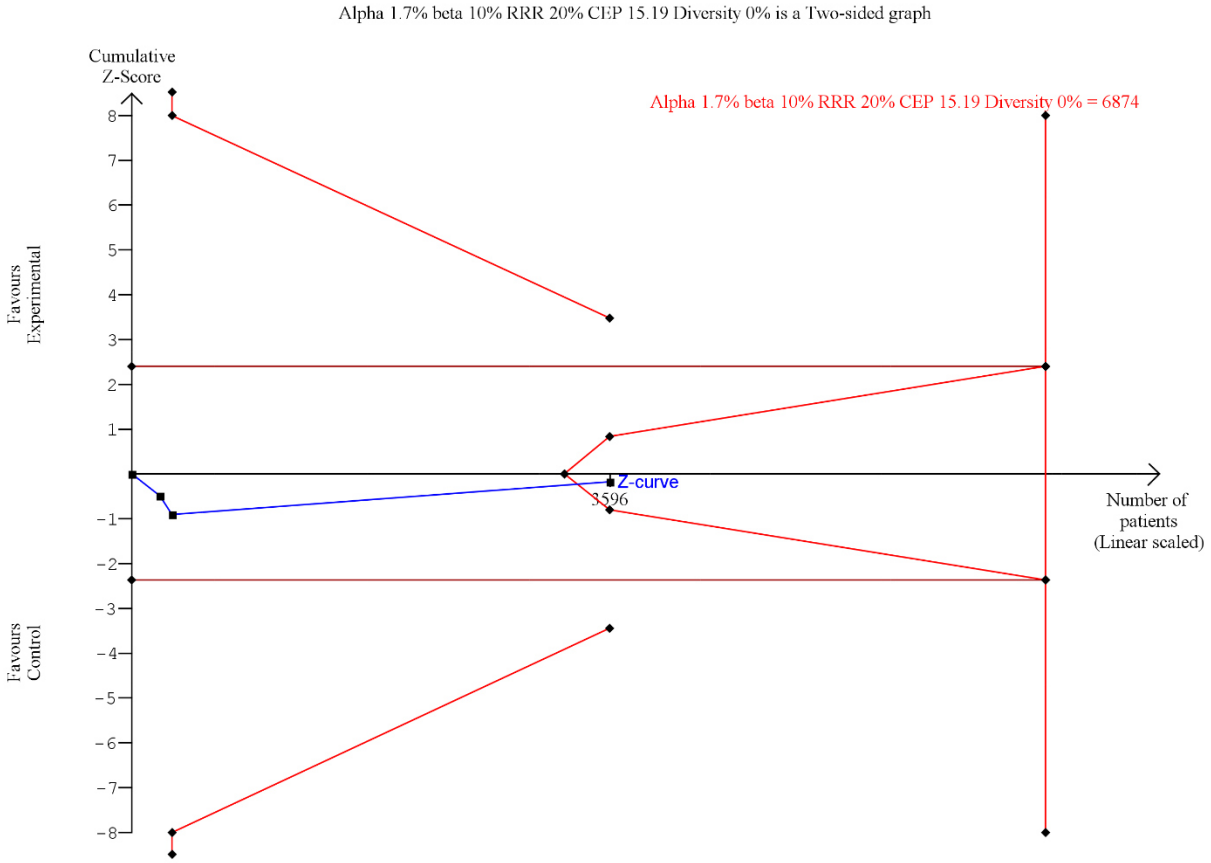
PNEUMONIA

Fig. S30. Forest plot (random effects model) on pneumonia



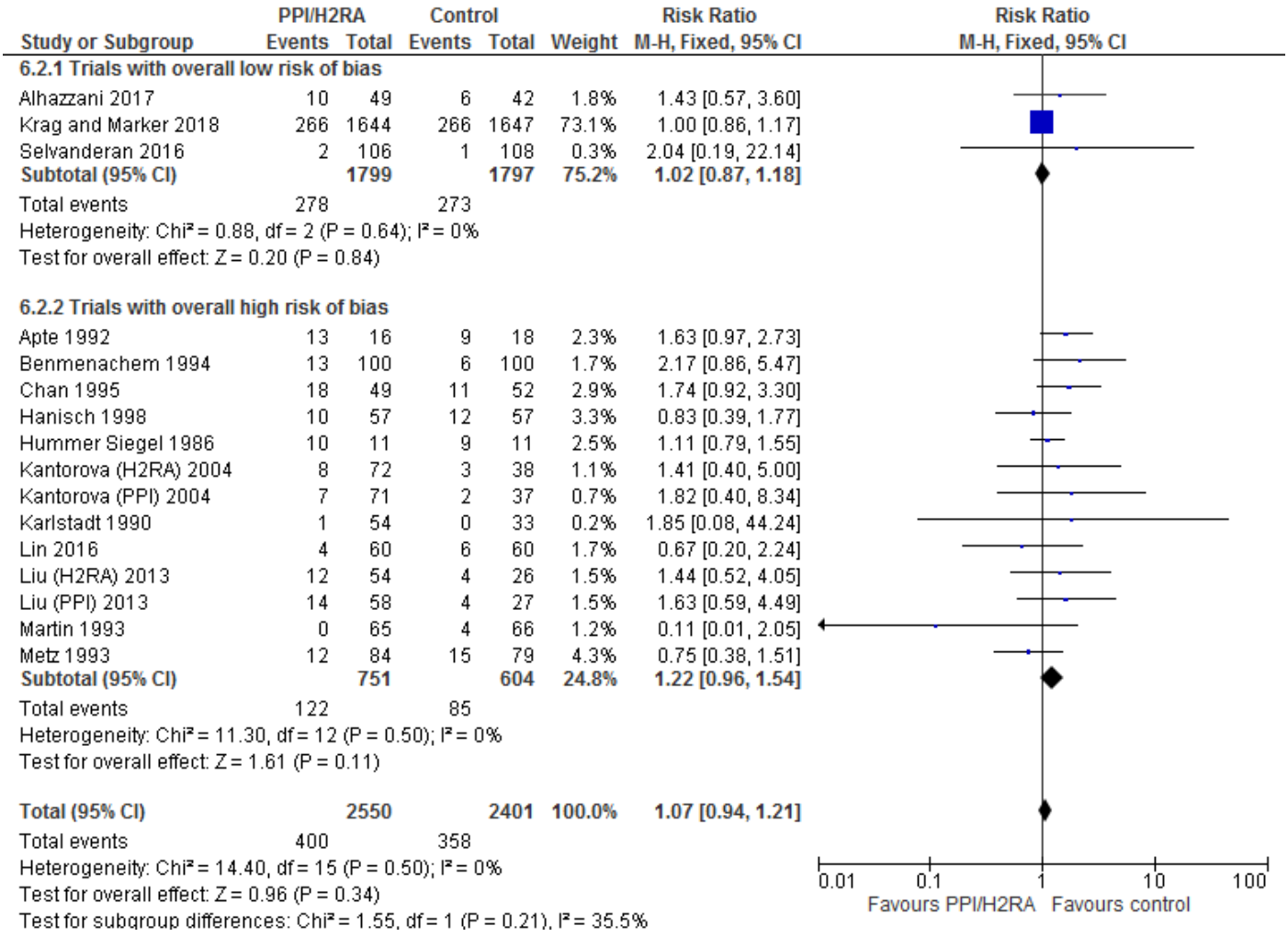
Random effects model. Meta-analysis showed that stress ulcer prophylaxis with PPI/H2RA versus placebo/no prophylaxis does not reduce/increase pneumonia.

Fig. S31. TSA of trials with overall low risk of bias for a 20% RRR/RRI on pneumonia



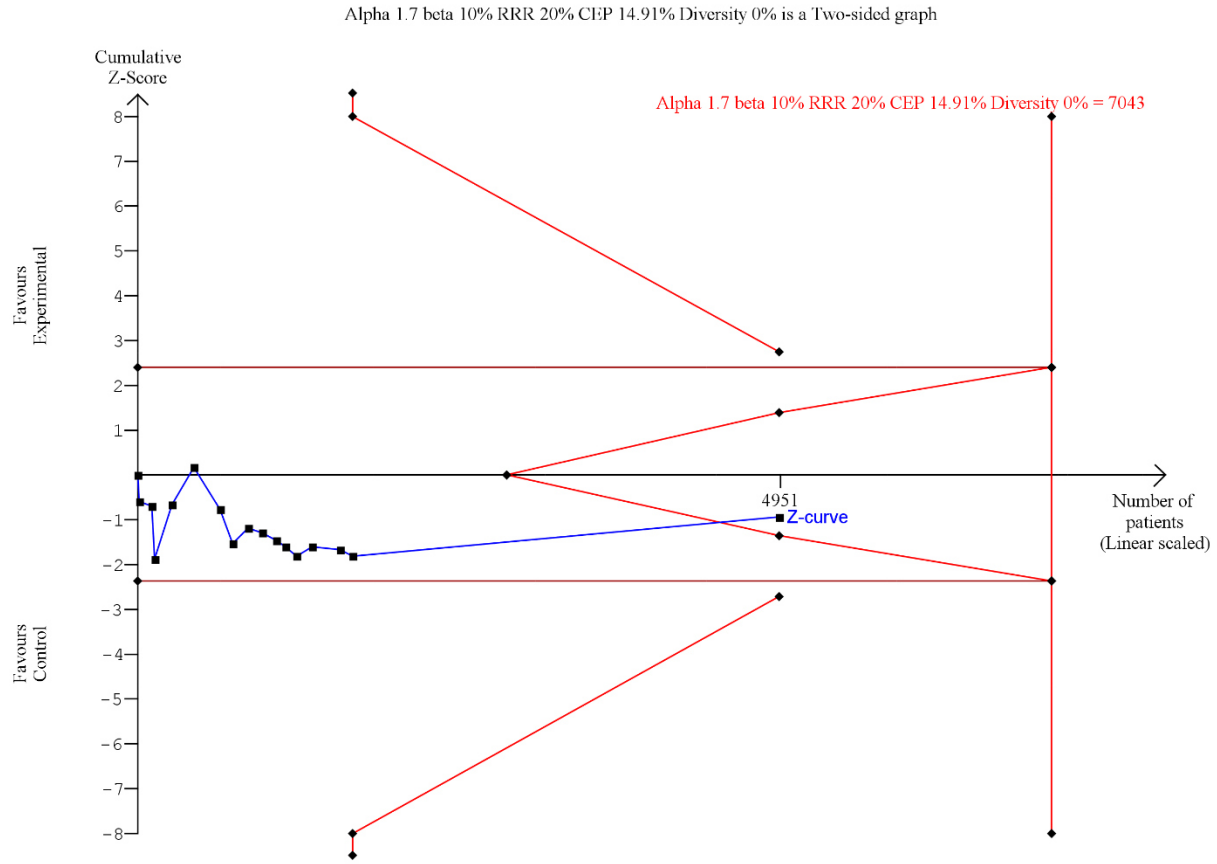
Control event proportion of 15.9%, diversity (D2) of 0%, alpha of 1.7%, power of 90% and RRR/RRI of 20%. We used an adjusted maximal type 1 error risk (alpha) of 0.017 due to two co-secondary outcomes. The RR was 1.01 with a naive 95% CI of 0.87, 1.18 in a random effects model and the TSA-adjusted CI 0.77, 1.33. As the cumulative Z-curve reaches futility area for non-inferiority we may exclude a 20% RRR/RRI.

Fig. S32. Forest plot (fixed effect model) on pneumonia



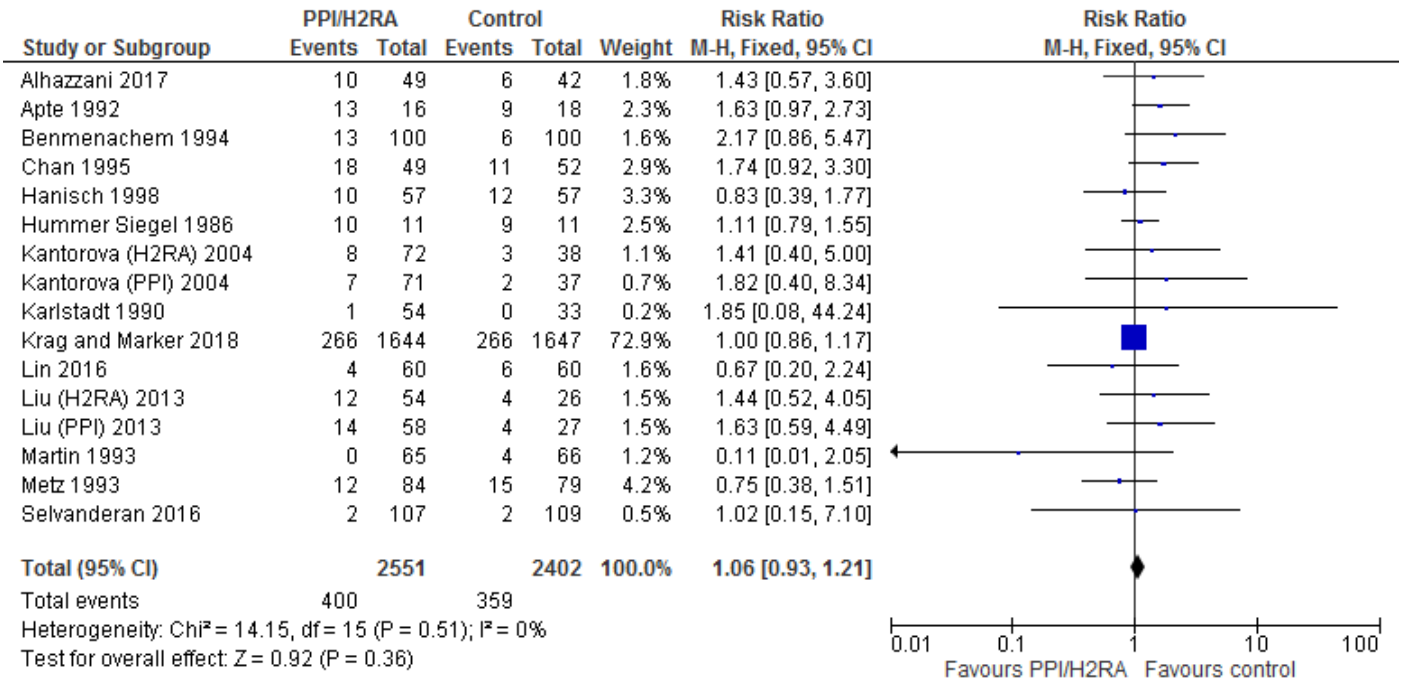
Fixed effects model. Meta-analysis showed that stress ulcer prophylaxis with PPI/H2RA versus placebo/no prophylaxis does not reduce/increase pneumonia.

Fig. S33. TSA of all trials for a 20% RRR/RRI on pneumonia



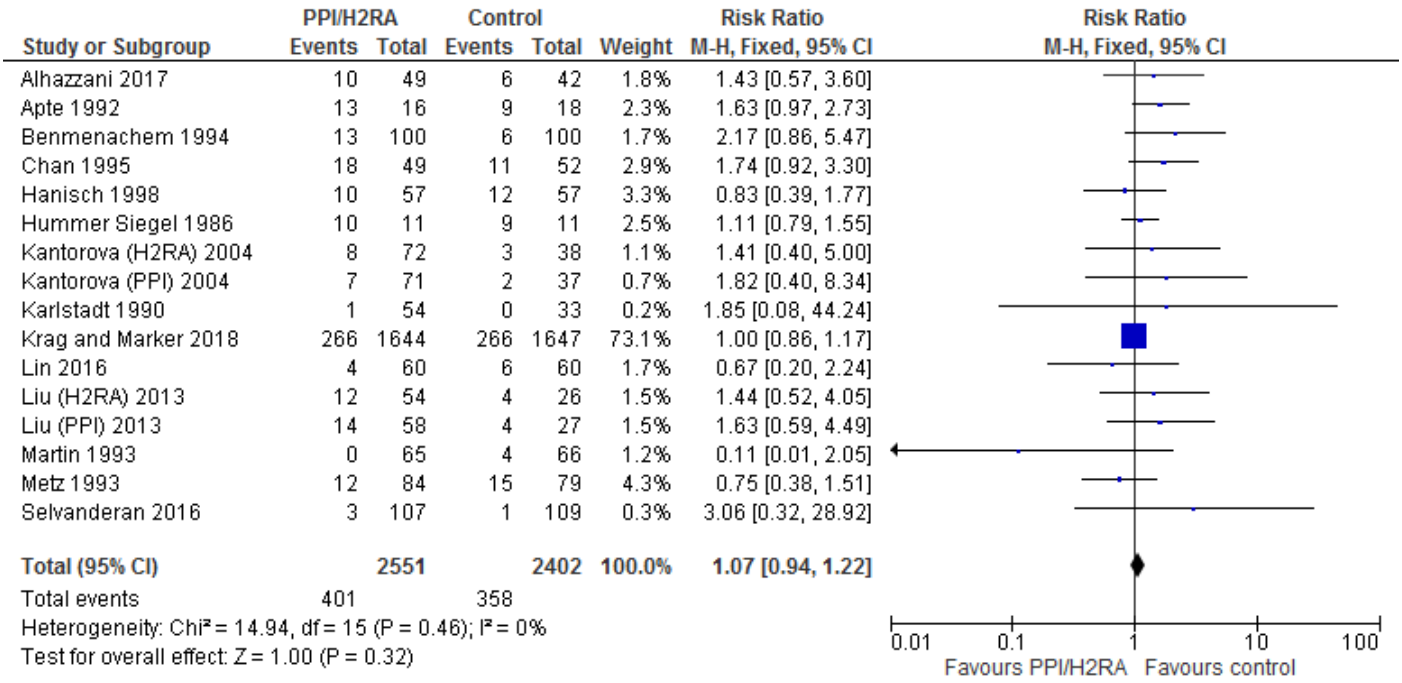
Control event proportion of 14.91%, diversity (D2) of 0%, alpha of 1.7%, power of 90% and RRR/RRI of 20%. We used an adjusted maximal type 1 error risk (alpha) of 0.017 due to five co-secondary outcomes. The RR was 1.07 with a naive 95% CI of 0.94, 1.21 in a fixed effects model and the TSA-adjusted CI 0.89, 1.27. As the cumulative Z-curve reaches futility area for non-inferiority we may exclude a 20% RRR/RRI.

Fig. S34. Forest plot of best worst-case scenario sensitivity analysis for missing data on pneumonia



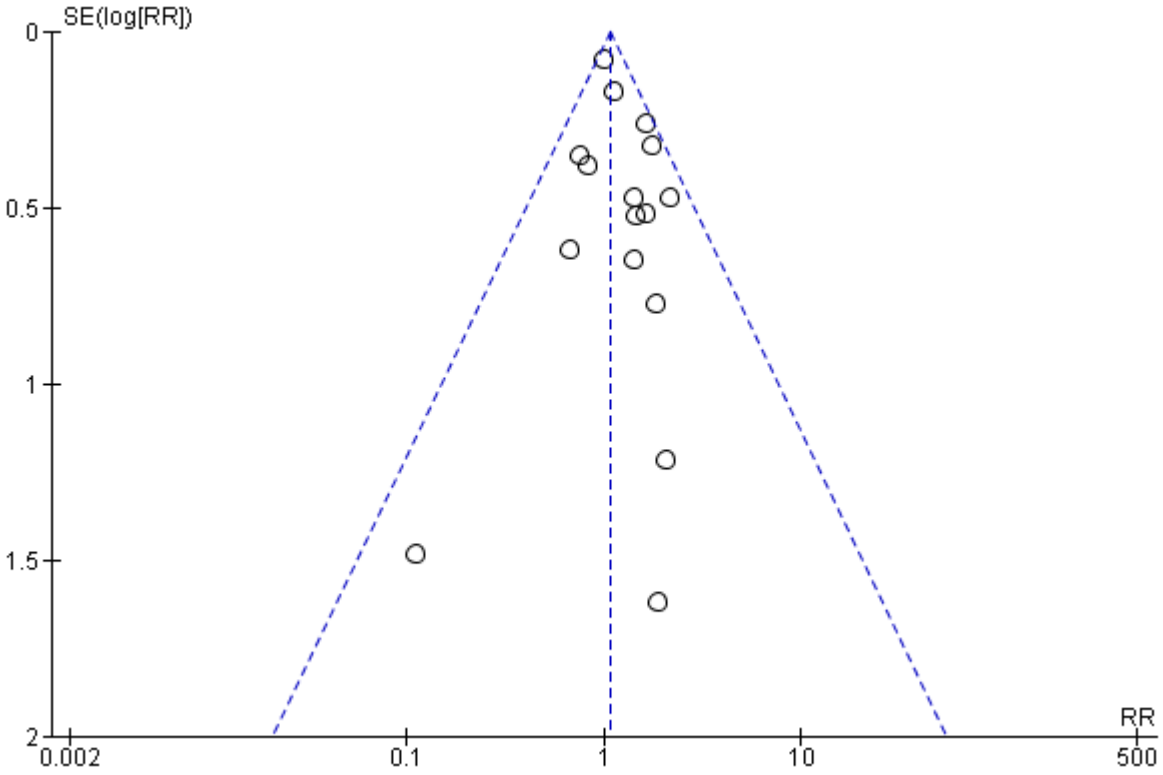
In this analysis it is assumed that all participants lost to follow-up in the PPI/H2RA group did not have pneumonia and all those with missing outcomes in the control group have had pneumonia. The primary meta-analysis result (RR 1.07, 95% CI 0.94, 1.21) and the result from this sensitivity analysis (RR 1.06, 95% CI 0.93, 1.21) show similar P values and CIs.

Fig. S35. Forest plot of worst best-case scenario sensitivity analysis for missing data on pneumonia



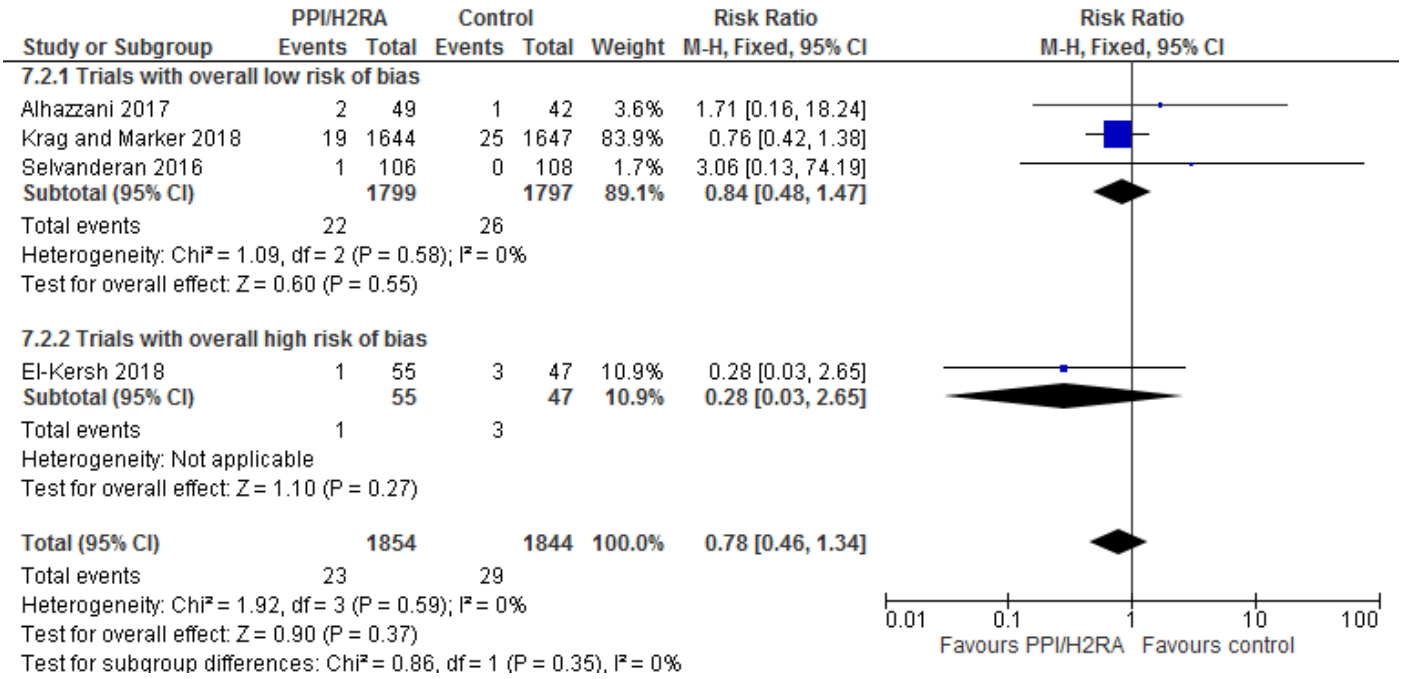
In this analysis it is assumed that all participants lost to follow-up in the PPI/H2RA group had pneumonia and all those with missing outcomes in the control group did not have pneumonia. The primary meta-analysis result (RR 1.07, 95% CI 0.94, 1.21) and the result from this sensitivity analysis (RR 1.07, 95% CI 0.94, 1.22) show similar P values and CIs.

Fig. S36. Funnel plot on pneumonia



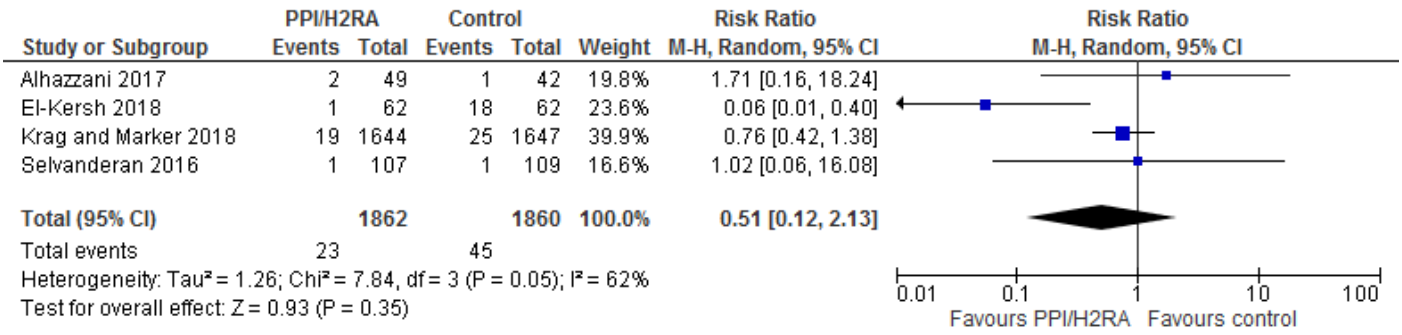
CLOSTRIDIUM DIFFICILE ENTERITIS

Fig. S37. Forest plot on *Cl. difficile* enteritis



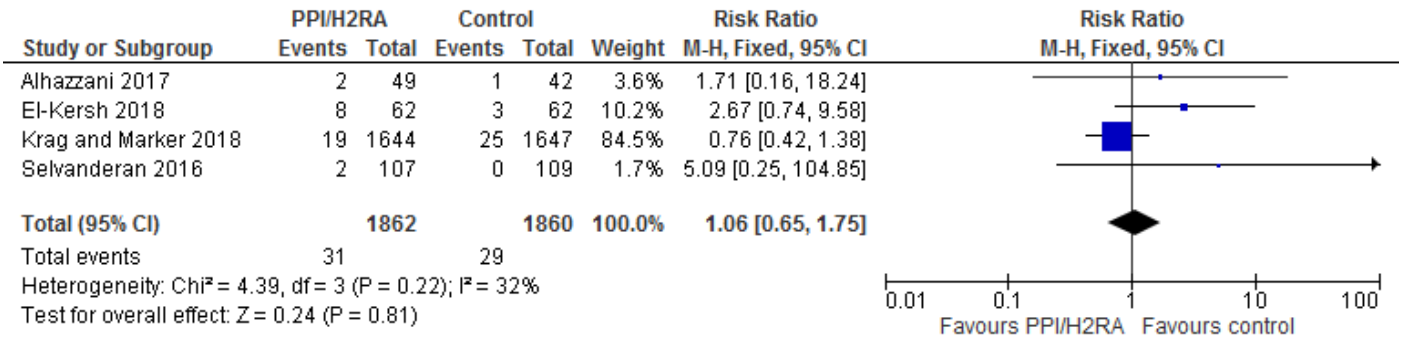
Meta-analysis showed that stress ulcer prophylaxis with PPI/H2RA versus placebo/no prophylaxis does not reduce/increase *Cl. difficile* enteritis.

Fig. S38. Forest plot of best worst-case scenario sensitivity analysis for missing data on *Cl. difficile* enteritis



In this analysis it is assumed that all participants lost to follow-up in the PPI/H2RA group did not have *Cl. difficile* enteritis and all those with missing outcomes in the control group have had *Cl. difficile* enteritis. The primary meta-analysis result (RR 0.78, 95% CI 0.46, 1.34) and the result from this sensitivity analysis (RR 0.51, 95% CI 0.12, 2.13) show similar P values and CI.

Fig. S39. Forest plot of worst best-case scenario sensitivity analysis for missing data on *Cl. difficile* enteritis



In this analysis it is assumed that all participants lost to follow-up in the PPI/H2RA group had *Cl. difficile* enteritis and all those with missing outcomes in the control group did not have Cl. difficile enteritis. The primary meta-analysis result (RR 0.78, 95% CI 0.46, 1.34) and the result from this sensitivity analysis (RR 1.06, 95% CI 0.65, 1.75) show similar P values and CIs.

Reference List

1. Apte NM, Karnad DR, Medhekar TP, Tilve GH, Morye S, Bhave GG (1992) Gastric Colonization and Pneumonia in Intubated Critically Ill Patients Receiving Stress Ulcer Prophylaxis: A Randomized, Controlled Trial. Crit Care Med; 20(5): 590-3.
2. Halloran LG, Zfass AM, Gayle WE, Wheeler CB, Miller JD (1980) Prevention of Acute Gastrointestinal Complications after Severe Head Injury: A Controlled Trial of Cimetidine Prophylaxis. Am J Surg; 139(1): 44-8.
3. Macdougall BR, Bailey RJ, Williams R (1977) H2-Receptor Antagonists and Antacids in the Prevention of Acute Gastrointestinal Haemorrhage in Fulminant Hepatic Failure. Two Controlled Trials. Lancet; 1(8012): 617-9.
4. Alhazzani W, Guyatt G, Alshahrani M, Deane AM, Marshall JC, Hall R, Muscedere J, English SW, Lauzier F, Thabane L, Arabi YM, Karachi T, Rochwerg B, Finfer S, Daneman N, Alshamsi F, Zytaruk N, Heel-Ansdell D, Cook D (2017) Withholding Pantoprazole for Stress Ulcer Prophylaxis in Critically Ill Patients: A Pilot Randomized Clinical Trial and Meta-Analysis. Crit Care Med; 45(7): 1121-1129.
5. Basso N, Bagarani M, Materia A, Fiorani S, Lunardi P, Speranza V (1981) Cimetidine and Antacid Prophylaxis of Acute Upper Gastrointestinal Bleeding in High Risk Patients. Controlled, Randomized Trial. Am J Surg; 141(3): 339-41.
6. Ben-Menachem T, Fogel R, Patel RV, Touchette M, Zarowitz BJ, Hadzijahic N, Divine G, Verter J, Bresalier RS (1994) Prophylaxis for Stress-Related Gastric Hemorrhage in the Medical Intensive Care Unit. A Randomized, Controlled, Single-Blind Study. Ann Intern Med; 121(8): 568-75.
7. Van Den Berg B, Van Blankenstein M (1985) Prevention of Stress-Induced Upper Gastrointestinal Bleeding by Cimetidine in Patients on Assisted Ventilation. Digestion; 31(1): 1-8.
8. Burgess P, Larson GM, Davidson P, Brown J, Metz CA (1995) Effect of Ranitidine on Intragastric Ph and Stress-Related Upper Gastrointestinal Bleeding in Patients with Severe Head Injury. Dig Dis Sci; 40(3): 645-50.
9. Cartier F, Gauthier-Lafaye P, Lareng L, Mottin J, Cara M, Passelecq J, (1980) Cimetideine in Patients at Risk of Stress Ulcers: A Multi-Centre Controlled Trial. Intensive Care Medicine; 6: 54.
10. Chan KH, Lai EC, Tuen H, Ngan JH, Mok F, Fan YW, Fung CF, Yu WC (1995) Prospective Double-Blind Placebo-Controlled Randomized Trial on the Use of Ranitidine in Preventing Postoperative Gastroduodenal Complications in High-Risk Neurosurgical Patients. J Neurosurg; 82(3): 413-7.
11. Darlong V, Jayalakhsmi TS, Kaul HL, Tandon R (2003) Stress Ulcer Prophylaxis in Patients on Ventilator. Trop Gastroenterol; 24(3): 124-8.
12. El-Kersh K, Jalil B, Mcclave SA, Cavallazzi R, Guardiola J, Guilkey K, Persaud AK, Furmanek SP, Guinn BE, Wiemken TL, Alhariri BC, Kellie SP, Saad M (2018) Enteral Nutrition as Stress Ulcer Prophylaxis in Critically Ill Patients: A Randomized Controlled Exploratory Study. J Crit Care; 43: 108-113.
13. Friedman CJ, Oblinger MJ, Suratt PM, Bowers J, Goldberg SK, Sperling MH, Blitzer AH (1982) Prophylaxis of Upper Gastrointestinal Hemorrhage in Patients Requiring Mechanical Ventilation. Crit Care Med; 10(5): 316-9.
14. Groll A, Simon JB, Wigle RD, Taguchi K, Todd RJ, Depew WT (1986) Cimetidine Prophylaxis for Gastrointestinal Bleeding in an Intensive Care Unit. Gut; 27(2): 135-40.
15. Gundogan K, Karakoc E, Teke T, Zerman A, Coruh A, Sungur M (2017) Effects of Enteral Nutrition on Stress Ulcer Hemorrhage in Critically Ill Patients: Multicenter Randomized Controlled Trial. Intensive Care Medicine Experimental; 5(2): 44.
16. Gursoy O, Memis D, Sut N (2008) Effect of Proton Pump Inhibitors on Gastric Juice Volume, Gastric Ph and Gastric Intramucosal Ph in Critically Ill Patients : A Randomized, Double-Blind, Placebo-Controlled Study. Clin Drug Investig; 28(12): 777-82.
17. Hanisch EW, Encke A, Naujoks F, Windolf J (1998) A Randomized, Double-Blind Trial for Stress Ulcer Prophylaxis Shows No Evidence of Increased Pneumonia. Am J Surg; 176(5): 453-7.
18. Hummer-Sigiel M, Jacquier A, Girard A, Garric J, M.C. L, J.Y. M (1986) Ranitidine Pour La Prophylaxie De L'ulcére De Stress Chexz Les Traumatisés Crâniens Graves Annales médicales de Nancy et de l'Est; 25: 101-103.
19. Jakob SM, Parviainen I, Ruokonen E, Uusaro A, Takala J (2005) Lack of Effect of Ranitidine on Gastric Luminal Ph and Mucosal Pco2 During the First Day in the Icu. Acta Anaesthesiol Scand; 49(3): 390-6.
20. Kam J, Modi C, Doraiswamy V, Abdul-Jawad S, Dixit D, Spira T, Adelman M (2011) Role of Gastrointestinal Ulcer Prophylaxis in Critically Ill Patients. The American Journal of Gastroenterology; 106(suppl. 2): s420.
21. Kantorova I, Svoboda P, Scheer P, Doubek J, Rehorkova D, Bosakova H, Ochmann J (2004) Stress Ulcer Prophylaxis in Critically Ill Patients: A Randomized Controlled Trial. Hepatogastroenterology; 51(57): 757-61.
22. Karlstadt RG, Iberti TJ, Silverstein J, Lindenberg L, Bright-Asare P, Rockhold F, Young MD (1990) Comparison of Cimetidine and Placebo for the Prophylaxis of Upper Gastrointestinal Bleeding Due to Stress-Related Gastric Mucosal Damage in the Intensive Care Unit. J Intensive Care Med; 5: 26-32.

23. Koelz HR, Aeberhard P, Hassler H, Kunz H, Wagner HE, Roth F, Halter F (1987) Prophylactic Treatment of Acute Gastroduodenal Stress Ulceration. Low-Dose Antacid Treatment without and with Additional Ranitidine. Scand J Gastroenterol; 22(9): 1147-52.
24. Krag M, Marker S, Perner A, Wetterslev J, Wise MP, Schefold JC, Keus F (2018) Pantoprazole in Patients at Risk for Gastrointestinal Bleeding in the Icu. New England Journal of Medicine.
25. Larson GM, Davidson P, Brown J, Wilson T, Bishop A (1989) Comparison of Ranitidine Versus Placebo on 24-Hour Gastric Ph and Upper Gastrointestinal (Ugi) Bleeding in Head Injury Patients. Abstr Am J Gastroenterol; 84: 1165.
26. Lin CC, Hsu YL, Chung CS, Lee TH (2016) Stress Ulcer Prophylaxis in Patients Being Weaned from the Ventilator in a Respiratory Care Center: A Randomized Control Trial. J Formos Med Assoc; 115(1): 19-24.
27. Liu BL, Li B, Zhang X, Fei Z, Hu SJ, Lin W, Gao DK, Zhang L (2013) A Randomized Controlled Study Comparing Omeprazole and Cimetidine for the Prophylaxis of Stress-Related Upper Gastrointestinal Bleeding in Patients with Intracerebral Hemorrhage. J Neurosurg; 118(1): 115-20.
28. Luk GD, Summer WR, Messersmith JF (1982) Cimetidine and Antacid in Prophylaxis of Acute Gastrointestinal Bleeding. A Randomized, Double-Blind, Controlled Study. Gastroenterology; 82: 1121.
29. Martin LF, Booth FV, Karlstadt RG, Silverstein JH, Jacobs DM, Hampsey J, Bowman SC, D'ambrosio CA, Rockhold FW (1993) Continuous Intravenous Cimetidine Decreases Stress-Related Upper Gastrointestinal Hemorrhage without Promoting Pneumonia. Crit Care Med; 21(1): 19-30.
30. Metz CA, Livingston DH, Smith JS, Larson GM, Wilson TH (1993) Impact of Multiple Risk Factors and Ranitidine Prophylaxis on the Development of Stress-Related Upper Gastrointestinal Bleeding: A Prospective, Multicenter, Double-Blind, Randomized Trial. The Ranitidine Head Injury Study Group. Crit Care Med; 21(12): 1844-9.
31. Nielsen HJ, Witt K, Moesgaard F, Kehlet H (1989) Ranitidine for Improvement of Delayed Hypersensitivity Response in Patients with Sepsis. Acta Chir Scand; 155(9): 445-9.
32. Peura DA, Johnson LF (1985) Cimetidine for Prevention and Treatment of Gastroduodenal Mucosal Lesions in Patients in an Intensive Care Unit. Ann Intern Med; 103(2): 173-7.
33. Powell H, Morgan M, Li SK, Baron JH (1993) Inhibition of Gastric Acid Secretion in the Intensive Care Unit after Coronary Artery Bypass Graft. Theor Surg; 8: 125-130.
34. Rohde H, Lorenz W, Fischer M (1980) Eine Randomisierte Klinische Studie Zur Stressulkusprophylaxe Mit Cimetidin Beim Schweren Polytrauma. Z Gastroenterol; 18(6): 328-9.
35. Ruiz-Santana S, Ortiz E, Gonzalez B, Bolanos J, Ruiz-Santana AJ, Manzano JL (1991) Stress-Induced Gastroduodenal Lesions and Total Parenteral Nutrition in Critically Ill Patients: Frequency, Complications, and the Value of Prophylactic Treatment. A Prospective, Randomized Study. Crit Care Med; 19(7): 887-91.
36. Selvanderan SP, Summers MJ, Finnis ME, Plummer MP, Ali Abdelhamid Y, Anderson MB, Chapman MJ, Rayner CK, Deane AM (2016) Pantoprazole or Placebo for Stress Ulcer Prophylaxis (Pop-up): Randomized Double-Blind Exploratory Study. Crit Care Med; 44(10): 1842-50.
37. Spapen H, Diltoer M, Nguyen DN, Ingels G, Ramet J, Huyghens L (1995) One Week Treatment with Cimetidine Does Not Attenuate the Cortisol Response to a Short Corticotropin Test in Stable Intensive Care Patients: A Prospective, Randomized, and Controlled Study. Acta Anaesthesiol Belg; 46(3-4): 133-40.
38. Vlatten A, Wiedeck H, Reinelt H, Stanescu A, Georgieff M (1998) Stressulkus-Prophylaxe Bei Hoch-Risiko-Intensivpatienten. Vergleich Von Omeprazol, Pirenzepin Und Plazebo. Wiener Klinische Wochenschrift; 110(suppl. 1): 38.
39. Zinner MJ, Zuidema GD, Smith P, Mignosa M (1981) The Prevention of Upper Gastrointestinal Tract Bleeding in Patients in an Intensive Care Unit. Surg Gynecol Obstet; 153(2): 214-20.
40. Domingues SHS, Stoeber GH, Modena AC (1985) Ranitidina Injetável Em Pacientes De Alto Risco. A folha medica; 91(3): 225-228.
41. Rigaud D, Accary JP, Chastre J, Mignon M, Laigneau JP, Reinberg A, Bonfils S (1988) Persistence of Circadian Rhythms in Gastric Acid, Gastrin, and Pancreatic Polypeptide Secretions Despite Loss of Cortisol and Body Temperature Rhythms in Man under Stress. Gastroenterol Clin Biol; 12(1): 12-8.

SUPPLEMENTARY MATERIAL

This supplementary material has been provided by the authors to give readers additional information about their work.

Supplement to: Barbateskovic M, Schjørring OL, Krauss SR, Meyhoff CS, Jakobsen JJ, Rasmussen BS, Perner A, Wetterslev J. Higher versus lower levels of oxygen supplementation in critically ill adults. A systematic review with meta-analysis and Trial Sequential Analysis.

Correspondence: Barbateskovic M, Copenhagen Trial Unit, Centre for Clinical Intervention Research, DK2100 Copenhagen, Denmark, Tel: +45 3545 7113, E-mail: marija.barbateskovic@ctu.dk

Supplementary material for

Higher versus lower levels of oxygen supplementation in critically ill adults. A systematic review with meta-analysis and Trial Sequential Analysis

PRISMA CHECKLIST

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	3-4
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	4
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	4
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	**5**
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	ESM
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	**5**
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	5
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	ESM
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	5
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	5

Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	6
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	5
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	6-7
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	8
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Table 1, ESM
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	8, ESM
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	Figure 1, 3, 4, ESM
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	8-11, ESM
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	8
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	8-11, ESM
DISCUSSION			
Summary of evidence	**24**	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	Table 1
Limitations	**25**	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	11-12
Conclusions	**26**	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	12-13
FUNDING			
Funding	**27**	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	27

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

SEARCH STRATEGIES

Cochrane Central Register of Controlled Trials (CENTRAL) – 15,248 hits

From inception to 17 October 2019

#1 MeSH descriptor: [Hyperoxia] explode all trees
#2 MeSH descriptor: [Anoxia] explode all trees
#3 MeSH descriptor: [Oxygen Inhalation Therapy] explode all trees
#4 MeSH descriptor: [Oxygen] explode all trees
#5 (inspir* or inhal* or fraction* or concentrat* or arterial* or saturation or level* or tension* or supply* or supplement* or supplie* or therap* or administr* or dosag* or dose* or dosing*) near/3 (oxygen):ti,ab,kw
#6 (hyperoxia or hyperoxemia or hyperoxaemia or hypoxia or hypoxemia or hypoxaemia or anoxia or anoxemia or anoxaemia or arterial oxygen or high oxygen or oxygenat* or blood gas or oxygen saturation or pao2 or sao2 or spo2 or fio2):ti,ab,kw
#7 (#1 or #2 or #3 or #4 or #5 or #6)
#8 MeSH descriptor: [Critical Illness] explode all trees
#9 MeSH descriptor: [Critical Care] explode all trees
#10 MeSH descriptor: [Intensive Care Units] explode all trees
#11 MeSH descriptor: [Emergency Medicine] explode all trees
#12 MeSH descriptor: [Emergency Service, Hospital] explode all trees
#13 (emergency department* or ED or emergency room* or ER or high dependency unit* or HDU or prehospital* or critically ill or acutely ill or intensive care or critical care or ICU* or coronary care unit or neurological intermediate care unit):ti,ab,kw
#14 MeSH descriptor: [Heart Arrest] explode all trees
#15 MeSH descriptor: [Myocardial Ischemia] explode all trees
#16 (cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct* or myocardial ischemia or acute coronary syndrome):ti,ab,kw
#17 MeSH descriptor: [Shock] explode all trees
#18 (shock):ti,ab,kw
#19 MeSH descriptor: [Meningitis] explode all trees
#20 (meningitis):ti,ab,kw
#21 MeSH descriptor: [Pneumonia] explode all trees
#22 (pneumonia):ti,ab,kw
#23 MeSH descriptor: [Pulmonary Disease, Chronic Obstructive] explode all trees
#24 (COPD or chronic obstructive pulmonary disease):ti,ab,kw
#25 MeSH descriptor: [Acute Lung Injury] explode all trees
#26 (acute lung injury):ti,ab,kw
#27 MeSH descriptor: [Respiratory Distress Syndrome, Adult] explode all trees
#28 (adult respiratory distress syndrome or ARDS):ti,ab,kw
#29 MeSH descriptor: [Pulmonary Embolism] explode all trees
#30 (pulmonary embolism or pulmonary infarct*):ti,ab,kw
#31 MeSH descriptor: [Multiple Trauma] explode all trees
#32 (severe trauma or multiple trauma):ti,ab,kw
#33 MeSH descriptor: [Craniocerebral Trauma] explode all trees
#34 (traumatic brain injury or TBI or head trauma or craniocerebral trauma):ti,ab,kw
#35 MeSH descriptor: [Stroke] explode all trees
#36 (stroke):ti,ab,kw
#37 MeSH descriptor: [Sepsis] explode all trees
#38 MeSH descriptor: [Shock, Septic] explode all trees
#39 (sepsis or septic shock):ti,ab,kw
#40 MeSH descriptor: [Intracranial Hemorrhages] explode all trees
#41 intracranial hemorrhage or subarachnoid hemorrhage or cerebral hemorrhage or intracranial bleeding or life-threatening bleeding:ti,ab,kw
#42 MeSH descriptor: [Poisoning] explode all trees
#43 (severe poisoning):ti,ab,kw
#44 MeSH descriptor: [Diabetic Ketoacidosis] explode all trees
#45 (diabetic ketoacidosis):ti,ab,kw

#46 MeSH descriptor: [Liver Failure, Acute] explode all trees
#47 (acute hepatic failure or fulminating hepatic failure):ti,ab,kw
#48 MeSH descriptor: [Acute Kidney Injury] explode all trees
#49 (acute kidney failure or acute renal injuries):ti,ab,kw
#50 MeSH descriptor: [Intestinal Perforation] explode all trees
#51 MeSH descriptor: [Appendicitis] explode all
#52 (intestinal perforation or appendicitis):ti,ab,kw
#53 (acute or emergency) near/2 (surgery or operat* or resection):ti,ab,kw
#54 (#8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36 or #37 or #38 or #39 or #40 or #41 or #42 or #43 or #44 or #45 or #46 or #47 or #48 or #49 or #50 or #51 or #52 or #53)
#55 (#7 and #54)

MEDLINE (Ovid) – 7099 hits

From 1950 to 17 October 2019

1. exp Hyperoxia/
2. exp Anoxia/
3. exp Oxygen Inhalation Therapy/
4. exp Oxygen/
5. ((inspir* or inhal* or fraction* or concentrat* or arterial* or saturation or level* or tension* or supply* or supplement* or supplie* or therap* or administr* or dosag* or dose* or dosing*) adj3 oxygen).tw.
6. (hyperoxia or hyperoxemia or hyperoxaemia or hypoxia or hypoxemia or hypoxaemia or anoxia or anoxemia or anoxaemia or arterial oxygen or high oxygen or oxygenat* or blood gas or oxygen saturation or pao2 or sao2 or spo2 or fio2).tw.
7. (1 or 2 or 3 or 4 or 5 or 6)
8. exp Critical Illness/
9. exp Critical Care/
10. exp Intensive Care Units/
11. exp Emergency Medicine/
12. exp Emergency Service, Hospital/
13. (emergency department* or ED or emergency room* or ER or high dependency unit* or HDU or prehospital* or critically ill or acutely ill or intensive care or critical care or ICU* or coronary care unit or neurological intermediate care unit).tw.
14. exp Heart Arrest/
15. exp Myocardial Ischemia/
16. (cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct* or myocardial ischemia or acute coronary syndrome).tw.
17. exp Shock/
18. shock.tw.
19. exp Meningitis/
20. meningitis.tw.
21. exp Pneumonia/
22. pneumonia.tw.
23. exp Pulmonary Disease, Chronic Obstructive/
24. (COPD or chronic obstructive pulmonary disease).tw.
25. exp Acute Lung Injury/
26. acute lung injury.tw.
27. exp Respiratory Distress Syndrome, Adult/
28. (adult respiratory distress syndrome or ARDS).tw.
29. exp Pulmonary Embolism/
30. (pulmonary embolism or pulmonary infarct*).tw.
31. exp Multiple Trauma/
32. (severe trauma or multiple trauma).tw.
33. exp Craniocerebral Trauma/
34. (traumatic brain injury or TBI or head trauma or craniocerebral trauma).tw.
35. exp Stroke/

36. stroke.tw.
37. exp Sepsis/
38. exp Shock, Septic/
39. (sepsis or septic shock).tw.
40. exp Intracranial Hemorrhages/
41. (intracranial hemorrhage or subarachnoid hemorrhage or cerebral hemorrhage or intracranial bleeding or life-threatening bleeding).tw.
42. exp Poisoning/
43. severe poisoning.tw.
44. exp Diabetic Ketoacidosis/
45. diabetic ketoacidosis.tw.
46. exp Liver Failure, Acute/
47. (acute hepatic failure or fulminating hepatic failure).tw.
48. exp Acute Kidney Injury/
49. (acute kidney failure or acute renal injuries).tw.
50. exp Intestinal Perforation/
51. exp Appendicitis/
52. (intestinal perforation or appendicitis).tw.
53. ((acute or emergency) adj2 (surgery or operat* or resection)).tw.
54. (8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53)
55. (7 and 54)
56. randomized controlled trial.pt.
57. controlled clinical trial.pt.
58. randomized.ab.
59. placebo.ab.
60. clinical trial.sh.
61. randomly.ab.
62. trial.ti.
63. (56 or 57 or 58 or 59 or 60 or 61 or 62)
64. exp animals/not humans.sh.
65. (63 not 64)
66. (55 and 65)

PubMed (NCBI) – 9428 hits

1966 to 17 October 2019 - 9428 hits

Search ((((((((((randomized controlled trial[Publication Type]) OR controlled clinical trial[Publication Type]) OR randomized[Title/Abstract]) OR placebo[Title/Abstract]) OR clinical trials[MeSH Terms]) OR randomly[Title/Abstract]) OR trial[Title])) NOT ((animals[MeSH Terms]) NOT humans[MeSH Terms]))) AND ((((((((Hyperoxia[MeSH Terms]) OR Anoxia[MeSH Terms]) OR Oxygen Inhalation Therapy[MeSH Terms]) OR Oxygen[MeSH Terms]) OR (((inspir*[Title/Abstract] OR inhal*[Title/Abstract] OR fraction*[Title/Abstract] OR concentrat*[Title/Abstract] OR arterial*[Title/Abstract] OR saturation[Title/Abstract] OR level*[Title/Abstract] OR tension*[Title/Abstract] OR supply*[Title/Abstract] OR supplement*[Title/Abstract] OR supplie*[Title/Abstract] OR therap*[Title/Abstract] OR administr*[Title/Abstract] OR dosag*[Title/Abstract] OR dose*[Title/Abstract] OR dosing*[Title/Abstract])) AND oxygen[Title/Abstract])) OR ((hyperoxia[Title/Abstract] OR hyperoxemia[Title/Abstract] OR hyperoxaemia[Title/Abstract] OR hypoxia[Title/Abstract] OR hypoxemia[Title/Abstract] OR hypoxaemia[Title/Abstract] OR anoxia[Title/Abstract] OR anoxemia[Title/Abstract] OR anoxaemia[Title/Abstract] OR arterial oxygen[Title/Abstract] OR high oxygen[Title/Abstract] OR oxygenat*[Title/Abstract] OR blood gas[Title/Abstract] OR oxygen saturation[Title/Abstract] OR pao2[Title/Abstract] OR sao2[Title/Abstract] OR

spo2[Title/Abstract] OR fio2[Title/Abstract]))))) AND ((((((((((((((((((((((((Critical Illness[MeSH Terms]) OR Critical Care[MeSH Terms]) OR Intensive Care Units[MeSH Terms]) OR Emergency Medicine[MeSH Terms]) OR Emergency Service, Hospital[MeSH Terms]) OR ((emergency department*[Title/Abstract] OR ED[Title/Abstract] OR emergency room*[Title/Abstract] OR ER[Title/Abstract] OR high dependency unit*[Title/Abstract] OR HDU[Title/Abstract] OR prehospital*[Title/Abstract] OR critically ill[Title/Abstract] OR acutely ill[Title/Abstract] OR intensive care[Title/Abstract] OR critical care[Title/Abstract] OR ICU*[Title/Abstract] OR coronary care unit[Title/Abstract] OR neurological intermediate care unit[Title/Abstract])) OR Heart Arrest[MeSH Terms]) OR Myocardial Ischemia[MeSH Terms]) OR ((cardiac arrest[Title/Abstract] OR cardiac failure[Title/Abstract] OR CPR[Title/Abstract] OR heart arrest[Title/Abstract] OR heart failure[Title/Abstract] OR myocardial infarct*[Title/Abstract] OR myocardial ischemia[Title/Abstract] OR acute coronary syndrome[Title/Abstract])) OR Shock[MeSH Terms]) OR shock[Title/Abstract]) OR ((Meningitis[MeSH Terms]) OR meningitis[Title/Abstract])) OR ((Pneumonia[MeSH Terms]) OR pneumonia[Title/Abstract])) OR ((Pulmonary Disease, Chronic Obstructive[MeSH Terms]) OR (COPD[Title/Abstract] OR chronic obstructive pulmonary disease[Title/Abstract]))) OR ((Acute Lung Injury[MeSH Terms]) OR lung injury[Title/Abstract])) OR ((Respiratory Distress Syndrome, Adult[MeSH Terms]) OR (adult respiratory distress syndrome[Title/Abstract] OR ARDS[Title/Abstract]))) OR ((Pulmonary Embolism[MeSH Terms]) OR (pulmonary embolism[Title/Abstract] OR pulmonary infarct*[Title/Abstract]))) OR ((Multiple Trauma[MeSH Terms]) OR (severe trauma[Title/Abstract] OR multiple trauma[Title/Abstract]))) OR ((Craniocerebral Trauma[MeSH Terms]) OR (traumatic brain injury[Title/Abstract] OR TBI[Title/Abstract] OR head trauma[Title/Abstract] OR craniocerebral trauma[Title/Abstract]))) OR ((Stroke[MeSH Terms]) OR stroke[Title/Abstract])) OR (((Sepsis[MeSH Terms]) OR Shock, Septic[MeSH Terms]) OR (sepsis[Title] OR septic shock[Title]))) OR ((Intracranial Hemorrhages[MeSH Terms]) OR (intracranial hemorrhage[Title/Abstract] OR subarachnoid hemorrhage[Title/Abstract] OR cerebral hemorrhage[Title/Abstract] OR intracranial bleeding[Title/Abstract] OR life-threatening bleeding[Title/Abstract]))) OR ((Poisoning[MeSH Terms]) OR severe poisoning[Title/Abstract])) OR ((Diabetic Ketoacidosis[MeSH Terms]) OR diabetic ketoacidosis[Title/Abstract])) OR ((Liver Failure, Acute[MeSH Terms]) OR (acute hepatic failure[Title/Abstract] OR fulminating hepatic failure[Title/Abstract]))) OR ((Acute Kidney Injury[MeSH Terms]) OR (acute kidney failure[Title/Abstract] OR acute renal injuries[Title/Abstract]))) OR ((((Intestinal Perforation[MeSH Terms]) OR Appendicitis[MeSH Terms]) OR (intestinal perforation[Title/Abstract] OR appendicitis[Title/Abstract])) OR (((acute[Title/Abstract] OR emergency)[Title/Abstract] AND (surgery[Title/Abstract] OR operat*[Title/Abstract] OR resection))[Title/Abstract])))

Embase (Ovid) – 9301 hits

From 1974 to 17 October 2019

1. *hyperoxia/
2. *hypoxia/
3. *oxygen therapy/
4. *oxygen/
5. *arterial oxygen saturation/
6. *oxygen blood level/
7. *arterial oxygen tension/
8. *blood oxygen tension/
9. ((inspir* or inhal* or fraction* or concentrat* or arterial* or saturation or level* or tension* or supply* or supplement* or supplie* or therap* or administr* or dosag* or dose* or dosing*) adj3 oxygen).tw.
10. (hyperoxia or hyperoxemia or hyperoxaemia or hypoxia or hypoxemia or hypoxaemia or anoxia or anoxemia or anoxaemia or arterial oxygen or high oxygen or oxygenat* or blood gas or oxygen saturation or pao2 or sao2 or spo2 or fio2).tw.
11. (1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10)
12. *critical illness/
13. *intensive care/
14. *intensive care unit/
15. *emergency medicine/

16. *emergency health service/
17. *coronary care unit/
18. (emergency department* or ED or emergency room* or ER or high dependency unit* or HDU or prehospital* or critically ill or acutely ill or intensive care or critical care or ICU* or coronary care unit or neurological intermediate care unit).tw.
19. *heart arrest/
20. *acute heart infarction/
21. (cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct* or myocardial ischemia or acute coronary syndrome).tw.
22. *shock/
23. shock.tw.
24. *meningitis/
25. meningitis.tw.
26. *pneumonia/
27. pneumonia.tw.
28. *chronic obstructive lung disease/
29. (COPD or chronic obstructive pulmonary disease).tw.
30. *acute lung injury/
31. acute lung injury.tw.
32. *adult respiratory distress syndrome/
33. (adult respiratory distress syndrome or ARDS).tw.
34. *lung embolism/
35. (pulmonary embolism or pulmonary infarct*).tw.
36. *multiple trauma/
37. (severe trauma or multiple trauma).tw.
38. *head injury/
39. *brain injury/
40. (traumatic brain injury or TBI or head trauma or craniocerebral trauma).tw.
41. *cerebrovascular accident/
42. *stroke unit/
43. stroke.tw.
44. *sepsis/
45. *septic shock/
46. (sepsis or septic shock).tw.
47. *brain hemorrhage/
48. (intracranial hemorrhage or subarachnoid hemorrhage or cerebral hemorrhage or intracranial bleeding or life-threatening bleeding).tw.
49. *intoxication/
50. severe poisoning.tw.
51. *diabetic ketoacidosis/
52. diabetic ketoacidosis.tw.
53. *acute liver failure/
54. (acute hepatic failure or fulminating hepatic failure).tw.
55. *acute kidney failure/
56. (acute kidney failure or acute renal injuries).tw.
57. *intestine perforation/
58. *appendicitis/
59. (intestinal perforation or appendicitis).tw.
60. ((acute or emergency) adj2 (surgery or operat* or resection)).tw.
61. (12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51

or 52 or 53 or 54 or 55 or 56 or 57 or 58 or 59 or 60)
62. (11 and 61)
63. CROSSOVER PROCEDURE.sh.
64. DOUBLE-BLIND PROCEDURE.sh.
65. SINGLE-BLIND PROCEDURE.sh.
66. (crossover* or cross over*).ti,ab.
67. placebo*.ti,ab.
68. (doubl* adj blind*).ti,ab.
69. allocat*.ti,ab.
70. trial.ti.
71. RANDOMIZED CONTROLLED TRIAL.sh.
72. random*.ti,ab.
73. (63 or 64 or 65 or 66 or 67 or 68 or 69 or 70 or 71 or 72)
74. (exp animal/ or exp invertebrate/ or animal.hw. or nonhuman/) not (exp human/ or human cell/ or (human or humans or man or men or wom?n).ti.)
75. (73 not 74)
76. (62 and 75)

Science Citation Index - Expanded (web of science) and Conference proceedings

From 1900 to 17 October 2019

#27 (#26 AND #25)
#26 TOPIC: (((random* OR control* OR RCT OR placebo OR group* OR trial*)))
#25 (#24 AND #3)
#24 (#23 OR #22 OR #21 OR #20 OR #19 OR #18 OR #17 OR #16 OR #15 OR #14 OR #13 OR #12 OR #11 OR #10 OR #9 OR #8 OR #7 OR #6 OR #5 OR #4)
#23 TITLE: (((acute or emergency) and (surgery or operat* or resection)))
#22 TOPIC: ((intestinal perforation or appendicitis))
#21 TOPIC: ((acute kidney failure or acute renal injuries))
#20 TOPIC: ((acute hepatic failure or fulminating hepatic failure))
#19 TOPIC: ((diabetic ketoacidosis))
#18 TOPIC: ((severe poisoning))
#17 TOPIC: ((intracranial hemorrhage or subarachnoid hemorrhage or cerebral hemorrhage or intracranial bleeding or life-threatening bleeding))
#16 TOPIC: ((sepsis or septic shock))
#15 TOPIC: (stroke)
#14 TOPIC: ((traumatic brain injury or TBI or head trauma or craniocerebral trauma))
#13 TOPIC: ((severe trauma or multiple trauma))
#12 TOPIC: ((pulmonary embolism or pulmonary infarct*))
#11 TOPIC: ((adult respiratory distress syndrome or ARDS))
#10 TOPIC: (acute lung injury)
#9 TOPIC: ((COPD or chronic obstructive pulmonary disease))
#8 TOPIC: (pneumonia)
#7 TOPIC: (meningitis)
#6 TOPIC: (shock)
#5 TOPIC: ((cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct* or myocardial ischemia or acute coronary syndrome))
#4 TOPIC: ((emergency department* or ED or emergency room* or ER or high dependency unit* or HDU or prehospital* or critically ill or acutely ill or intensive care or critical care or ICU* or coronary care unit or neurological intermediate care unit))
#3 (#2 OR #1)

#2 TITLE: (((hyperoxia or hyperoxemia or hyperoxaemia or hypoxia or hypoxemia or hypoxaemia or anoxia or anoxemia or anoxaemia or arterial oxygen or high oxygen or oxygenat* or blood gas or oxygen saturation or pao2 or sao2 or spo2 or fio2)))
#1 TITLE: ((((inspir* or inhal* or fraction* or concentrat* or arterial* or saturation or level* or tension* or supply* or supplement* or supplie* or therap* or administr* or dosag* or dose* or dosing*) and oxygen)))

BIOSIS Previews (web of science) – 4017 hits

From 1969 to 17 October 2019

#27 (#26 AND #25)
#26 TOPIC: ((random* OR control* OR RCT OR placebo OR group* OR trial*))
#25 (#24 AND #3)
#24 (#23 OR #22 OR #21 OR #20 OR #19 OR #18 OR #17 OR #16 OR #15 OR #14 OR #13 OR #12 OR #11 OR #10 OR #9 OR #8 OR #7 OR #6 OR #5 OR #4)
#23 TITLE: ((((acute or emergency) and (surgery or operat* or resection))))
#22 TOPIC: (((intestinal perforation or appendicitis)))
#21 TOPIC: (((acute kidney failure or acute renal injuries)))
#20 TOPIC: (((acute hepatic failure or fulminating hepatic failure)))
#19 TOPIC: (((diabetic ketoacidosis)))
#18 TOPIC: (((severe poisoning)))
#17 TOPIC: (((intracranial hemorrhage or subarachnoid hemorrhage or cerebral hemorrhage or intracranial bleeding or life-threatening bleeding)))
#16 TOPIC: (((sepsis or septic shock)))
#15 TOPIC: ((stroke))
#14 TOPIC: (((traumatic brain injury or TBI or head trauma or craniocerebral trauma)))
#13 TOPIC: (((severe trauma or multiple trauma)))
#12 TOPIC: (((pulmonary embolism or pulmonary infarct*)))
#11 TOPIC: (((adult respiratory distress syndrome or ARDS)))
#10 TOPIC: ((acute lung injury))
#9 TOPIC: (((COPD or chronic obstructive pulmonary disease)))
#8 TOPIC: ((pneumonia))
#7 TOPIC: ((meningitis))
#6 TOPIC: ((shock))
#5 TOPIC: (((cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct* or myocardial ischemia or acute coronary syndrome)))
#4 TOPIC: (((emergency department* or ED or emergency room* or ER or high dependency unit* or HDU or prehospital* or critically ill or acutely ill or intensive care or critical care or ICU* or coronary care unit or neurological intermediate care unit)))
#3 (#2 OR #1)
#2 TITLE: (((hyperoxia or hyperoxemia or hyperoxaemia or hypoxia or hypoxemia or hypoxaemia or anoxia or anoxemia or anoxaemia or arterial oxygen or high oxygen or oxygenat* or blood gas or oxygen saturation or pao2 or sao2 or spo2 or fio2)))
#1 TITLE: (((inspir* or inhal* or fraction* or concentrat* or arterial* or saturation or level* or tension* or supply* or supplement* or supplie* or therap* or administr* or dosag* or dose* or dosing*) and oxygen))

CINAHL (EBSCO) – 6314 hits

From inception to 17 October 2019

S66 (S53 AND S65)
S65 (S54 or S55 or S56 or S57 or S58 or S59 or S60 or S61 or S62 or S63 or S64)

S64 TX allocat* random*
S63 (MH "Quantitative Studies")
S62 (MH "Placebos")
S61 TX placebo*
S60 TX random* allocat*
S59 (MH "Random Assignment")
S58 TX randomi* control* trial*
S57 TX ((singl* n1 blind*) or (singl* n1 mask*)) or TX ((doubl* n1 blind*) or (doubl* n1 mask*)) or TX ((tripl* n1 blind*) or (tripl* n1 mask*)) or TX ((trebl* n1 blind*) or (trebl* n1 mask*))
S56 TX clinic* n1 trial*
S55 PT Clinical trial
S54 (MH "Clinical Trials+")
S53 (S7 AND S52)
S52 (S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35 OR S36 OR S37 OR S38 OR S39 OR S40 OR S41 OR S42 OR S43 OR S44 OR S45 OR S46 OR S47 OR S48 OR S49 OR S50 OR S51)
S51 AB ((acute or emergency)) AND AB ((surgery or operat* or resection))
S50 AB (intestinal perforation or appendicitis)
S49 MW Appendicitis
S48 MW Intestinal Perforation
S47 AB (acute kidney failure or acute renal injuries)
S46 MW acute kidney failure
S45 AB (acute hepatic failure or fulminating hepatic failure)
S44 MW Liver Failure, Acute
S43 AB diabetic ketoacidosis
S42 MW Diabetic Ketoacidosis
S41 AB severe poisoning
S40 MW Poisoning
S39 AB (intracranial hemorrhage or subarachnoid hemorrhage or cerebral hemorrhage or intracranial bleeding or life-threatening bleeding)
S38 MW Intracranial Hemorrhage
S37 AB (sepsis or septic shock)
S36 MW Shock, Septic
S35 MW Sepsis
S34 AB stroke
S33 MW Stroke
S32 AB (traumatic brain injury or TBI or head trauma or craniocerebral trauma)
S31 AB (severe trauma or multiple trauma)
S30 MW Multiple Trauma
S29 AB (pulmonary embolism or pulmonary infarct*)
S28 MW Pulmonary Embolism
S27 AB (adult respiratory distress syndrome or ARDS)
S26 MW Respiratory Distress Syndrome
S25 AB acute lung injury
S24 MW Acute Lung Injury
S23 MW (COPD or chronic obstructive pulmonary disease)
S22 MW Pulmonary Disease, Chronic Obstructive
S21 AB pneumonia
S20 MW Pneumonia
S19 AB meningitis
S18 MW Meningitis

S17 AB shock
S16 MW Shock
S15 AB (cardiac arrest or cardiac failure or CPR or heart arrest or heart failure or myocardial infarct* or myocardial ischemia or acute coronary syndrome)
S14 MW Myocardial Ischemia
S13 MW heart arrest
S12 AB (emergency department*) or (ED) or (emergency room*) or (ER) or (high dependency unit*) or (HDU) or (prehospital*) or (critically ill) or (acutely ill) or (intensive care) or (critical care) or (ICU*) or (coronary care unit) or (neurological intermediate care unit)
S11 MW emergency medicine
S10 MW intensive care units
S9 MW critical care
S8 MW critical illness
S7 (S1 OR S2 OR S3 OR S4 OR S5 OR S6)
S6 AB (hyperoxia) or (hyperoxemia) or (hyperoxaemia) or (hypoxia) or (hypoxemia) or (hypoxaemia) or (anoxia) or (anoxemia) or (anoxaemia) or (arterial oxygen) or (high oxygen) or (oxygenat*) or (blood gas) or (oxygen saturation) or (pao2) or (sao2) or (spo2) or (fio2)
S5 AB (((inspir*) or (inhal*) or (fraction*) or (concentrat*) or (arterial*) or (saturation) or (level*) or (tension*) or (supply*) or (supplement*) or (supplie*) or (therap*) or (administr*) or (dosag*) or (dose*) or (dosing*))) AND AB (oxygen)
S4 MW oxygen
S3 MW oxygen therapy
S2 MW anoxia
S1 MW hyperoxia

Latin American Caribbean Health Science Literature (LILACS) – 4103 hits

From inception to 17 October 2019

tw:((tw:((hyperoxia OR hyperoxemia OR hyperoxaemia OR hypoxia OR hypoxemia OR hypoxaemia OR anoxia OR anoxemia OR anoxaemia OR arterial oxygen OR high oxygen OR oxygenation OR blood gas OR oxygen saturation OR pao2 OR sao2 OR spo2 OR fio2) OR ((inspiratory OR inhalation OR fraction OR concentration OR arterial OR saturation OR level OR tension OR supply OR supplement OR supplied OR therapy OR administration OR dosage OR dose OR dosing) AND (oxygen)))) AND (tw:((acute surgery OR acute operation OR acute resection OR emergency surgery OR emergency operation OR emergency resection) OR (intestinal perforation OR appendicitis) OR (acute kidney failure OR acute renal injuries) OR (acute hepatic failure OR fulminating hepatic failure) OR (diabetic ketoacidosis) OR (severe poisoning) OR (intracranial hemorrhage OR subarachnoid hemorrhage OR cerebral hemorrhage OR intracranial bleeding OR life-threatening bleeding) OR (sepsis OR septic shock) OR (stroke) OR (traumatic brain injury OR tbi OR head trauma OR craniocerebral trauma) OR (severe trauma OR multiple trauma) OR (pulmonary embolism OR pulmonary infarction) OR (adult respiratory distress syndrome OR ards) OR (acute lung injury) OR (copd OR chronic obstructive pulmonary disease) OR (pneumonia) OR (meningitis) OR (shock) OR (cardiac arrest OR cardiac failure OR cpr OR heart arrest OR heart failure OR myocardial infarction OR myocardial ischemia OR acute coronary syndrome) OR (emergency department OR ed OR emergency room OR er OR high dependency unit OR hdu OR prehospital OR critically ill OR acutely ill OR intensive care OR critical care OR icu OR coronary care unit OR neurological intermediate care unit))) AND (tw:((randomized OR randomised OR random OR randomly OR control OR controlled OR rct OR placebo OR group OR trial))))

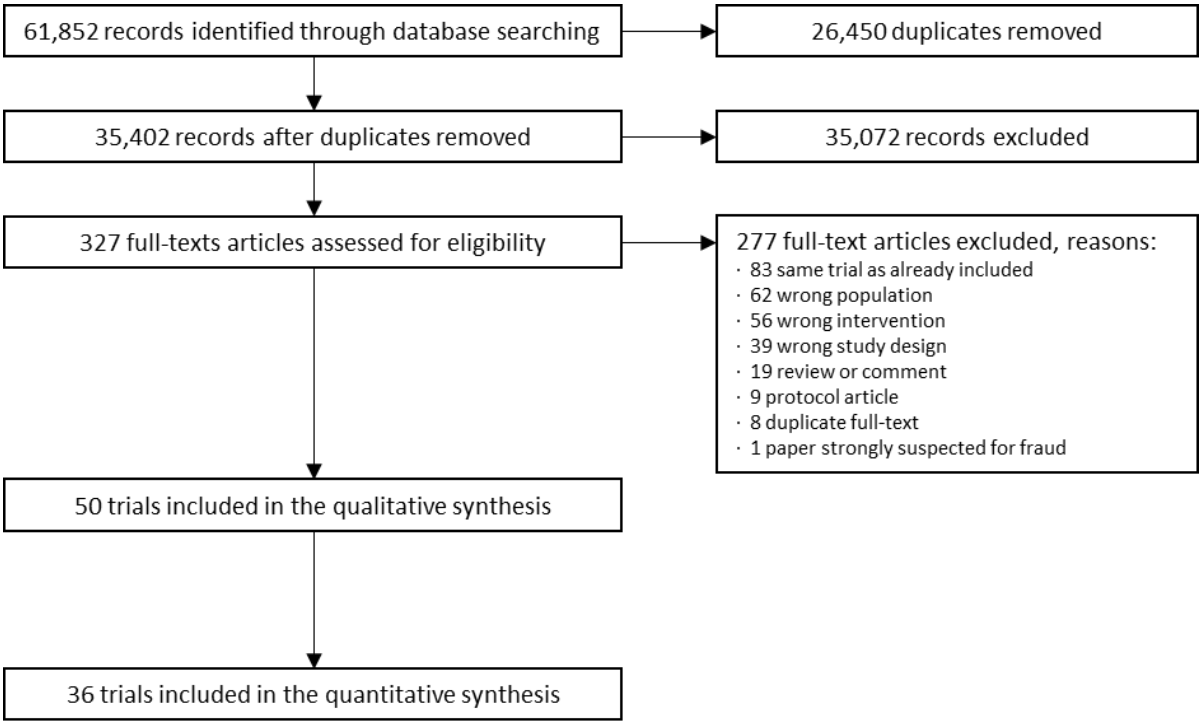
DATA COLLECTION FORM

General			Interventions				Sub.gr analyses											
Trial id	**Year**	**Publ. type**	**Exp gr. Intervention**	**Control gr. Intervention**	**Intervention period**	**Max follow-up**	Sub.Gr.1 **Overall RoB**	Sub.Gr.2 **Saturation/target used**	Sub.Gr.3 **Level of saturation/target in cont.gr.**	Sub.Gr4. Sub-pop. **ICU**	Sub.Gr.4. Sub-pop. **Random prior to hosp adm.**	Sub.Gr.4. Sub-pop. **Any cerebral disease**	Sub.Gr.4. Sub-pop. **Any cardiac disease**	Sub.Gr.4. Sub-pop. Any **trauma**	Sub.Gr.4. Sub-pop. **Out of hosp cardiac arrest**	Sub.Gr.4. Sub-pop. **Lung disease**	Sub.Gr.4. Sub-pop. **COPD**	Sub.Gr.5. **Above or below median duration**
Trial X																		
Trial Y																		
Trial Z																		

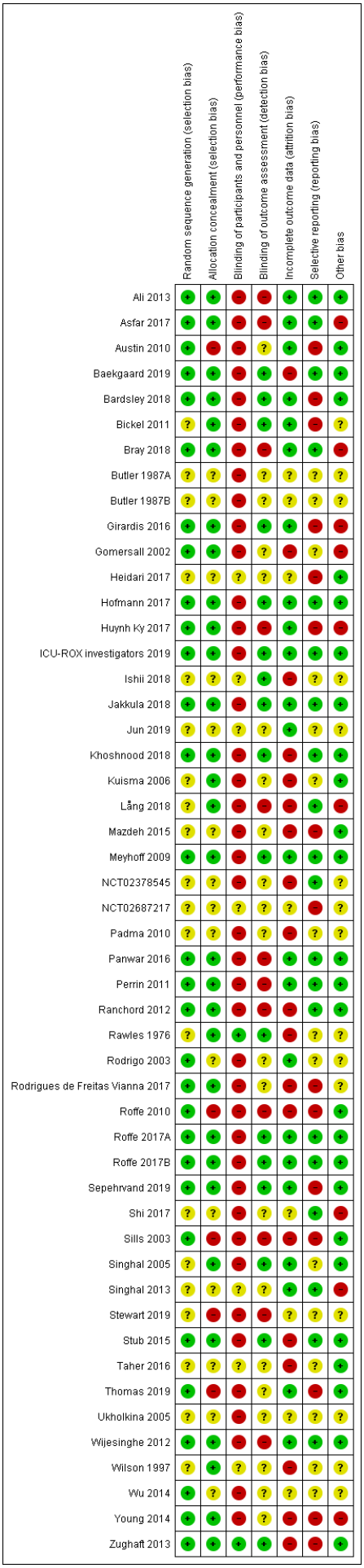
Randomisation and follow-up									Outcome X									
E: No randomised	C: No randomised	Total randomised	*E: Lost to follow-up*	*C: Lost to follow-up*	E: No analysed	C: No analysed	Total analysed	Notes	E: No events	E: No analysed	C: No events	C: No analysed	Blind pt+pers	Blind outcome assessor	Incomepl. outcome data	Selective outcome reporting	Time point used	Notes

The following data were collected: 1) Trial: country, date of publication; 2) Participants: numbers randomised, numbers analysed, numbers lost to follow up/withdrawn, type of population, age, sex, disease severity, setting, inclusion criteria, and exclusion criteria; 3) Interventions: intervention, comparator, duration and co-interventions; 4) Outcomes: predefined primary, secondary outcomes and timing of outcome measurement [1].

PRISMA FLOWCHART



RISK OF BIAS SUMMARY



DETAILS OF INCLUDED TRIALS AND RISK OF BIAS ASSESSMENT

Ali 2013 [2]		
Methods	Randomised clinical trial	
Participants	**Sample size:** n = 301 (experimental 155 (148 analysed), control 146 (141 analysed). **Sex (male gender):** experimental 44%, control 51% **Age (mean):** experimental 73, control 71 **Country:** UK **Setting:** patients with acute stroke admitted to a stroke unit **Inclusion criteria:** adult patients with a clinical diagnosis of acute stroke as defined by the World Health Organization if they were admitted to the University Hospital of North Staffordshire within the preceding 24 hours, were able to give informed consent, or a relative was contactable and willing to give assent, and if there was no clear indication for or against oxygen treatment. **Exclusion criteria:** patients with contraindications to fixed-dose oxygen treatment at a rate of 2 or 3 L/min (e.g. type II respiratory failure), patients where stroke was not the primary clinical problem, and patients with other serious life-threatening illnesses likely to lead to death within a few months **Disease severity:** not reported	
Interventions	**Experimental:** oxygen via nasal cannulae at a flow rate of 2 L/min if baseline oxygen saturation (SpO_2) was greater than 93% or 3 L/min if baseline SpO_2 was 93% or less for a period of 72 hours. **Control:** oxygen only when clinically indicated **Co-intervention:** participants who developed indications for oxygen, or needed a higher concentration of oxygen than the protocol prescribed, were given the appropriate concentration of oxygen by the treating clinician, irrespective of the treatment group. **Duration:** 72 hours.	
Outcomes	▪ Functional and quality of life outcomes Mortality **Timing of outcome measurement:** all at six months	
Notes	Quality of life was measured with both EQ-5D and EQ-VAS. We used the results form the EQ-VAS as the other trials reported quality-of-life using this scale. We used the RevMan calculater to calculated SD's. Email sent to Dr Roffe 16 August 2019 and reply was received.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computerised randomisation algorithm
Allocation concealment (selection bias)	Low risk	Telephone or web portal access to a remote centre
Blinding of participants and personnel (performance bias)	High risk	Single-blinded according to trial report and unblinded according to protocol. Patients and their doctors were aware of treatment allocation (confirmed by email).
Blinding of outcome assessment (detection bias)	High risk	Questionnaires were completed by the patient of their carer (confirmed by email). Unspecified for mortality.
Incomplete outcome data (attrition bias)	Low risk	4.5% withdrew consent in experimental group and 3.4% in control group. Reasons for withdrawal are given in trial report. EQ-5D: 110 completed the questionnaire in the experimental group (148-110-22 dead=16 did not), and 112 in the control group (141-112-21dead=8 did not). 11% in the experimental group did not complete EQ-5D and 6% in the control group. EQ-VAS: 81 completed the questionnaire in the experimental group (148-81-22dead=45 did not), and 96 in the control group (141-96-21dead=24 did not) 30% did not complete the EQ-VAS in the experimental group and 17% in the control group. 22 died in the experimental group and 21 died in the control group
Selective reporting (reporting bias)	Low risk	According to the protocol (supplementary material) the protocol was made before the initiation of the trial. According to the registration on ISRCTN, the trial was registered retrospectively.

Other bias	Low risk	The study was published by public grants
Asfar 2017 [3]		
Methods	Randomised clinical trial The trial was a 2-by-2 factorial trial randomizing to 4 groups. 2 groups were included in our analysis.	
Participants	**Sample size**: 442 randomized (219 experimental, 223 control) **Sex (male)**: experimental 63%, control 65% **Age (mean)**: experimental 67.8, control 66.3 **Country:** France **Setting**: patients with septic shock admitted to a multidisciplinary ICU **Disease severity score: SAPS III median 71** **Inclusion criteria:** ▪ Patients aged 18 years and older if they were mechanically ventilated, and exhibited septic shock refractory to fluid resuscitation as defined by an absence of response to 20 mL/kg of crystalloids or colloids and requiring vasopressor (norepinephrine or epinephrine, at a minimum infusion rate of 0·1 µg/kg per min); they also had to have been assessed within 6 hour after the initiation of vasopressors. Septic shock was defined by the presence of 2 or more diagnostic criteria of systemic inflammatory response syndrome, proven or suspected infection, and sudden dysfunction of at least 1 organ. **Exclusion criteria:** ▪ Severe hypoxaemia defined as PaO_2: FiO_2 ratio of less than 100 mm Hg for a minimum positive end-expiratory pressure of 5 cm H_2O ▪ Plasma sodium concentration of less than 130 mmol/L or more than 145 mmol/L ▪ Intracranial hypertension ▪ Patient admitted for cardiac arrest ▪ Overt cardiac failure ▪ Under legal guardianship ▪ No affiliation with the French health-care system ▪ Pregnancy ▪ Recent participation in another biomedical study or another interventional study with mortality as the primary endpoint ▪ An investigator's decision not to resuscitate	
Interventions	**Experimental**: hyperoxia group (mechanical ventilation with FiO_2 of 1.0 for 24 hours after inclusion; thereafter FiO_2 as in the normoxia group). Categorized by us as using a high target in the experimental group. **Control**: target SaO_2 of 88% to 95% using mechanical ventilation **Co-intervention**: not specified **Duration**: 24 hours	
Outcomes	**Primary outcome** ▪ Death from any cause at day 28 after inclusion **Secondary outcomes** ▪ 90-day mortality ▪ Daily sequential organ failure score (SOFA) from inclusion to day 7 ▪ 19 days alive and free from organ dysfunction at day 28 ▪ Length of stay in the ICU ▪ Alive at day 28 without organ support was defined as days alive without vasopressor infusion, mechanical ventilation, or renal replacement treatment ▪ Safety data (as specified on https://clinicaltrials.gov/) **Not prespecified outcomes** ▪ Patients with at least one serious adverse event ▪ Chest radiograph scores ▪ Atelectasis ▪ Pneumothorax ▪ Ventricular arrhythmias ▪ Mesenteric ischaemia ▪ Digital ischaemia	

	▪ ICU-acquired weakness ▪ Patients with ≥1 nosocomial infection during ICU stay ▪ Patients with ≥1 nosocomial pneumonia during ICU stay	
Notes	Email sent to Dr Asfar 5 December 2018 and reply was received.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated randomization list stratified by site and presence or absence of ARDS by use of permuted blocks of random sizes (nQuery Advisor 6.0).
Allocation concealment (selection bias)	Low risk	The pharmacists assigned a random number to each therapeutic package. The attribution of a given therapeutic package to a patient in accordance to the randomization list was done with a web-based secured randomization system (Clinsight software)
Blinding of participants and personnel (performance bias)	High risk	Unblinded
Blinding of outcome assessment (detection bias)	High risk	Unblinded. Unspecified for mortality.
Incomplete outcome data (attrition bias)	Low risk	2.7% in the experimental group and 0.9% in the control were excluded from analysis
Selective reporting (reporting bias)	Low risk	Protocol (clinicaltrials.gov) was pre-published, and all outcomes were reported on
Other bias	High risk	Early stopping bias: the trial was stopped after a pre-planned interim analysis, due to no prespecified criteria. The trial was funded by public grants (the French ministry of health).
Austin 2010 [4]		
Methods	Randomised clinical trial	
Participants	**Sample size:** 405 (experimental 226, control 179) **Sex (% male):** experimental 50%, control 46% **Age (mean):** experimental 69, control 69 **Country:** Australia **Setting:** Pre-hospital. Patients randomised by paramedics. **Inclusion criteria:** people aged 35 years or older with breathlessness and a history or risk of chronic obstructive pulmonary disease **Exclusion criteria:** asthma patients **Disease severity:** not reported	
Interventions	**Experimental:** high flow oxygen treatment (8-10 l/min) administered by a non-rebreather face mask and bronchodilators delivered by nebulisation with oxygen at flows of 6-8 l/min. **Control:** titrated oxygen treatment delivered by nasal prongs to achieve arterial oxygen saturations between 88% and 92%, with concurrent bronchodilator treatment administered by a nebuliser driven by compressed air **Co-intervention:** all patients received other standard treatment according to Tasmanian Ambulance Service guidelines, including basic support, nebulised bronchodilators (salbutamol 5 mg made up in 2.5 ml normal saline, ipratropiumbromide 500 µg made up with 2.5 ml normal saline), dexamethasone 8 mg intravenously, and, where necessary, salbutamol 200-300 mg intravenously or 500 mg intramuscularly. **Duration:** during prehospital transport (mean 47 minutes)	
Outcomes	**Primary outcome:** ▪ Hospital mortality	

	Secondary outcomes: ▪ Ventilation - invasive and non-invasive required during treatment by Ambulance officers and during hospital stay ▪ Arterial blood gas (ABG) results assessed within 30 min of arrival after treatment by ambulance for acute exacerbation of COPD ▪ Hospital admission within 30min of arrival ▪ Length of Hospital Stay **Timing of outcome measurements:** during hospital admission	
Notes	Email sent to Dr Austin 15 August 2019. Reminder sent 26 August 2019. No reply was received.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computerised random number generation
Allocation concealment (selection bias)	High risk	Closter randomised; personnel were aware of which group the next randomised patient was allocated to
Blinding of participants and personnel (performance bias)	High risk	Unblinded
Blinding of outcome assessment (detection bias)	Unclear risk	Unblinded. Unspecified for mortality.
Incomplete outcome data (attrition bias)	Low risk	Results given for both intention-to-treat and per-protocol analysis
Selective reporting (reporting bias)	High risk	Protocol was registered retrospectively
Other bias	Low risk	The Australian College of Ambulance Professionals (ACAP) provided funding. FlaemNova, Milan, Italy, donated Walkie nebulisation air compressors. Neither of the study sponsors had a role in study design; data collection, analysis, or interpretation; or the writing of the report.
Baekgaard 2019 [5]		
Methods	Randomised clinical trial	
Participants	**Sample size:** 41 (experimental 20 (18 analysed), control 21 (20 analysed)) **Sex (% male):** experimental 80%, control 76.2% **Age (mean):** experimental 60%, control 50% **Country:** Denmark **Setting:** Trauma centre **Inclusion criteria:** patients above 18 years of age with blunt or penetrating trauma, that generated a trauma team activation and were directly transferred from the scene of accident to our trauma centre **Exclusion criteria:** patients in cardiac arrest before/on admission, patients with a suspicion of smoke inhalation, and patients not admitted to a hospital ward after the initial treatment in the trauma bay **Disease severity:** First GCS in the trauma bay 13.5 in experimental group and 13.0 in control group	
Interventions	**Experimental:** Non-intubated patients received 15 L/min via a non-rebreather mask and intubated patients received a FiO_2 of 1.0 in the trauma bay and during intra-hospital transportation. In the operating room, ICU, post-anaesthesia care unit and ward the FiO_2 could be reduced to 0.8 if an arterial oxygen saturation ≥98% was obtained. **Control:** The lowest dosage of oxygen (≥21%) that ensured an arterial oxyhaemoglobin saturation (SpO_2) target of 94% either using mechanical ventilation, a non-rebreather mask, a nasal cannula or no supplementary oxygen was applied. Supplemental oxygen was not given unless the SpO_2 was below 94% and thus, only spontaneously breathing patients without supplementary oxygen	

	could saturate above 94%. In case the SpO_2 became unmeasurable, the intervention was interrupted, and standard (liberal) treatment was applied **Co-intervention:** **Duration:** 24 hours
Outcomes	**Primary outcome:** ▪ Evaluate feasibility of maintenance of normoxia within the first 24 hours after trauma **Secondary outcomes:** ▪ 30-day mortality ▪ Major pulmonary complications (combined endpoint). Major pulmonary complications included pneumonia, acute respiratory distress syndrome (ARDS) and acute lung injury (ALI). ▪ In-hospital sepsis ▪ Surgical site infection ▪ Number of days on mechanical ventilation ▪ Hospital- and intensive care unit length of stay (LOS) ▪ Glasgow Outcome Scale Extended (GOSE) score at day 30 **Timing of outcome measurements:** 30 days
Notes	Email sent to Dr Baekgaard 10 October 2019 and reply was received

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomisation module in Research Electronic Data Capture (REDCap) was used. The randomisation table was generated outside of REDCap in the statistical software R by a statistician otherwise not involved in the study
Allocation concealment (selection bias)	Low risk	The randomisation table was generated outside of REDCap in the statistical software R by a statistician otherwise not involved in the study
Blinding of participants and personnel (performance bias)	High risk	Open label
Blinding of outcome assessment (detection bias)	Low risk	In-hospital pulmonary complications was evaluated independently by two attending anaesthesiologists blinded to the patients' allocation. Blinding was ensured by providing the assessors with patient charts, imaging studies and laboratory values masked to treatment allocation. Unspecified for mortality.
Incomplete outcome data (attrition bias)	High risk	3/41 patients withdrew consent and 5/41 (12%) (not similar in both groups: 4 in the control group and 1 in the experimental group) were lost to 30-day follow-up (except for mortality)
Selective reporting (reporting bias)	Low risk	The protocol was pre-registered, and all outcomes were reported on
Other bias	Low risk	The trial was funded by public funds (clarified by email)

Bardsley 2018 [6]

Methods	Randomised clinical trial
Participants	**Sample size:** 90 (experimental 45, control 45) **Sex (% male):** experimental 38%, control 53% **Age (mean):** experimental 70, control 72 **Country:** New Zealand **Setting:** hospitalised patients with exacerbations of chronic obstructive pulmonary disease **Inclusion criteria:** hospital inpatients, ≥40 years of age, with an admission diagnosis of AECOPD **Exclusion criteria:** requirement for ≥4 L/min of oxygen via nasal cannulae to maintain SpO_2 between 88 to 92%; current requirement for non-invasive ventilation (NIV); baseline transcutaneous partial pressure of carbon dioxide ($PtCO_2$) > 60 mmHg; inability to provide written informed consent; and any other condition which at the Investigator's discretion, was believed

	may present a safety risk or impact on the feasibility of the study results **Disease severity:** not reported
Interventions	**Experimental:** 8 L/min oxygen by nebuliser mask **Control:** 8L/min air by nebuliser mask **Co-intervention:** 2.5 mg salbutamol **Duration:** 15 minutes
Outcomes	**Primary outcome:** ▪ $PtCO_2$ (transcutaneous CO_2) **Secondary outcomes:** ▪ $PcapCO_2$ ▪ Proportion of participants who had a rise in $PtCO_2$ or $PcapCO_2$ of ≥4 and ≥ 8 mmHg ▪ Capillary pH ▪ Heart rate ▪ SpO_2 **Timing of outcome measurements:** 35 minutes No relevant outcomes reported
Notes	Email sent to Dr Beasley 11 October 2019. Reply was received.

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer-generated random numbers
Allocation concealment (selection bias)	Low risk	Sealed opaque envelopes
Blinding of participants and personnel (performance bias)	High risk	Participants were blinded. An initial 15 min wash-in and titration period was administered by an unblinded investigator using nasal cannulae, if required, to ensure that participant's SpO_2 were within 88 to 92%
Blinding of outcome assessment (detection bias)	Low risk	Investigator who recorded heart rate and $PtCO_2$ were blinded.
Incomplete outcome data (attrition bias)	Low risk	2/90 (1 in each group withdrew)
Selective reporting (reporting bias)	High risk	The protocol was pre-registered. Primary outcome changed due to difficulties in obtaining adequate amounts of blood to fill the capillary tubes from some participants was difficult (missing data)
Other bias	Low risk	The trial was supported by public grants

Bickel 2011 [7]

Methods	Randomised clinical trial
Participants	**Sample size:** 210 (experimental 107, control 103) **Sex (male%):** experimental 75%, control 71% **Age (mean):** experimental 28,5, control 27,6 **Country:** Israel **Setting:** Acute surgery for appendicitis **Inclusion criteria:** adult patients (aged 15 years) having an open appendectomy for acute appendicitis **Exclusion criteria:** patients with chronic obstructive pulmonary disease, severe malnutrition (serum albumin concentration 3 g/dL; to convert to grams per litre, multiply by 10), or

	immunodeficiency disease **Disease severity:** not reported
Interventions	**Experimental:** FIO_2 80% (combined with 20% air). In the recovery room following completion of the operation, the patients received high-flow oxygen (10 L/min) through a nonrebreathing mask with a reservoir for 2 hours **Control:** FiO_2 30% and 70% nitrogen. In the recovery room following completion of the operation, the patients received oxygen (4 L/min) by nasal cannula for 2 hours **Co-intervention:** preoperative antibiotics against gram-negative and anaerobic bacteria were given to all patients, including intravenous aminoglycosides (gentamicin sulfate, 5 mg/kg) and metronidazole (500 mg). When intraoperative findings indicated gangrenous or perforated appendicitis, antibiotic treatment lasted for 5 days. Anesthesia was introduced with fentanyl citrate (1-5 µg/kg), propofol (2 mg/kg) or thiopental sodium (4 mg/kg), and rocuronium bromide (0.5 mg/kg) or atracurium besylate (0.5 mg/kg), following pre-oxygenation by mask. Midazolam maleate was additionally prescribed (1-2 mg). Anesthesia was maintained with nitrous oxide with oxygen (in a ratio depending on group selection), isoflurane, 1%, rocuronium bromide or atracurium besylate (0.10-0.15 µg/kg), and fentanyl citrate (5-10 µg on demand). **Duration:** during surgery and until 2 hours after (mean 2,5h hour)
Outcomes	▪ Surgical site infection 14 days of surgery ▪ Duration of postoperative hospitalisation No relevant outcomes reported
Notes	Email sent to Dr Bickel 15 August 2019. Reminder sent 26 August 2019. No reply was received

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described
Allocation concealment (selection bias)	Low risk	Sealed envelopes
Blinding of participants and personnel (performance bias)	High risk	Patients and the surgical team (including the investigators) were blinded, however other personnel were not blinded
Blinding of outcome assessment (detection bias)	Low risk	Surgical wounds were evaluated daily by the residents and senior surgeons, all blinded to the FIO_2 assignment
Incomplete outcome data (attrition bias)	Low risk	No patient was lost to follow-up
Selective reporting (reporting bias)	High risk	Protocol was registered retrospectively
Other bias	Unclear risk	It was unclear how the trial was funded

Bray 2018 [8]	
Methods	Randomised clinical trial
Participants	**Sample size:** 62 (experimental 25, control 37) **Sex (male %):** experimental 75%, control 86% **Age (mean):** experimental 60,5, control 64 **Country:** Australia **Setting:** out-of-hospital cardiac arrest **Inclusion criteria:** adults (age ≥18 years), unconscious (Glasgow Coma Scale<9) with an advanced airway (endotracheal tube [ETT] or supraglottic airways [SGA]) in situ and sustained ROSC following an OHCA of presumed cardiac cause, and an initial monitored rhythm assessed as shockable (ventricular fibrillation or pulseless ventricular tachycardia) **Exclusion criteria:** paramedic witnessed OHCAs, $Sp0_2$ <95% or no pulse oximetry trace, known or

	suspected to be pregnant; dependant on others for activities of daily living (i.e. facilitated care or nursing home residents); and evidence of a "Not for Resuscitation" order
Interventions	Following ROSC, oxygen was delivered at a flow rate of ≥10 L/min via bag-valve reservoir (BVR) connected to the airway until a satisfactory pulse oximeter trace and reading was achieved. If eligibility criteria were met, the paramedics randomised the patients to: **Experimental:** continue to receive oxygen at ≥10 L/min **Control:** oxygen reduced to either 2 or 4 L/min (target 90-94%) **Co-intervention:** all the standard post resuscitation treatments were given as per current ambulance Clinical Practice Guidelines, except for the amount of oxygen delivered **Duration:** duration of pre-hospital transport (50 minutes)
Outcomes	▪ SpO_2 ≥94% on arrival at hospital ▪ SpO_2 ≥90% on arrival at hospital ▪ Re-arrest during ambulance transport Survival at hospital discharge
Notes	Email sent to Dr Bray 15 August 2019. Reminder sent 26 August 2019 and reply was received.

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated allocation
Allocation concealment (selection bias)	Low risk	Opaque envelopes
Blinding of participants and personnel (performance bias)	High risk	Paramedics and data collectors not blinded
Blinding of outcome assessment (detection bias)	High risk	Paramedics and data collectors not blinded. Unspecified for mortality.
Incomplete outcome data (attrition bias)	Low risk	No patients were lost in follow-up, however, one patient in the experimental group requested data withdrawal
Selective reporting (reporting bias)	Low risk	A study protocol was pre-registered prior to randomisation (NCT02499042)
Other bias	High risk	Prof Finn, Dr Hein and Dr Bray received funding from the National Health and Medical Research Council (NHMRC) Centre of Research Excellence: Australian Resuscitation Outcomes Consortium (Aus-ROC). At the time of this study, Drs Bray and Stub received a Heart Foundation Fellowship and Prof Cameron received a NHMRC Fellowship. Dr Stub also received a Viertel Charitable Foundation Grant. The trial appeared to be free of other components that could put it at risk of other bias Early stopping bias: the study was stopped early due to high numbers of desaturation in the titrated arm.

Butler 1987A [9]

Methods	Randomised clinical trial (skin oxygen study)
Participants	**Sample size:** 20 (experimental 10, control 10) **Sex (% male):** not reported **Age (mean):** not reported **Country:** UK **Setting:** patients undergoing below knee amputation **Inclusion criteria:** patients admitted to a vascular unit and requirering major amputation for ischemia **Exclusion criteria:** visible ischaemic demarcation above a suitable level for below-knee

	amputation or severe disease of the ipsilateral knee joint precluding satisfactory prosthetic fitting **Disease severity:** not reported
Interventions	**Experimental:** FiO_2 0,28 oxygen by ventimask postoperatively **Control:** No supplemental oxygen **Co-intervention:** light gauze dressings were used. The patients had physiotherapy **Duration:** 48 hours
Outcomes	**Primary outcome:** ▪ Transcutaneous pO_2 measurements **Timing of outcome measurements:** 1 day prior to surgery and 1, 2, 7 and 14 days post-operatively No relevant outcomes reported
Notes	Contact details were not identified; thus, email was not sent. The publication by Butler et al. reports on two trials; this extraction concerns the *Skin oxygen trial*

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	High risk	Unblinded
Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	Unclear risk	It is unclear whether all patients completed the trial
Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	Unclear risk	It was unclear how the trial was funded

Butler 1987B [9]	
Methods	Randomised clinical trial (healing study)
Participants	**Sample size:** 39 (experimental 17, control 22) **Sex (% male):** experimental 65%, control 59% **Age (mean):** experimental 71, control 66 **Country:** UK **Setting:** patients undergoing below knee amputation **Inclusion criteria:** patients admitted to a vascular unit and requirering major amputation for ischemia **Exclusion criteria:** visible ischaemic demarcation above a suitable level for below-knee amputation or severe disease of the ipsilateral knee joint precluding satisfactory prosthetic fitting **Disease severity:** not reported
Interventions	**Experimental:** FiO_2 0,28 oxygen by ventimask postoperatively **Control:** No supplemental oxygen **Co-intervention:** light gauze dressings were used. The patients had physiotherapy **Duration:** 48 hours
Outcomes	**Primary outcome:** ▪ Stump healing

	Timing of outcome measurements: 1 year	
Notes	Contact details were not identified; thus, email was not sent. The publication by Butler et al. reports on two trials; this extraction concerns the *Healing trial*	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	High risk	Unblinded
Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	Unclear risk	It is unclear how many patients were lost to follow-up/analysed (Table 2). We used no lost to follow-up in our analyses
Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	Unclear risk	It was unclear how the trial was funded
Girardis 2016 [10]		
Methods	Randomised clinical trial	
Participants	**Sample size**: 480 (experimental 244, control 236) **Sex (male %)**: experimental 57%, control 56% **Age (median)**: experimental 65, control 63 **Country:** Italy **Setting**: multidisciplinary ICU **Disease severity score**: SAPS II score median 38 **Inclusion criteria** ▪ All patients aged 18 years or older and admitted to the ICU with an expected length of stay of 72 hours or longer **Exclusion criteria** ▪ Age younger than 18 years ▪ Pregnancy ▪ ICU readmission ▪ A decision to withhold life-sustaining treatment ▪ Immunosuppression or neutropenia ▪ Enrolment in another study Patients with acute decompensation of chronic obstructive pulmonary disease and acute respiratory distress syndrome with a PaO_2:FiO_2 ratio le	
Interventions	**Experimental**: oxygen therapy was administered according to standard ICU practice; FiO_2 of at least 0.4, allowing PaO_2 values up to 150 mm Hg and an SpO_2 between 97% and 100%. If the SpO_2 decreased below 95% to 97%, the FiO_2 was increased to reach the target value of SpO_2. Patients received FiO_2 of 1.0 during intubation, airway suction, or hospital transfer. **Control**: oxygen therapy was administered at the lowest possible FiO_2 to maintain the PaO_2 between 70 and 100 mm Hg or SpO_2 values between 94% and 98%. FiO_2 was gradually reduced or oxygen supplementation discontinued whenever the PaO_2 or SpO_2 exceeded 100 mm Hg or 98%. Supplemental oxygen was administered only if SpO_2 decreased below 94%. Categorized by us as using a high target in the control group.	

	Co-intervention: not specified **Duration**: until ICU discharge
Outcomes	▪ ICU mortality ▪ New-onset respiratory, cardiovascular, liver, and renal failure (defined as a SOFA score ≥ 3 for the corresponding organ) occurring 48 hours or more after ICU admission ▪ Need for reoperation in surgical patients ▪ Bloodstream, respiratory, and surgical site infections (defined according to Centres for Disease Control and Prevention definitions). Only microbiologically documented bloodstream and respiratory tract infections were considered **Un-prespecified secondary outcomes** ▪ Hospital mortality ▪ Ventilation-free hours during the ICU stay
Notes	Email sent 6 December 2018 to Dr Girardis and reply was received.

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated random numbers
Allocation concealment (selection bias)	Low risk	The randomisation sequence was concealed from the researchers by use of sequentially numbered, closed, opaque envelopes that were opened after patient study inclusion
Blinding of participants and personnel (performance bias)	High risk	Open label
Blinding of outcome assessment (detection bias)	Low risk	Not described; however, blinding of outcome assessment was clarified by email. Unspecified for mortality.
Incomplete outcome data (attrition bias)	Low risk	Results from Intention-to-treat analyses are given in the supplementary. 2 patients withdrew consent, randomization group for these 2 patients were not reported; thus can't be included in the sensitivity analysis on losses to follow-up. Outcome respiratory failure: 18 in experimental and 15 in control group were lost to follow-up
Selective reporting (reporting bias)	High risk	The trial was registered retrospectively (NCT01319643)
Other bias	High risk	Early stopping bias: the trial was stopped after a not pre-planned interim analysis. The trial was funded by public grants

Gomersall 2002 [11]

Methods	Randomised clinical trial
Participants	**Sample size**: 36 (experimental 19, control 17) **Sex (males %)**: experimental 82%, control 76% **Age (mean)**: experimental 68, control 69 **Country:** Hong Kong **Setting**: patients with acute exacerbation of chronic obstructive pulmonary disease admitted to a multidisciplinary ICU **Disease severity score:** not reported **Inclusion criteria**

	▪ Patients admitted with a clinical diagnosis of an acute exacerbation of chronic obstructive pulmonary disease and a PaO_2 < 6.6 kPa (50 mm Hg), and $PaCO_2$ > 6.6 kPa (50 mm Hg) on air. **Exclusion criteria** ▪ Chest radiologic signs of pulmonary oedema, lung cancer, pneumothorax, or pneumonia ▪ If the participant already met study criteria for mechanical ventilation ▪ Mechanical ventilation for respiratory failure twice in the preceding 6 months ▪ Inability to walk more than 20 yards on flat ground ▪ Co-existing terminal disease
Interventions	Oxygen therapy was provided via a Venturi-type mask and adjusted according to the results of arterial blood samples with the aim of reaching the desired target oxygen tension within 1 hour of entry to the trial. **Experimental**: target PaO_2 above 9.0 kPa (70 mm Hg) (categorized by us as using a low target in the experimental group) **Control:** target PaO_2 of >6.6 kPa (50 mm Hg). Categorized by us as using a low target in the control group. **Co-intervention**: patients in the low-oxygen tension group also received doxapram if they developed an acidosis with pH < 7.2, whereas those in the high-oxygen tension group received doxapram if they developed symptomatic acidosis. Bronchodilator, steroid, and antibiotic therapy was standardized. **Duration**: treatment protocols, including oxygen therapy, were continued after discharge from the ICU until oxygen therapy was no longer considered necessary
Outcomes	▪ Need for mechanical ventilation ▪ Duration of hospital stay ▪ Cardiac arrhythmia ▪ Mortality ▪ Coma **Timing of outcome measurements:** not specified
Notes	Email sent to Dr Gomersall 6 December 2018, but no reply was received.

Risk of bias assessment

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer generated numbers
Allocation concealment (selection bias)	Low risk	Unmarked sealed envelopes
Blinding of participants and personnel (performance bias)	High risk	Only patients were blinded
Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	High risk	2/19 (11%) of the patients in the experimental group were excluded from analysis due to protocol violation
Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	High risk	Difference between doxapram co-intervention between groups. The trial was funded by public grants.

Heidari 2017 [12]	
Methods	Randomised clinical trial

Participants	**Sample size:** 79 (experimental 39, control 40) **Sex (% male):** experimental 68%, control 46% **Age (mean):** experimental 59, control 60 **Country:** Iran **Setting:** patients with non-ST-segment elevation acute coronary syndrome admitted to the emergency ward **Inclusion criteria:** age between 18-84 years; diagnosed with acute coronary syndrome with non ST segment elevation according to Branvald criteria; no clinical evidence of heart failure; no chronic lung disease or other respiratory problems; lack of cardiac arrest or cardiogenic shock before entering the hospital; oxygen saturation above 90% on admission; absence of congenital heart disease **Exclusion criteria:** need for inotropic support; having ST elevation acute myocardial infarction; oxygen saturation less than 90% during hospitalization; emergency coronary angioplasty or emergency coronary artery bypass during hospitalization; death **Disease severity:** not reported
Interventions	**Experimental:** 4-6L/min oxygen with nasal cannula **Control:** 4-6L/min room air with nasal cannula **Co-intervention:** not reported **Duration:** 6 hours
Outcomes	**Primary outcome:** ▪ Cardiac dysrhythmias; timepoint: continues over 24 hours ▪ Chest pain; timepoint: 24, 28, 32, 36, 40, 44, 48 hours ▪ The amount of narcotic analgesic; timepoint: 48 hours **Secondary outcomes:** ▪ Readmission due to cardiac problems; timepoint: end of weeks 1, 2, 3, 4 ▪ Visit due to cardiac problems; timepoint: end of weeks 1, 2, 3, 4 **Timing of outcome measurements:** see individual outcomes (timepoints as specified in protocol)
Notes	Email sent to Dr Rahzani 11 October 2019. Reminder sent 18 October 2019. Reply was not received.

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	Unclear risk	Stated that the trial was triple blinded, but it was unclear who was blinded and how blinding was maintained
Blinding of outcome assessment (detection bias)	Unclear risk	The trial was described as triple blinded, but it was unclear who was blinded and how
Incomplete outcome data (attrition bias)	Unclear risk	Unclear how many patients were analysed. We assume the authors report by protocol.
Selective reporting (reporting bias)	High risk	Protocol retrospectively registered. The authors did not report on the secondary outcomes. Outcomes were only reported at 24 hours.
Other bias	Low risk	The trial was funded by public grants

Hofmann 2017 [13,14]	
Methods	Randomised clinical trial

Participants	**Sample size:** 6629 (experimental 3311, control 3318) **Sex (male %):** experimental 68%, control 71% **Age (median):** experimental 68, control 68 **Country:** Sweden **Setting:** patients with suspected acute myocardial infarction **Inclusion criteria:** patients who presented to the ambulance services, emergency departments, coronary care units, or catheterization laboratories of participating hospitals were evaluated for eligibility. Trial participants were required to be 30 years of age or older and to have symptoms suggestive of myocardial infarction (defined as chest pain or shortness of breath) for less than 6 hours, an oxygen saturation of 90% or higher on pulse oximetry, and either electrocardiographic changes indicating ischemia or elevated cardiac troponin T or I levels on admission (i.e., above the locally defined decision limit for the identification of myocardial infarction) **Exclusion criteria:** patients who were receiving ongoing oxygen therapy, as well as those who presented with a cardiac arrest or had a cardiac arrest between presentation and enrolment (for whom high-flow oxygen therapy would normally be provided)
Interventions	**Experimental:** oxygen therapy (6 litres per minute for 6 to 12 hours delivered through an open face mask) **Control:** ambient air **Co-intervention:** not specified **Duration:** 6-12 hours
Outcomes	▪ Death from any cause within 365 days after randomisation ▪ Death from any cause within 30 days after randomisation ▪ Re-hospitalisation with myocardial infarction ▪ Re-hospitalisation with heart failure ▪ Cardiovascular death ▪ Composites of these end points **Timing of outcome measurement:** 30 days and 365 days
Notes	Email sent to Dr Hofmann 15 August 2019 and reply was received. Dr Hofmann clarified that they had data on several of our pre-defined outcomes, but that it would take some work and time to retrieve them. Due to our strict deadline, we declined to wait for these data.

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer-generated list was performed with the use of an online randomisation module
Allocation concealment (selection bias)	Low risk	Online randomisation module
Blinding of participants and personnel (performance bias)	High risk	Open label
Blinding of outcome assessment (detection bias)	Low risk	It was unclear whether the personnel diagnosing the patients with myocardial infarction, atrial fibrillation, atrioventricular block, cardiogenic shock and cardiac arrest were blinded. It was clarified by email that these data, including mortality, come from registries (SWEDEHEART, National Population Registries) and that all investigators including the statistician were blinded until code break.
Incomplete outcome data (attrition bias)	Low risk	Less than 5% were lost to follow-up
Selective reporting (reporting bias)	Low risk	All pre-specified outcomes were reported on
Other bias	Low risk	Supported by the Swedish Heart–Lung Foundation, the Swedish Research Council, and the Swedish Foundation for Strategic Research. The funding agencies had no access to the trial data and no role in the trial design, implementation, or reporting. No sponsorship or funding from industry or for-profit sources was received for the trial.

Huynh Ky 2017 [15]		
Methods	Randomised clinical trial	
Participants	**Sample size:** 60 (experimental SpO_2 92% 20, experimental SpO_2 97% 20 (1 lost to follow-up), control 20) **Sex:** overall 73% men **Age:** overall 63 years **Country:** Canada **Setting:** acute phase of acute coronary syndrome **Inclusion criteria:** patients with acute coronary syndrome **Exclusion criteria:** severe COPD patients	
Interventions	**Experimental 1:** automated oxygen titration with $FreeO_2$ targeting SpO_2 92% **Experimental 2:** automated oxygen titration with $FreeO_2$ targeting SpO_2 97% **Control:** manual administration of oxygen (target unknown) **Co-intervention:** not specified **Duration:** maximum 24 hours	
Outcomes	▪ Frequency of desaturation (SpO_2 < 90% for 30 s) ▪ Frequency of arrhythmias ▪ Rate of tachycardia episodes ▪ Level of cardiac enzymes in patients with acute coronary disease **Timing of outcome measurements:** not stated	
Notes	The two experimental groups are extracted and compare - as target in the control is not reported. Email sent to Dr Lellouche 15 August 2019. Reminder sent 26 August 2019 and reply was received	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated random number (random.org) was used (clarified by email)
Allocation concealment (selection bias)	Low risk	Sealed envelopes (clarified by email)
Blinding of participants and personnel (performance bias)	High risk	Single blinded
Blinding of outcome assessment (detection bias)	High risk	Not blinded (clarified by email). Unspecified for mortality.
Incomplete outcome data (attrition bias)	Low risk	Less than 5% of the patients were withdrawn
Selective reporting (reporting bias)	High risk	Protocol was not published (clarified by email)
Other bias	High risk	Funded by local funds, the $FreeO_2$ prototypes were provided by Oxynov. Competing interests Dr Lellouche: co-development of $FreeO_2$, co-founder and administrator of Oxynov, the company that commercialize $FreeO_2$.
ICU-ROX investigators 2019 [16]		
Methods	Randomised clinical trial	
Participants	**Sample size**: 1000 (experimental 501, control 499) **Sex (male)**: experimental 63%, control 63% **Age**: experimental 58, control 58 **Country**: New Zealand **Setting**: mechanically ventilated adults admitted to a multidisciplinary ICU **Disease severity score:** APACHE II score median 23.5 **Inclusion criteria:** People at least 18 years of age who require invasive mechanical ventilation in	

	the ICU and are expected to be receiving mechanical ventilation beyond the next calendar day **Exclusion criteria:** ▪ Greater than 2 hours of invasive mechanical ventilation or non-invasive ventilation, or both, in an ICU during this hospital admission (includes time ventilated in another hospital's ICU) ▪ In the view of the treating clinician, hyperoxia is clinically indicated for reasons including (but not limited to) carbon monoxide poisoning or a requirement for hyperbaric oxygen therapy ▪ In the view of the treating clinician, avoidance of hyperoxia is clinically indicated for reasons including (but not limited to) chronic obstructive airways disease (COPD), paraquat poisoning, previous exposure to bleomycin, or chronic hypercapnic respiratory failure ▪ Pregnancy ▪ Death is deemed to be inevitable as a result of the current acute illness and either the treating clinician, the participant, or the substitute decision maker are not committed to full active treatment ▪ Adults with a life expectancy of less than 90 days due to a chronic or underlying medical condition ▪ Admitted following a drug overdose (including alcohol intoxication) ▪ Long-term dependence on invasive ventilation prior to this acute illness ▪ Confirmed or suspected diagnosis of any of the following: Guillain-Barré syndrome, cervical cord injury above C5, muscular dystrophy, or motor neuron disease ▪ Enrolment not considered in the participant's best interests ▪ Enrolled in any other trial of targeted oxygen therapy ▪ Previously enrolled in the ICU-ROX study
Interventions	**Experimental**: no specific measures taken to avoid high FiO_2 or SpO_2, FiO_2<0.30 discouraged (thus, we could not categorize the experimental group as either using a low or a high target). Patients assigned to the 'higher group' received 'standard care' both while ventilated and after extubation with no specific measures taken to avoid high FiO_2 or high SpO_2. The use of upper alarm limits for SpO_2 in the 'higher group' was prohibited as upper alarm limits for SpO_2 were not used as part of standard care. The lower limit alarm for SpO_2 was set at 90% (or lower if clinically appropriate). If the PaO_2 or the SaO_2 were lower than the acceptable limit, inspired oxygen might be increased if clinically appropriate, irrespective of the SpO_2 reading. The use of an FiO_2 of less than 0.3 whilst ventilated was discouraged. **Control**: target SaO_2/SpO_2 91% to 96%. When a participant was allocated to conservative oxygen therapy, the inspired oxygen concentration was decreased to room air as rapidly as possible provided that the SpO_2 measured by peripheral pulse oximetry was greater than the acceptable lower limit. SpO_2 levels of greater than 96% were strictly avoided and an upper SpO_2 alarm limit of 97% applied whenever supplemental oxygen was administered in the ICU to minimise the risk of hyperoxaemia. After extubation, in the conservative oxygen group, the upper monitored alarm limit of acceptable SpO_2 of 97% was applied whenever supplemental oxygen was being administered. In the event that the SpO_2 exceeded the acceptable upper limit, downward titration of supplemental oxygen was undertaken as a high priority and supplemental oxygen was discontinued as soon possible. The lower limit alarm for SpO_2 was set at 90% (or lower if clinically appropriate). If the PaO_2 or the SaO_2 were lower than the acceptable limit, inspired oxygen might be increased if clinically appropriate, irrespective of the SpO_2 reading. Categorized by us as using a low target in the control group. **Co-intervention**: no restrictions to concomitants treatments provided to participants. If an increase in FiO_2 for procedures performed in the ICU included (but not limited to) bronchoscopy, suctioning, tracheostomy or preparation for extubation, this was permitted in both groups **Duration**: until death or discharge from the ICU, or day 28 post randomization
Outcomes	**Primary outcome:** ▪ Ventilator free days to day 28 **Secondary outcomes:** ▪ All-cause mortality (day 90 and 180) ▪ Duration of survival ▪ Proportion of participants in paid employment at baseline who were unemployed at 180 days ▪ Cognitive function at day 180

	▪ Quality of life at day 180 ▪ Cause-specific mortality ▪ Functional outcome assessed by the extended Glasgow outcome scale (in patients with acute brain pathology)	
Notes		
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Secure web-based randomisation
Allocation concealment (selection bias)	Low risk	Central randomization
Blinding of participants and personnel (performance bias)	High risk	Not blinded
Blinding of outcome assessment (detection bias)	Low risk	Centralised assessors masked to study-group assignments undertook day 180 assessments of cognitive function, quality of life, and function. Unspecified for mortality.
Incomplete outcome data (attrition bias)	Low risk	Less than 5% (3.5%) were lost to follow-up
Selective reporting (reporting bias)	Low risk	The trial was registered prior to randomization. Only feasibility outcomes were reported, and outcomes pre-specified in protocol will be reported in final trial report including 1000 participants. However, mortality is reported in total (30.3%), but is not specified to treatment group.
Other bias	Low risk	The trial was supported by public funds. The trial appeared to be free of other components that could put it at risk of bias
Ishii 2018 [17]		
Methods	Randomised clinical trial	
Participants	**Sample size**: 44 (experimental 21, control 23) **Sex**: not specified **Age**: not specified **Country**: Japan **Setting**: surgical ICU **Disease severity score:** not reported **Inclusion criteria**: mechanically ventilated patients admitted to surgical ICU for more than 12 hours **Exclusion criteria**: not specified	
Interventions	**Experimental**: FiO_2 of 1.0 using high flow nasal cannula. Categorized by us as using a high target in the experimental group. **Control**: expected FiO_2 to achieve a PaO_2 of 100 mmHg (13.3 kPa) using high flow nasal cannula The interventions are "non-invasive", as the interventions are initiated after extubation (of the mechanical ventilated) where after oxygen are administered via high-flow-nasal cannula. Categorized by us as using a low target in the control group. **Co-intervention**: not specified **Duration**: one hour	
Outcomes	▪ Atelectasis **Timing of outcome measurement:** not specified No relevant outcomes reported	
Notes	Email sent 6 December 2018 to Dr Ishii. Reminder sent 15 August 2019. Reply was not received.	

Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	It was stated that the trial was randomised, but the method of sequence generation was not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	Unclear risk	The trial was described as double-blinded, but it was unclear who was blinded and how blinding was maintained
Blinding of outcome assessment (detection bias)	Low risk	Radiologist was blinded
Incomplete outcome data (attrition bias)	High risk	14% were lost to follow-up. To which group these patients were randomised to was not reported
Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	Unclear risk	It was unclear how the trial was funded
Jakkula 2018 [18]		
Methods	Randomised clinical trial The trial was a RCT with a 2 factorial design. We only extracted data from the normoxia and moderate hyperoxia groups	
Participants	**Sample size**: 123 (experimental 60, control 63) **Sex (male)**: experimental 48%, control 50% **Age**: experimental 60, control 59 **Country**: Finland **Setting**: adults admitted to the ICU after out-of-hospital cardiac arrest (OHCA) **Disease severity score:** APACHE II score median 28 **Inclusion criteria:** ▪ Adults resuscitated from witnessed OHCA with ventricular fibrillation (VF) or ventricular tachycardia (VT) as the initial rhythm. In addition, all of the following inclusion criteria had to be met: ・ return of spontaneous circulation (ROSC) 10 to 45 minutes from the onset of cardiac arrest. ・ confirmed or suspected cardiac origin of the arrest. ・ mechanical ventilation upon ICU arrival. ・ markedly impaired level of consciousness defined as no response to verbal commands and Glasgow coma scale (GCS) motor score < 5 (withdrawal to painful stimuli at best). ・ deferred consent from next of kin possible or likely; and ・ active intensive care and targeted temperature management (TTM) initiated. **Exclusion criteria:** ▪ Adults with confirmed or suspected acute or pre-existing intracranial pathology or suspicion of increased intracranial pressure, or both ▪ Adults with severe oxygenation failure defined as PaO_2/FiO_2 (fraction of inspired oxygen) < 100 mmHg upon arrival to ICU and no improvement in oxygenation after adding sufficient PEEP level ▪ Severe chronic obstructive pulmonary disease ▪ Age < 18 or > 80 years ▪ Pregnancy	

Interventions	**Experimental**: target 20 to 25 kPa (150 to 187.5 mmHg). Categorized by us as using a high target in the experimental group. **Control**: target PaO_2 target 10 to 15 kPa (75 to 112.5 mmHg) or target SpO_2 of 95% to 98%. Categorized by us as using a high target in the control group. **Co-intervention**: all adults received targeted temperature management (TTM) at 33 °C or 36 °C and were sedated according to the treating clinicians' instructions. All adults received standard care, monitoring and assessments based on the protocol of the ICU, including direct blood pressure monitoring via an arterial catheter. **Duration**: 36 hours
Outcomes	**Primary outcome:** ▪ NSE serum concentration at 48 hours after cardiac arrest **Secondary outcomes:** ▪ NSE serum concentration at 24 and 72 hours after cardiac arrest ▪ S100 protein serum concentration at 24, 48, and 72 hours after cardiac arrest ▪ Cardiac troponin (TnT) concentration at 24, 48, and 72 hours after cardiac arrest ▪ Results of NIRS monitoring during the first 48 hours after admission to the ICU ▪ Results of continuous EEG monitoring for 48 hours after arrival at the ICU and a statement of the findings by an experienced senior neurologist or neurophysiologist ▪ Cerebral performance category (CPC) at 6 months after cardiac arrest ▪ Total duration of intensive care ▪ Total duration of mechanical ventilation ▪ Length of hospital stay ▪ Discharge destination ▪ Vital status at hospital discharge (dead or alive) **Feasibility outcomes:** ▪ Difference in $PaCO_2$ between groups targeting low to normal (4.5 to 4.7 kPa) and high to normal (5.8 to 6.0) $PaCO_2$ ▪ Difference in PaO_2 between groups targeting low to normal (10 to 15 kPa) and high to normal (20 to 25 kPa) PaO_2 ▪ Difference in MAP between groups targeting low to normal (65 to 75 mmHg) and high to normal (80 to 100 mmHg) MAP ▪ Distribution of values for primary and secondary outcomes ▪ Randomized or screened participant ratio ▪ Consent rate ▪ Data completion rate ▪ Recruitment duration
Notes	Email sent 6 December 2019 to Dr Jakkula. Reminder was sent 15 August 2019. Reply was not received.

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated random numbers
Allocation concealment (selection bias)	Low risk	Web based system
Blinding of participants and personnel (performance bias)	High risk	The treating personnel was not blinded from the treatment targets

Blinding of outcome assessment (detection bias)	Low risk	The neurophysiologist analysing the EEG results and the neurologist evaluating the neurologic recovery of the participants were blinded to the study group allocations. Unspecified for mortality.
Incomplete outcome data (attrition bias)	Low risk	Less than 5% were lost to follow-up
Selective reporting (reporting bias)	Low risk	The trial was registered prior to randomization
Other bias	Low risk	The trial was funded by public and private funds. The funding bodies had no input regarding the design, management, or reporting of the trial. The trial appeared to be free of other components that could put it at risk of bias
Jun 2019 [19]		
Methods	Randomised clinical trial The trial was a three-arm trial comparing oxygen via nasal catheter in one group and invasive mechanical ventilation in two groups. We only extracted data from the two groups of mechanical ventilation.	
Participants	**Sample size:** 58 (experimental 29, control 29) **Sex (male %):** not specified **Age:** not specified **Country:** not specified **Setting:** patients with acute exacerbations of chronic obstructive pulmonary disease and acute myocardial infarction **Inclusion criteria:** elderly with acute exacerbations of chronic obstructive pulmonary disease and acute myocardial infarction who could not receive percutaneous coronary intervention **Exclusion criteria:** patients with shock and malignant arrhythmia	
Interventions	**Experimental:** invasive mechanical ventilation with FiO_2 50-70% the first 48 hours, hereafter gradually decreased to 40-50% **Control:** invasive mechanical ventilation with FiO_2 30-50% **Co-intervention:** heparin sodium continuous venous pump, anti-anxiety and expansion of coronary therapy **Duration:** not specified	
Outcomes	▪ 14-day mortality ▪ malignant arrhythmia ▪ myocardial infarction recurrence rate **Timing of outcome measurement:** not specified	
Notes	Results were published in an abstract. Only results on recurrent myocardial infarction was reported with the number of events. Contact details were not identified, thus, email requesting additional information about the trial was not sent.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method od sequence generation was not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	Unclear risk	Not described
Blinding of outcome assessment (detection bias)	Unclear risk	Not described

Incomplete outcome data (attrition bias)	Low risk	No patient was lost to follow-up
Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	Unclear risk	It was unclear how the trial was funded
Khoshnood 2018 [20,21]		
Methods	Randomised clinical trial	
Participants	**Sample size:** 94 (experimental 49, control 45) **Sex (male %):** experimental 63%, control 71% **Age:** experimental 63, control 66 **Country:** Sweden **Setting:** Patients with ST elevation myocardial infarction randomised in ambulance **Inclusion criteria:** normoxic patients (SpO_2 ≥ 94% on room air) with a first time STEMI accepted for acute percutaneous coronary intervention (PCI), and with symptom duration of less than 6 h **Exclusion criteria:** patients with a previous AMI or inability to make a decision to participate	
Interventions	**Experimental:** 10L/min via open design OxyMask (vent) **Control:** room air vis open design OxyMask **Co-intervention:** apart from the study intervention, the patients received standard care in the ambulance and were treated with aspirin, ticagrelor, heparin, β-blockers, and morphine as needed **Duration:** from randomisation in ambulance to end of primary PCI (experimental 87 minutes, control 86 minutes)	
Outcomes	▪ MSI on CMR ▪ IS on CMR ▪ MaR on CMR ▪ Ejection fraction on CMR ▪ Microvascular obstruction on CMR ▪ Pain difference (visual analog scale) at randomization vs. at PCI balloon inflation start ▪ Doses of opioids (substance and mg) and β-blockers (substance and mg) given before and during the PCI ▪ SpO_2 change from inclusion to PCI start ▪ IS as measured in hospital with the area under the troponin T curve (first 24 h) ▪ ST-segment elevation resolution ▪ TIMI flow during PCI ▪ Use of heart failure medications (e.g. β-blockers, angiotensin converting enzyme inhibitors, angiotensin receptor blockers, diuretics, digoxin, and sinus node inhibitors) at 6 months ▪ Subjectively perceived health (EQ-5D) at 6 months WMSI on echocardiography ▪ Assessment of remodelling by quantification of LV volumes, LVEF, and WMSI at index hospitalization to 6 months ▪ Mortality is reported, although not pre-specified in protocol **Timing of outcome measurement:** mortality and quality of life reported at 6 months follow-up	
Notes	Email sent to Dr Khoshnood 15 August 2019 and reply was received.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated random numbers
Allocation concealment (selection bias)	Low risk	Sealed envelopes
Blinding of participants and personnel (performance bias)	High risk	Only patients were blinded to the intervention

Blinding of outcome assessment (detection bias)	Low risk	The observers for MaR and IS were blinded to all clinical data. All analyses were performed by researchers blinded to the group allocation (according to protocol). Unspecified for mortality.
Incomplete outcome data (attrition bias)	High risk	Additional information from the trial found in the 2018 sub-publication in European Journal of Emergency Medicine, revealed that a high proportion (39 and 37%) of the randomised patients were excluded from the analyses.
Selective reporting (reporting bias)	Low risk	All outcomes stated in the protocol were assessed
Other bias	Low risk	The trial was supported by public grants
Kuisma 2006 [22]		
Methods	Randomised clinical trial	
Participants	**Sample size:** 28 (experimental 14, control 14) **Sex (male):** experimental 71%, control 93% **Age:** experimental 64, control 62 **Country:** Finland **Setting:** early post-resuscitation **Inclusion criteria:** patients with a bystander witnessed out-of-hospital ventricular fibrillation **Exclusion criteria:** not specified	
Interventions	**Experimental:** FiO_2 100% **Control:** FiO_2 30% **Co-intervention:** during CPR all patients were ventilated with 100% oxygen. **Duration:** 60 minutes	
Outcomes	▪ Serum NSE and S-100 levels at 24 and 48h after ROSC ▪ Adequacy of oxygenation at 10 and 60 min after ROSC ▪ The need for to raise the Fi02 to avoid hypoxaemia in the group which was ventilated with 30% oxygen ▪ Mortality at hospital discharge	
Notes	Email sent to Dr Kuisma 15 August 2019. Reminder sent 26 August 2019. No reply was received	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	It was stated the trial was randomised, but the sequence generation was not described
Allocation concealment (selection bias)	Low risk	Envelopes were used
Blinding of participants and personnel (performance bias)	High risk	Unblinded
Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	High risk	13% of the randomised participants were excluded from analyses - and the participants were not specified by randomisation group
Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	Low risk	The trial was supported by public grants
Lång 2018 [23]		
Methods	Randomised clinical trial	
Participants	**Sample size**: 65 (experimental 38, control 27) **Sex (male)**: experimental 82%, control 85%	

	Age: experimental 45, control 43 **Country**: Finland **Setting**: mechanically ventilated adults with traumatic brain disease admitted to the ICU **Disease severity score:** APACHE II score median 22 **Inclusion criteria:** ▪ Isolated non-penetrating TBI or adults with multiple trauma with TBI with Glasgow coma scale (GCS) eight or less (inclusive), expected need for intubation and mechanical ventilation > 24 hour ▪ Recruitment within 18 hours after admission to ICU ▪ Time from TBI < 36 hours ▪ Informed consent from next of kin **Exclusion criteria:** ▪ Age < 18 or > 65 years ▪ Anticipated brain death in 12 hours or otherwise moribund adults expected to die in 24 hours ▪ Expected need for mechanical ventilation < 24 hours ▪ Insufficient oxygenation assessed by a clinician ▪ Adults with multiple trauma with brain injury and severe abdominal, thoracic or pelvic injury possibly affecting oxygenation ▪ No consent ▪ Insufficient oxygenation with the treatment modality of the lower oxygenation group ($Pa0_2$ < 13 kPa or SpO_2 < 95% with $Fi0_2$ 0.40 and PEEP of 10) ▪ Oxygenation failure probable during ICU care ▪ Penetrating TBI ▪ Suspected pregnancy (perform urinary or serological pregnancy test if suspected)
Interventions	**Experimental:** FiO_2 of 0.70. Categorized by us as using a high target in the experimental group. **Control**: FiO_2 of 0.40. Categorized by us as using a low target in the control group. **Co-intervention**: not specified **Duration**: maximum 14 days
Outcomes	▪ Laboratory markers during the first 3 days ▪ Pulmonary function (P/F ratio, ARDS, atelectasis, pneumonia) ▪ Length of mechanical ventilation ▪ Length of ICU stay ▪ Length of hospital stay ▪ Death at six months ▪ Extended Glasgow outcome scale at six months **Timing of outcome measurements:** maximum 6 months
Notes	Email sent 6 December 2018 to Dr Lång and a reply was received.

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomized, but the method of sequence generation was not described
Allocation concealment (selection bias)	Low risk	Sealed, opaque envelopes
Blinding of participants and personnel (performance bias)	High risk	Open label
Blinding of outcome assessment (detection bias)	High risk	Only the neurologist assessing the neurological outcomes was blinded. Unspecified for mortality.

Incomplete outcome data (attrition bias)	High risk	8% were lost to follow-up and were not specified to allocation group in publication. Number of participants lost to follow-up in each group was clarified by email.
Selective reporting (reporting bias)	Low risk	The trial was registered on clinicaltrials.gov prior to randomization, however quality of life is not reported; however trial authors are planning to publish these results
Other bias	High risk	Unplanned trial stop. It was unclear how the trial was funded. The trial was supported by Kuopio University hospital (according to protocol).
Mazdeh 2015 [24]		
Methods	Randomised clinical trial	
Participants	**Sample size**: 51 (experimental 26, control 25) **Sex (men)**: experimental 54%, control 56% **Age**: not specified **Country**: Iran **Setting**: adults with stroke initially referred to the department of neurology, but admitted to the ICU **Disease severity score:** not reported **Inclusion criteria:** ▪ Age between 40 and 70 years ▪ GCS > 12 and adults with isolated brain damage and intact airway control ▪ Ischaemic and haemorrhagic stroke with no need for surgical intervention ▪ Less than 12 hours have passed from the accident ▪ NIHSS square between 7 and 9 Participants were selected from adults referred to the department of neurology in Farshchian hospital, an affiliated hospital of Hamadan university of medical sciences. The participants were admitted to the ICU and monitored carefully by expert nurses. Due to participants being transferred from the department of neurology and transferred to the ICU to be monitored, we do not regard these adults as typical adults admitted to the ICU. **Exclusion criteria:** ▪ Adults under 40 and older than 70 years ▪ Adults with diabetes mellitus and ischaemic heart disease, renal failure, acute pulmonary oedema, history of massive myocardial infarction and heart failure ▪ Adults who need intubation on arrival to the hospital ▪ Adults with a baseline blood pressure of less than 90/60, or hypoxia ▪ Adults requiring surgical intervention (i.e. acute subdural haematoma and cerebral haemorrhage) ▪ Adults with blood pressure greater than 170/90 in the first 12 hours of the incident ▪ Adults with successful cardiopulmonary resuscitation (CPR) within 12 hours ▪ History of previous stroke or unconsciousness, resulting in the need for intubation and mechanical ventilation ▪ Death or lost to follow-up ▪ Adults in the control group where oxygen therapy was inevitable for them	
Interventions	**Experimental**: FiO_2 of 0.5 - oxygen therapy with Venturi mask (categorized by us as using a low target in the experimental group) **Control**: no supplemental oxygen was administered. Categorized by us as using a low target in the control group. **Co-intervention**: routine medication (as stated in protocol) **Duration**: oxygen administration was given in the first 12 hours of admission	
Outcomes	▪ Good recovery and number of complications in the first day of admission, before discharge, and 6 months after discharge using ranking scale and Barthel index (as stated in protocol)	

	▪ Not pre-specified outcome: mortality	
Notes	Email sent 6 December 2018 to Dr Seifirad who forwarded the email on to Dr. Mazdeh, however, no reply was received. Overall poor reporting quality.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomized, but the method of sequence generation was not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	High risk	Unblinded
Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	High risk	1 out of 52 (2%) randomized participants were lost to follow-up and not described in the manuscript. It is not stated to which group this person was allocated. Participants in the control group in which oxygen therapy was inevitable were excluded
Selective reporting (reporting bias)	High risk	We judged the trial to be registered retrospectively. It was registered 13 November 2013 and submitted to journal 30 December 2013
Other bias	Low risk	The trial was funded by a public hospital (Vic-chancellor of research and Technology, Hamadan medical university). The trial appeared to be free of other components that could put it at risk of bias
Meyhoff 2009 [25]		
Methods	Randomised clinical trial Only data from the acute operated patients are used	
Participants	**Sample size:** 385 (experimental 190, control 195) **Sex (% men):** we do not have this information solely for the acute operated patients. Total distribution: experimental 42%, control 42% **Age (median):** we do not have this information solely for the acute operated patients. Total distribution: experimental 64 years, control 64 years **Country:** Denmark **Setting:** patients undergoing acute abdominal surgery **Inclusion criteria:** Overall patients were eligible if they were 18 years or older and scheduled to undergo acute or elective laparotomy. We only included patient undergoing acute laparotomy. When laparotomy was indicated for a gynaecological disease, only patients with suspected malignancy were included **Exclusion criteria:** operations performed under general anesthesia within 30 days, chemotherapy for malignancy within 3 months, inability to provide informed consent, and preoperative arterial hemoglobin oxygen saturation below 90% without supplemental oxygen assessed by pulse oximetry	
Interventions	**Experimental:** FiO_2 of 0.80 **Control:** FiO_2 of 0.30 In both groups, FIO_2 was increased if hypoxia was detected or suspected to ensure arterial oxygen saturation greater than 94% and arterial oxygen tension greater than 9 kPa. **Co-intervention:** The trial protocol included several important aspects of perioperative care, including epidural analgesia, control of temperature and glucose level, absence of preoperative oral bowel preparation, and standardized anesthesia without nitrous oxide. The protocol recommended cefuroxime (1.5 g) and	

	metronidazole (1 g) given intravenously as standard antibiotic choice, but ampicillin (2 g) or benzylpenicillin (2 million IU) in combination with gentamicin (0.240 g) and metronidazole (1 g) were also allowed. Fewer antibiotics were required in the case of elective cholecystectomy or laparotomies with no potential contamination. We considered "timely" administration of antibiotics as administration of the first and second antibiotic within 60 minutes prior to skin incision. Perioperative fluids were given only to replace measured or calculated deficits (no thirdspace loss), aiming at a postoperative body weight increase of less than 1 kg. Blood loss was replaced 1:1 with colloids, and blood transfusion was initiated if blood loss exceeded 20 mL/kg. Anesthesia was either inhalational or total intravenous anesthesia, determined entirely by the attending anesthetist. **Duration:** during and two hours after surgery
Outcomes	**Primary outcome:** ▪ Surgical site infection within 14 days, defined according to the Centers for Disease Control and Prevention. **Secondary outcomes:** ▪ Atelectasis ▪ Pneumonia ▪ Respiratory failure ▪ Mortality **Additional outcomes:** ▪ Re-operation ▪ Sepsis ▪ Myocardial infarction ▪ Lung embolism ▪ Stroke ▪ Serious adverse events **Timing of outcome measurements:** 30 and 90 days
Notes	The PROXI trial included both acute and elective patients. Randomisation included stratification for acute and elective surgery. We therefore contacted corresponding author to ask for data/results on only the acute patients. These were received and used in the analyses. Email sent to Dr Meyhoff 21 June 2019 and reply was received.

Risk of bias assessment

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated allocation list using study center, diabetes mellitus, acute or elective operations, and body mass index as stratification variables
Allocation concealment (selection bias)	Low risk	Central interactive voice-response system
Blinding of participants and personnel (performance bias)	High risk	*"Cardboard shields were placed on the side of the anaesthesia machines to keep the surgical team blinded to group allocation. In the post anaesthesia care unit, opaque bags covered the flow meters. Information about perioperative FiO_2, arterial oxygen partial pressure (PaO_2) as well as flow of oxygen and air was collected on a separate paper form, placed in a sealed opaque envelope when patients were discharged from the post anaesthesia care unit. The patients were not informed of their group allocation during the trial or follow-up".* Surgical team blinded. Participants blinded. Personnel in the post anaesthesia care unit blinded. Outcome assessor blinded. Anaesthetists not blinded. The trial was conducted according to highest standards; however, incidence of mortality and lung outcomes may have been

		influenced by the actions of the unblinded anaesthesiologist during surgery.
Blinding of outcome assessment (detection bias)	Low risk	The Steering Committee was blinded and had no access to patient allocation during the trial. An independent statistician analysed the PROXI data under code (treatment A and B) and prepared a blinded version of the results. Mortality data retrieved by register.
Incomplete outcome data (attrition bias)	Low risk	All patients were included in the analyses
Selective reporting (reporting bias)	Low risk	Study protocol was published prior to randomisation and all pre-specified outcomes were reported on
Other bias	Low risk	Funding/Support: The trial was supported by the Danish Medical Research Council (271-05-0206), the Council, the Novo Nordisk Foundation, the Aase and Ejnar Danielsens Foundation (105728), the A. P. Møller Foundation for the Advancement of Medical Science, the Danish Society of Anaesthesiology and Intensive Care Medicine's Research Initiative, the Beckett-Foundation, the Brødrene Hartmanns Foundation, and the Etly and Jørgen Stjerngrens Foundation. Role of the Sponsors: The funding sources had no role in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; or the preparation, review, or approval of the manuscript.
NCT02378545 [26]		
Methods	Randomised clinical trial	
Participants	**Sample size: 50** (experimental 25, control 25) **Sex (% male):** experimental 28%, control 52% **Age (mean):** experimental 69, control 59 **Country:** UK **Setting:** patients with sepsis presenting to the emergency department by ambulance **Inclusion criteria:** ▪ Adult patients aged 18 years or above ▪ Diagnosed with presumed 'sepsis' ▪ Arrive at Emergency Department by ambulance ▪ Provision of informed consent ▪ Willing to allow their General Practitioner and consultant, if appropriate, to be notified of participation in the study **Exclusion criteria:** ▪ Female participants who are pregnant ▪ Existing diagnosis of chronic obstructive pulmonary disease (COPD) ▪ A primary diagnosis (or suspected diagnosis) of an acute cerebral vascular event, acute coronary syndrome, acute pulmonary oedema, asthmatic major cardiac arrhythmia (as part of primary diagnosis), seizure, drug overdose, injury from burn or trauma ▪ Participants who require immediate intubation and ventilation on arrival in the Emergency Department ▪ Participants undergoing or have undergone cardiopulmonary resuscitation in the pre-hospital phase of their treatment ▪ Current participation in another Clinical Trial of an Investigational Medicinal Product (CTIMP). **Disease severity:** mean Glasgow coma scale score: experimental 14.5, control 14.6	
Interventions	**Experimental:** 15L/min using a non-re-breathe oxygen mask **Control:** target SpO_2 of 94% **Co-intervention:** not reported **Duration:** during emergency department stay	
Outcomes	**Primary outcome:**	

	▪ 90-day mortality **Timing of outcome measurements:** 90 day	
Notes	Email sent to Dr Nutbeam 23 September 2019. Dr Nutbeam replied and forwarded a statistical report on the trial. Email was sent 11 October 2019 asking for additional information about the trial. Reminder sent 18 October 2019. Reply was not received.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated the trial was randomised, but the method of sequence generation was not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	High risk	Open label
Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	High risk	2/25 in the normoxia group were lost to follow-up - reason was not reported.
Selective reporting (reporting bias)	Low risk	The protocol was pre-registered prior to randomisation and all outcomes were reported
Other bias	Unclear risk	It was unclear how the trial was funded
NCT02687217 [27]		
Methods	Randomised clinical trial	
Participants	**Sample size: 60** (experimental 30, control 30) **Sex (% male):** experimental 77%, control 77% **Age (mean):** mean age not reported **Country:** India **Setting:** patients with acute appendicitis who presented to the surgical emergency **Inclusion criteria:** ▪ Clinical diagnosis or Radiological diagnosis of acute appendicitis ▪ Appendectomy through the Mc Burney incision **Exclusion criteria:** ▪ Patients with chronic obstructive pulmonary diseases ▪ Immunodeficiency disease ▪ Patients requiring midline incision ▪ Patients requiring general anaesthesia after failure of spinal anaesthesia ▪ Patients requiring higher oxygen in perioperative period **Disease severity:** not specified	
Interventions	**Experimental:** FiO_2 0.50 throughout the surgery and FiO_2 0.31 via venturi mask 2 hours postoperatively **Control:** no supplemental oxygen throughout the surgery and FiO_2 0.28 via venturi mask 2 hours postoperatively **Co-intervention:** not reported **Duration:** 2 hours	
Outcomes	**Primary outcome:**	

	▪ ASEPSIS Score **Secondary outcomes:** ▪ Number of Patients Requiring Additional Investigations ▪ Number of Patients Requiring Additional Treatment **Timing of outcome measurements:** 14-days No relevant outcomes reported	
Notes	Email sent to Dr Sattavan 11 October 2019. Reminder sent 18 October 2019. Reply was not received.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	Unclear risk	Not described
Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	Unclear risk	Doubt that outcomes are actually reported at all
Selective reporting (reporting bias)	High risk	The protocol was registered retrospectively, and the registered outcomes were not reported
Other bias	Unclear risk	We are unsure about the validity of the trial results. It was unclear how the trial was funded
Padma 2010 [28]		
Methods	Randomised clinical trial	
Participants	**Sample size:** 40 analysed (experimental 20, control 20) **Sex:** not reported **Age:** not reported **Country:** India **Setting:** oxygen therapy in acute Ischaemic stroke **Inclusion criteria:** anterior circulation Ischaemic stroke presenting within 12 h of stroke onset ineligible for thrombolysis, minimum NIHSS score of ≥4 **Exclusion criteria:** active chronic obstructive airway disease, patients requiring >2 L/min of oxygen to maintain peripheral arterial oxygen saturation (SaO_2) > 95%, NIHSS < 4, medically unstable, pregnancy and contraindication to magnetic resonance imaging (MRI)	
Interventions	**Experimental:** humidified oxygen via a simple face mask at flow rates of 10 L/min for 12 hours **Control:** room air or oxygen at 2 L/min via a simple face mask to maintain $SaO_2 \geq 95\%$ **Co-intervention:** not reported **Duration:** 12 hours	
Outcomes	▪ Mortality ▪ The NIHSS, modified Rankin Score (mRS), Barthel Index (BI) were measured at 0, 1, 7 day of admission and at 3 months follow-up MRI with DWI/PWI was performed at admission, 24 h later and at 3 months follow-up	
Notes	Email sent to Dr Padma 15 August 2019. Reminder sent 26 August 2019. Reply was not received	
Risk of bias assessment		

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	High risk	Unblinded
Blinding of outcome assessment (detection bias)	Unclear risk	Neuroradiologist was blinded, however it was not stated whether other outcome assessors were blinded. Unspecified for mortality.
Incomplete outcome data (attrition bias)	High risk	40 completed the trial - how many patients were randomised were not reported
Selective reporting (reporting bias)	Unclear risk	Protocol could not be found
Other bias	Unclear risk	No funding was received. Overall poor reporting quality

Panwar 2016 [29]	
Methods	Randomised clinical trial
Participants	**Sample size**: 104 (experimental 51, control 53 (1 lost to follow-up) **Sex (male)**: experimental 65%, control 62% **Age**: experimental 62 , control 62 **Country**: Australia, New Zealand and France **Setting**: mechanically ventilated adults admitted to a multidisciplinary ICU **Disease severity score:** APACHE III score median 80 (control) and 70 (experimental) **Inclusion criteria:** ▪ People admitted to the ICU ▪ Aged ≥18 years ▪ Receiving invasive MV for < 24 hours and their treating clinician expected MV to continue for at least next 24 hours. The reason for the inclusion criterion of "invasive MV for < 24 hours" was to ensure that participants who would be assigned to the conservative oxygen group did not get exposed to standard liberal oxygen therapy for prolonged periods prior to randomization. **Exclusion criteria:** ▪ Known pregnancy ▪ Imminent risk of death ▪ If the treating clinician lacked equipoise for the participant to be enrolled in this trial The exclusion criterion of "lack of equipoise‟ included those clinical situations where the most appropriate approach (conservative versus liberal) to oxygen therapy is well established. For example, in hypercapnic participants with chronic respiratory failure or exacerbation of chronic obstructive pulmonary disease (COPD), there is level I evidence supporting a conservative approach to oxygen therapy (1) and in participants with carbon monoxide poisoning or necrotizing fasciitis a liberal approach is preferred. However, among participants who had COPD listed as one of the prior co-morbid conditions, the treating clinicians could allow enrolment of those adults who were admitted for reasons unrelated to COPD
Interventions	**Experimental**: SpO_2 target ≥ 96%. Categorized by us as using a high target in the experimental group. **Control**: target SpO_2 of 88% to 92%. When FiO_2 requirement was < 0.50 an SpO_2 of 90% to 92% was recommended, and when FiO_2 requirement was ≥ 0.50 an SpO_2 of 88% to 90% was

	recommended. Categorized by us as using a low target in the control group. **Co-intervention**: participating sites were requested to adhere to best practice guidelines in relation to other potentially confounding co-interventions such as adjustment of tidal volume, PEEP, fluid management, blood transfusion, muscle relaxation, sedation interruption, ventilator weaning, nutrition, use of steroids, early mobilization and physiotherapy. **Duration**: entire duration of mechanical ventilation
Outcomes	**Primary outcomes** ▪ Proportion of time spent in the assigned SpO_2 range in each arm ▪ Area under the curve for PaO_2, FiO_2 and SpO_2 on day 0 to day 7 in each arm **Secondary outcomes** ▪ Incidence of circulation-related events ▪ Incidence of respiration-related events ▪ Incidence of acute kidney injury ▪ Incidence of other organ-systems related outcomes ▪ Time to successful extubation (alive and extubated for >48 hours) ▪ MV free days ▪ ICU mortality ▪ Hospital mortality ▪ All-cause mortality **Timing of outcome measurements:** not specified
Notes	Email sent to Dr Panwar 5 December 2018. Reminder sent 10 December 2018; reply was received.

Risk of bias assessment

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer generated randomization list
Allocation concealment (selection bias)	Low risk	Opaque sealed envelopes
Blinding of participants and personnel (performance bias)	High risk	Participants were unaware of their assigned group but blinding of treating clinicians was not considered feasible
Blinding of outcome assessment (detection bias)	High risk	Not described; however, Dr Panwar clarified in an email that outcome assessment was not blinded. Unspecified for mortality.
Incomplete outcome data (attrition bias)	Low risk	Only 1 (1/104) participant was lost to follow-up
Selective reporting (reporting bias)	Low risk	A study protocol was registered prior to randomization (ACTRN 12613000505707), and all outcomes were reported on
Other bias	Low risk	The trial was supported by public grants. The trial appeared to be free of other components that could put it at risk of bias.

Perrin 2011 [30]

Methods	Randomised clinical trial
Participants	**Sample size:** 106 (experimental 53 (3 withdrawn), control 53) **Sex (male):** experimental 54%, control 34% **Age:** experimental 35, control 33 **Country:** New Zealand **Setting:** patients presenting to the emergency department with asthma **Inclusion criteria:** previous doctor diagnosis of asthma, history consistent with a current acute exacerbation of asthma and a forced expiratory volume in 1 s (FEV1) below/at 50% of predicted

	values at the time of first assessment **Exclusion criteria:** patients with a diagnosis of COPD, or disorders associated with hypercapnic respiratory failure such as neuromuscular disease, chest wall restriction or obesity hypoventilation syndrome, were excluded from the study due to the potential for confounding. Patients who were unconscious, unable to speak or unable to perform spirometry were also excluded
Interventions	**Experimental:** flow rate of 8 l/min via a medium concentration mask (Hudson RCI, Durham, North Carolina, USA) which delivers an FiO_2 of between 0.4 and 0.78 **Control:** received oxygen only if their saturation was at/below 92% on room air, with oxygen titrated as required at 5 min intervals, to achieve an oxygen saturation of 93-95%. Flow rates up to 4 l/min were delivered via nasal cannulae (Hudson RCI) and those >4 l/min were delivered by medium concentration mask **Co-intervention:** all patients received salbutamol 2.5 mg and ipratropium bromide 0.5 mg via an air-driven nebuliser (Portaneb, Respironics, Murrysville, Pennsylvania, USA) on arrival. Patients with severe asthma (FEV1 30-50% predicted) received salbutamol 2.5 mg via a nebuliser every 20 min and prednisone 40 mg orally. Those with very severe asthma (FEV1 <30% predicted) received salbutamol 2.5 mg via a nebuliser every 15 min, hydrocortisone 200 mg intravenously and magnesium sulfate 2 g in 100 ml of normal saline intravenously over 20 min. **Duration:** 1 hour
Outcomes	▪ Measurements of $PtCO_2$ ▪ FEV1 ▪ Respiratory rate and heart rate were made at baseline (0 min) and at 20, 40 and 60 min ▪ The oxygen saturation was measured continuously throughout the study period and recorded at 5 min intervals. **Timing of outcome measurements:** not specified No relevant outcomes reported
Notes	Email sent to Dr Beasley 15 August 2018 and reply was received.

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computerised randomisation sequence
Allocation concealment (selection bias)	Low risk	Computerised database
Blinding of participants and personnel (performance bias)	High risk	Unblinded
Blinding of outcome assessment (detection bias)	High risk	Unblinded
Incomplete outcome data (attrition bias)	Low risk	3/106 (3%) participants were lost to follow-up and reasons are stated in the trial report.
Selective reporting (reporting bias)	Low risk	The trial was registered prior to randomisation and outcomes are reported on
Other bias	Low risk	The trial was funded by public funds

Ranchord 2012 [31]

Methods	Randomised clinical trial
Participants	**Sample size:** 148 (experimental 72 (68 analysed), control 76 (68 analysed) **Sex (men):** experimental 78%, control 71% **Age:** experimental 60, control 62 **Country:** New Zealand and UK **Setting:** patients with ST-elevated myocardial infarction

	Inclusion criteria: subjects 18 years or older who presented to hospital within 12 hours after the onset of ischemic symptoms, with ST-segment elevation of N0.1 mV in 2 contiguous limb leads or 0.2 mV in 2 or more precordial leads or new onset left bundle-branch block, were eligible for enrolment. **Exclusion criteria:** previous myocardial infarction, severe chronic obstructive pulmonary disease (COPD) or type II respiratory failure, cardiogenic shock or oxygen saturation b85% at the time of presentation, pregnancy, previous bleomycin treatment, or participation in another clinical trial. Subjects with cardiac arrest or ventricular fibrillation were not specifically excluded from the study but had to have recovered sufficiently to be able to give written informed consent. Subjects who were subsequently diagnosed to have a condition other than STEMI (e.g., pericarditis), who had an exclusion criterion recognized after randomization, or in whom no formal long consent was documented were withdrawn and not included in the study analysis
Interventions	**Experimental:** 6 L/min of oxygen delivered via a medium concentration mask. If saturations fell to <92%, then higher oxygen concentrations were delivered **Control:** oxygen delivered via nasal prongs or a medium concentration mask: the flow-rate was adjusted to achieve an oxygen saturation of 93% to 96%. If the oxygen saturations were ≥93% while breathing room air in subjects randomised to titrated oxygen, no supplemental oxygen was administered **Co-intervention:** Not stated **Duration:** 6 hours
Outcomes	▪ Infarct size ▪ Mortality ▪ Reinfarction ▪ Target vessel revascularization ▪ MACE (major adverse cardiac events) **Timing of outcome measurements:** 30 days
Notes	Email sent to Dr Beasley 16 August 2019 and reply was received.

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomisation was achieved by way of sealed envelopes in a locked study box, with a 50% chance of either treatment strategy (confirmed by email)
Allocation concealment (selection bias)	Low risk	Sealed envelopes (confirmed by email)
Blinding of participants and personnel (performance bias)	High risk	Unblinded
Blinding of outcome assessment (detection bias)	High risk	Unblinded (confirmed by email). Unspecified for mortality.
Incomplete outcome data (attrition bias)	High risk	12/148 (8%) were withdrawn from analyses, intention to treat data were not reported
Selective reporting (reporting bias)	Low risk	The trial was registered prior to randomisation
Other bias	Low risk	The trial was funded by public funds

Rawles 1976 [32]

Methods	Randomised clinical trial
Participants	**Sample size:** 200 (experimental 105, control 95) **Sex (male):** experimental 86%, control 74% **Age:** experimental 53, control 54 **Country:** UK **Setting:** patients with suspected myocardial infarction admitted to the coronary care unit **Inclusion criteria:** suspected myocardial infarction

	Exclusion criteria: clinical evidence of right or left heart failure, chronic bronchitis or emphysema or breathlessness from any other cause, or if they had been transferred from other wards or the treatment of arrhythmias or had undergone a cardiac arrest before admission or suffered from cardiogenic shock
Interventions	**Experimental:** 6 L/min oxygen with MC mask **Control:** compressed air with MC mask **Co-intervention:** not described **Duration:** 24 hours
Outcomes	▪ Severity of infarction ▪ Incidence of arrhythmias ▪ Use of analgesics ▪ Mortality **Timing of outcome measurement:** during hospital stay (mean 13 days)
Notes	Contact information was not identified; thus email was not sent.

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described
Allocation concealment (selection bias)	Low risk	Numbered sealed envelopes
Blinding of participants and personnel (performance bias)	Low risk	"The cylinders were shrouded so that the identity of the gas was not apparent to the medical staff or patients. If during the first 24 hours it became mandatory to give oxygen because of cardiac arrest or developing left ventricular failure, then the mask was disconnected from the cylinder and connected to the piped wall supply of oxygen without disclosing to the medical staff the identity of the former gas".
Blinding of outcome assessment (detection bias)	Low risk	Blinded
Incomplete outcome data (attrition bias)	High risk	43/200 (21,5%) was excluded from analysis
Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	Unclear risk	It was unclear how the trial was funded

Rodrigo 2003 [33]

Methods	Randomised clinical trial
Participants	**Sample size:** 77 (experimental 39 (1 withdrawn from analysis), control 28 (2 withdrawn from analysis)) **Sex (male):** experimental 63%, control 67% **Age:** experimental 36, control 38 **Country:** Uruguay **Setting:** patients with acute asthma admitted to the emergency department **Inclusion criteria:** diagnosis criteria of asthma of the American Thoracic Society; age from 18 to 50 years; peak expiratory flow rate < 60% of predicted value **Exclusion criteria:** temperature > 38C, or a history of cardiac, hepatic, renal disease, or other medical disease, or pregnancy; and an expressed willingness to participate in the study, with written informed consent obtained
Interventions	**Experimental:** 100% oxygen via a standard nonrebreathing facemask **Control:** 28% oxygen via a standard face mask **Co-intervention:** at the end of oxygen protocol, all patients received albuterol and ipratropium bromide (120 µg of albuterol sulfate and 21 µg of ipratropium bromide per actuation) delivered by a Inetered-dose inhaler into a spacer device in a dose of four puffs at 10 minutes intervals in

	accordance with previous evidence. Additionally, patients with a poor response received hydrocortisone 400 mg IV **Duration:** 20 minutes
Outcomes	▪ Heart and respiratory rates ▪ Pulmonary function ▪ Arterial blood gas levels after 20 minutes of oxygen administration **Timing of outcome measurement:** 20 minutes post intervention No relevant outcomes reported
Notes	Email sent to Dr Rodrigo 16 August 2019. Reminder sent 26 August 2019. Reply was not received

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated random numbers
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	High risk	Unblinded, face masks used are obvious different (reservoir bag is only present in the non-rebreathing face mask)
Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	Low risk	Less than 5% were lost to follow-up
Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	Unclear risk	It was unclear how the trial was funded

Rodrigues de Freitas Vianna 2017 [34]

Methods	Randomised clinical trial
Participants	**Sample size:** 76 (68 analysed) **Sex (male):** 60 (only overall reported) **Age:** 68 (only overall reported) **Country:** Brazil **Setting:** mechanically ventilated intensive care unit patients **Inclusion criteria:** patients undergoing endotracheal intubation and on mechanical ventilation for 12 h, hemodynamically stable, sedated or not, and requiring endotracheal suctioning according to American Association for Respiratory Care criteria **Exclusion criteria:** individuals using high doses of vasopressor amines and/or having severe cardiac arrhythmias; with hemoglobin 7 g/dL, impossibility of appropriate monitoring of SpO_2 , baseline FIO_2 0.60, requirement of PEEP of 10 cm H_2O, rib fractures, presence of a chest drain, severe bronchospasm, intracranial hypertension hypertension (intracranial pressure 10 mm Hg), hemorrhagic disorders, marked degree of gastroesophageal reflux, bullous lung disease, unilateral lung disease, use of a tracheostomy, closed suction system, peak pressure 35 cm H_2O, hemodynamic instability with mean arterial pressure 60 mm Hg, central venous pressure (CVP) 6 mm Hg, and no criteria indicating the need for endotracheal suctioning.
Interventions	Open endotracheal suctioning was performed using 2 different intervention sequences: **Experimental:** hyperoxygenation FIO_2 1.0 **Control:** hyperoxygenation of 0.20 above baseline (FIO_2 + 0.20) **Co-intervention:** not described **Duration:** during endotracheal suctioning procedure
Outcomes	▪ Oxygen (SpO_2) and ventilation ($ETCO_2$) measures ▪ Respiratory mechanic measures

	▪ Volumetric capnography measures **Timing of outcome measurement:** 30 minutes post intervention No relevant outcomes reported
Notes	Email sent to Dr Vianna 16 August 2019. Reminder sent 26 August 2019; however, no reply was received

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomisation was performed by drawing lots
Allocation concealment (selection bias)	Low risk	Opaque envelopes
Blinding of participants and personnel (performance bias)	High risk	Single-blinded
Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	High risk	8 patients died (and excluded from analysis); however, allocation group was not reported
Selective reporting (reporting bias)	High risk	The protocol was registered retrospectively
Other bias	Unclear risk	It was unclear how the trial was funded

Roffe 2010 [35]

Methods	Randomised clinical trial
Participants	**Sample size:** experimental 30 (1 excluded from analyses), control 33 (3 excluded from analyses) **Sex (male):** experimental 52%, control 67% **Age (mean years):** experimental 75, control 73 **Country:** UK **Setting:** patients with acute stroke **Inclusion criteria:** adult patients with a clinical diagnosis of acute stroke who were not moribund were recruited within 72 hours of admission **Exclusion criteria:** patients with definite indications for oxygen supplementation (SpO_2 below 90%, decompensated congestive cardiac failure, pneumonia with consolidation on the chest radiograph, known chronic hypoxia requiring long-term oxygen treatment; severe persistent disability from a prior stroke, confusion and restlessness making probe placement difficult, reduced peripheral perfusion leading to an unobtainable or poor oximetry trace, pregnancy, and refusal of consent
Interventions	**Experimental:** 2 L/min oxygen supplementation via nasal cannulae overnight (21:00-9:00) **Control:** room air **Co-intervention:** additional oxygen was given at the discretion of the clinical team, if medically indicated **Duration:** 12 hours
Outcomes	▪ Time spent with an SpO_2 below 90% during the night (corrected for an 8-hour recording) ▪ The lowest SpO_2 recorded during the night ▪ Feasibility (the proportion of patients prescribed oxygen who actually had oxygen in place when checked) ▪ Tolerability (sleep disturbance) ▪ Mortality **Timing of outcome measurements:** mortality at 14 days
Notes	Email sent to Dr Roffe 16 August 2019 and reply was received

Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated random numbers
Allocation concealment (selection bias)	High risk	Not concealed (confirmed by email)
Blinding of participants and personnel (performance bias)	High risk	Single-blind
Blinding of outcome assessment (detection bias)	High risk	Unblinded (confirmed by email) Unspecified for mortality.
Incomplete outcome data (attrition bias)	High risk	More than 5% were lost to follow-up, although reasons were justified
Selective reporting (reporting bias)	High risk	Protocol was not published or registered (confirmed by email)
Other bias	Low risk	The trial was funded by North Staffordshire Medical Institute (confirmed by email)

Roffe 2017A [36]	
Methods	Randomised clinical trial
Participants	**Sample size:** randomised: 2668+(2668/2) = 4002, analysed at 90 day follow-up: 2567+1275=3842 **Sex (men):** experimental 55%, control 55% **Age (mean years):** experimental 72, control 72 **Country:** UK **Setting:** patients with acute stroke **Inclusion criteria:** adults (aged ≥18 years) with a clinical diagnosis of acute stroke within 24 hours of hospital admission, who had no clinical indications for or contraindications to oxygen treatment or any concomitant condition likely to limit life expectancy to less than 12 months **Exclusion criteria:** if the responsible doctor considered the patient to have definite indications for or contraindications to oxygen treatment at a rate of 2-3 L/min. Potential indications for oxygen treatment were: oxygen saturation on air <90%, hypoxia associated with acute left ventricular failure, severe pneumonia, pulmonary embolus, and chronic respiratory failure patients treated with long term oxygen at home. Potential contraindications to fixed dose oxygen treatment were type 2 respiratory failure and very severe hypoxia. Patients were also excluded if the stroke was not the main clinical problem, or if he/she had another serious life-threatening illness likely to lead to death within the next few months.
Interventions	**Experimental:** continuous oxygen for 72 hours **Control:** oxygen only if clinically indicated Oxygen was given via nasal tubes at 3 L/min if baseline oxygen saturation was 93%or less and at 2 L/min if oxygen saturation was greater than 93% **Co-intervention:** not specified **Duration:** 72 hours
Outcomes	**Primary Outcome:** ▪ Modified Rankin Score at 3 months **Secondary outcomes at one week:** ▪ No of patients with neurological improvement (≥4-point decrease in the NIHSS) ▪ Any deaths ▪ Highest oxygen saturation during the first 72 hours ▪ Lowest oxygen saturation during the first 72 hours

	Secondary outcomes at 3 months: ▪ Mortality ▪ The percentage of patients living at home ▪ Ability to perform activities of daily living (Barthel index) ▪ Quality of life (EuroQuol and VAS) Extended activities of daily living (Nottingham EADL)	
Notes	The trial was designed as a multi-arm trial: continuous oxygen for 72 hours, nocturnal oxygen for three nights and oxygen only if clinically indicated. Quality of life was measured with both EQ5D-3L and EQ-VAS. We used the results from the EQ-VAS as the other trials reported quality-of-life using this scale. We used the RevMan calculator to calculate SD's based on the data reported in the Health Technology Assessment. Email sent to Dr Roffe 16 August 2019 and reply was received	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated
Allocation concealment (selection bias)	Low risk	Centralised web-based
Blinding of participants and personnel (performance bias)	High risk	Unblinded
Blinding of outcome assessment (detection bias)	Low risk	Ninety-day assessments were undertaken by the SO_2S study office, which was blind to treatment allocation Unspecified for mortality.
Incomplete outcome data (attrition bias)	Low risk	Less than 5% were lost to follow-up at 90-day assessment
Selective reporting (reporting bias)	Low risk	The protocol was registered prior to randomisation
Other bias	Low risk	The trial was supported by public grants
Roffe 2017B [36]		
Methods	Randomised clinical trial	
Participants	**Sample size:** randomised: 2667+2668/2= 4001, analysed: 2561+1274=3835 **Sex (men):** experimental 55%, control 55% **Age:** experimental 72, control 72 **Country:** UK **Setting:** patients with acute stroke **Inclusion criteria:** adults (aged ≥18 years) with a clinical diagnosis of acute stroke within 24 hours of hospital admission, who had no clinical indications for or contraindications to oxygen treatment or any concomitant condition likely to limit life expectancy to less than 12 months **Exclusion criteria:** if the responsible doctor considered the patient to have definite indications for or contraindications to oxygen treatment at a rate of 2-3 L/min. Potential indications for oxygen treatment were: oxygen saturation on air <90%, hypoxia associated with acute left ventricular failure, severe pneumonia, pulmonary embolus, and chronic respiratory failure patients treated with long term oxygen at home. Potential contraindications to fixed dose oxygen treatment were type 2 respiratory failure and very severe hypoxia. Patients were also excluded if the stroke was not the main clinical problem, or if he/she had another serious life-threatening illness likely to lead to death within the next few months.	
Interventions	**Experimental:** nocturnal oxygen (21:00 to 07:00 hours) for 3 nights **Control:** oxygen only if clinically indicated Oxygen was given via nasal tubes at 3 L/min if baseline oxygen saturation was 93%or less and at 2 L/min if oxygen saturation was greater than 93%	

	Co-intervention: NS **Duration:** 3 nights (10hours x 3)
Outcomes	**Primary Outcome:** ▪ Modified Rankin Score at 3 months **Secondary outcomes at one week:** ▪ No of patients with neurological improvement (≥4 point decrease in the NIHSS) ▪ Any deaths ▪ Highest oxygen saturation during the first 72 hours ▪ Lowest oxygen saturation during the first 72 hours **Secondary outcomes at 3 months:** ▪ Mortality ▪ The percentage of patients living at home ▪ Ability to perform activities of daily living (Barthel index) ▪ Quality of life (EuroQuol and VAS) ▪ Extended activities of daily living (Nottingham EADL)
Notes	The trial was designed as a multi-arm trial: continuous oxygen for 72 hours, nocturnal oxygen for three nights and oxygen only if clinically indicated. Quality of life was measured with both EQ5D-3L and EQ-VAS. We used the results from the EQ-VAS as the other trials reported quality-of-life using this scale. We used the RevMan calculator to calculate SD's based on the data reported in the Health Technology Assessment. It seems like the test results from the comparison continuous versus nocturnal had been adjusted, thus we used the SD calculated from the comparison continuous versus control (MD 0.10 (-1.93 to 2.12) P=0.90). Email sent to Dr Roffe 16 August 2019 and reply was received

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated
Allocation concealment (selection bias)	Low risk	Centralised web-based
Blinding of participants and personnel (performance bias)	High risk	Unblinded
Blinding of outcome assessment (detection bias)	Low risk	Ninety-day assessments were undertaken by the SO_2S study office, which was blind to treatment allocation. Unspecified for mortality
Incomplete outcome data (attrition bias)	Low risk	Less than 5% were lost to follow-up at 90-day assessment
Selective reporting (reporting bias)	Low risk	The protocol was registered prior to randomisation
Other bias	Low risk	The trial was supported by public grants

Sepehrvand 2019 [37]	
Methods	Randomised clinical trial
Participants	**Sample size:** 50 (experimental 25, control 25) **Sex (% male):** experimental 56%, control 60% **Age (mean):** experimental 73, control 74 **Country:** Canada

	Setting: patients presented at the emergency department with acute heart failure **Inclusion criteria:** patients >40 years of age presenting to the emergency department with objective acute heart failure (brain natriuretic peptide > 400 picogram/mL and/or chest X-ray with pulmonary congestion) and with a planned admission for the treatment of heart failure as the primary diagnosis. Patients were eligible for randomization within 16 hours of presenting to the emergency department. **Exclusion criteria:** Patients on home O_2, known prior hypercapnic failure ($PaCO_2$ > 50 mmHg), asthma, primary pulmonary hypertension, requiring urgent positive pressure ventilation or intubation, or on >10 L/min O_2 were excluded **Disease severity:** not reported	
Interventions	**Experimental:** SpO_2 ≥ 96% **Control:** SpO_2 90-92% **Co-intervention:** **Duration:** 72 hours	
Outcomes	**Primary outcome:** ▪ change in N-terminal pro-brain-type natriuretic peptide (NT-proBNP) from baseline to 72 hours **Secondary outcomes:** ▪ change in dyspnoea on visual analogue scale from baseline to 72 hours ▪ change in global symptoms using patient global assessment measure to 72 hours ▪ change in peak expiratory flow at 72 h ▪ worsening of heart failure at 7 days ▪ diuretic response at 72 hours ▪ clinical event at 30 days following hospital discharge (all-cause mortality and HF readmission) **Timing of outcome measurements:** not specified	
Notes	Email sent to Dr Ezekowitz 11 October 2019. Reminder sent 18 October 2019. Reply was not received.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer-generated random numbers
Allocation concealment (selection bias)	Low risk	Computerised response system
Blinding of participants and personnel (performance bias)	High risk	Open label
Blinding of outcome assessment (detection bias)	Low risk	Subjective outcomes were assessed by a research coordinator who was blinded to the patient's group allocation. Unspecified for mortality.
Incomplete outcome data (attrition bias)	Low risk	No patients were lost to follow-up
Selective reporting (reporting bias)	High risk	Retrospectively registered. Primary outcome changed.
Other bias	Low risk	The trial was supported by public funds
Shi 2017 [38]		
Methods	Randomised clinical trial	
Participants	**Sample size:** 18 (experimental 9, control 9) **Sex (% male):** experimental 67%, control 71%	

	Age (mean): experimental 62, control 58 **Country:** China **Setting:** patients with acute ischemic stroke **Inclusion criteria:** age >18, presenting <4.5 hours after witnessed symptom onset; eligible for intravenous thrombolysis; National Institutes of Health Stroke Scale (NIHSS) score<25; pre-admission modified Rankin Scale score <1; and acute ischemic stroke confirmed by computed tomography or magnetic resonance imaging the following day **Exclusion criteria:** active chronic obstructive pulmonary disease, >3 L/min oxygen required to maintain peripheral arterial oxygen saturation >95% per stroke management guidelines; rapidly improving neurological deficits; medically unstable; pregnancy; inability to obtain informed consent **Disease severity:** experimental admission NIHSS 12, control admission NIHSS 12.3
Interventions	**Experimental:** 10 L/min by oxygen facemask **Control:** room air **Co-intervention:** all enrolled patients with acute ischemic stroke received intravenous tPA thrombolytic therapy and standard clinical treatment (anticoagulant and antiplatelet) **Duration:** 4 hours
Outcomes	**Primary outcome:** ▪ Scores assessed by National Institutes of Health Stroke Scale (NIHSS) **Secondary outcomes:** ▪ Number of participants with adverse events that are related to treatment **Timing of outcome measurements:** 7 days
Notes	Email sent to Dr Liu and Dr Ji 11 October 2019. Reminder sent 18 October 2019. Reply was not received.

Risk of bias assessment

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	High risk	Un-blinded. Open label according to registration on clinicaltrials.gov
Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	Unclear risk	Number of analysed patients not adequately described. We used no lost to follow-up
Selective reporting (reporting bias)	Low risk	The protocol was pre-registered, and outcomes were reported
Other bias	High risk	The trial was supported by public grants No explanation of why the trial stopped enrolment of patients at n=18 (estimated enrolment in protocol: n=40)

Sills 2003 [39]

Methods	Randomised clinical trial
Participants	**Sample size:** 25 (15 analysed: experimental 8, control 7) **Sex:** not reported **Age:** not reported **Country:** UK

	Setting: patients with acute stroke **Inclusion criteria:** not reported **Exclusion criteria:** not reported	
Interventions	**Experimental:** 2L/min oxygen via nasal cannula overnight **Control:** no routine oxygen **Co-intervention:** not reported **Duration:** 8 hours (23:00-7:00)	
Outcomes	▪ Oxygen saturation **Timing of outcome measurements:** not specified No relevant outcomes reported	
Notes	Conference abstract Email sent to Dr Roffe 16 August 2019 and reply was received	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Random numbers list (confirmed by email)
Allocation concealment (selection bias)	High risk	Not concealed (confirmed by email)
Blinding of participants and personnel (performance bias)	High risk	Unblinded (confirmed by email)
Blinding of outcome assessment (detection bias)	High risk	Unblinded (confirmed by email)
Incomplete outcome data (attrition bias)	High risk	40% were lost to follow-up
Selective reporting (reporting bias)	High risk	No protocol was pre-published or registered (confirmed by email)
Other bias	Low risk	The trial was funded by North Staffordshire Medical Institute
Singhal 2005 [40]		
Methods	Randomised clinical trial	
Participants	**Sample size:** 16 (experimental 9, control 7) **Sex (male):** experimental 44%, control 43% **Age:** experimental 67, control 70 **Country:** US **Setting:** acute stroke **Inclusion criteria:** nonlacunar, anterior circulation ischemic stroke presenting 12 hours after witnessed symptom onset or 15 hours after last seen neurologically intact; (2) ineligible for intravenous/intra-arterial thrombolysis; (3) National Institutes of Health Stroke Scale (NIHSS) score 4; (4) pre-admission modified Rankin scale (mRS) score 1, and (5) mean transit time (MTT) lesion larger than DWI lesion (perfusion– diffusion "mismatch") with evidence for cortical hypoperfusion on MRI **Exclusion criteria:** (1) active chronic obstructive pulmonary disease; (2) 3 L/min oxygen required to maintain peripheral arterial oxygen saturation (SaO_2) 95% as per current stroke management guidelines;18 (3) rapidly improving neurological deficits; (4) medically unstable; (5) pregnancy; (6) inability to obtain informed consent; and (7) contraindication for MRI	
Interventions	**Experimental:** humidified oxygen via simple facemask at flow rates of 45 L/min **Control:** room air or nasal oxygen 1 to 3 L/min if necessary to maintain SaO_2 95% **Co-intervention:** not described **Duration:** 8 hours	
Outcomes	▪ Comparison of DWI lesion growth at 4 hours between groups	

	▪ Mean NIHSS scores and perfusion parameters at 4 hours ▪ Percentage of ADC voxels undergoing reversal at 4 hours or 24 hours ▪ Brain hemorrhage at 24 hours ▪ 3-month stroke lesion volumes ▪ NIHSS and mRS scores ▪ Mortality	
Notes	Authors initially planned to enroll 40 patients in this pilot study to allow formal power calculations. The presented results are interim analysis. Email was sent to Dr Singhal 16 August 2019. Reminder sent 26 August 2019. Reply received 28 August; however, no clarifications were achieved	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described
Allocation concealment (selection bias)	Low risk	Sealed envelopes
Blinding of participants and personnel (performance bias)	High risk	Unblinded
Blinding of outcome assessment (detection bias)	Low risk	The neuroradiologists were blinded to allocation group; however, the clinical investigator was unblinded. Unspecified for mortality.
Incomplete outcome data (attrition bias)	Low risk	No patients were lost to follow-up
Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	Low risk	The trial was supported by public funds
Singhal 2013 [41]		
Methods	Randomised clinical trial	
Participants	**Sample size:** 85 (experimental 43, control 42) **Sex:** experimental 37%, control 60% **Age:** experimental 74, control 73 **Country:** US **Setting:** acute ischemic stroke **Inclusion criteria:** age greater than or equal to 18 years; acute ischemic stroke in whom treatment can potentially be started within 9 hours after symptom onset. If the symptom onset time is unknown, the time of onset will be defined as the midpoint between the time when the subject was last seen neurologically intact, and when found to have a neurological deficit; national Institutes of Health Stroke Scale (NIHSS) score 4 or greater. **Exclusion criteria:** patients being actively considered for intravenous or intra-arterial thrombolysis will be excluded; patients likely to have acute stroke intervention such as carotid endarterectomy or stent or angioplasty, hemicraniectomy, etc.; rapidly improving neurological deficits (transient ischemic attack); known history of severe chronic obstructive pulmonary disease (Forced Expiratory Vital Capacity less than 1.0 or oxygen dependent); more than 3 L/min oxygen required to maintain peripheral arterial oxygen saturation above 92%; new York Heart Association Class III heart failure; endotracheal intubation prior to enrolment or impending need for artificial ventilation; coma (National Institutes of Health Stroke Scale item 1a score of 3); suspected seizure at or after onset of stroke, or a known active seizure disorder; blood glucose below 50 mg/dL or greater than 250 mg/dL prior to enrolment; concurrent severe non-stroke medical illness requiring admission to a non-neurological intensive care unit; expected survival less than 90 days; any condition that might limit neurological assessment or follow-up in the opinion of the investigator; pre-menopausal women with a positive pregnancy blood test performed at	

	admission; inability to obtain consent from the patient or legally authorized representative; active participation in another intervention study (e.g. investigational drug trial); proven alternate etiology for stroke-like symptoms
Interventions	**Experimental:** oxygen 30-45L/min via a facemask **Control:** room air, inhaled at 30-45L/min via a facemask for 8 hours **Co-intervention:** not described **Duration:** 8 hours
Outcomes	▪ Change in NIHSS ▪ MRI lesion growth ▪ Tissue reperfusion ▪ % mismatch lost SAE, brain edema and brain haemorrhage: no difference according to abstract; however, data were not reported. **Timing of outcome measurements:** not specified
Notes	The trial was prematurely terminated due to imbalance in deaths. Two abstracts identified, but no full trial report was identified. Results on serious adverse events were extracted from clinicaltrials.gov. Email was sent to Dr Singhal 16 August 2019. Reminder sent 26 August 2019. Reply received 28 Aug; however, no clarifications were achieved

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the methods of sequence generation was not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	Unclear risk	The trial was described as double blind, but it was unclear who was blinded and how blinding was maintained
Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	Low risk	One patient was withdrawn
Selective reporting (reporting bias)	Low risk	According to trial registration on clinialtrials.gov (NCT00414726), the trial was registered prospectively
Other bias	High risk	The trial was funded by public funds. Early stopping: the trial was stopped due to an imbalance in deaths, which was not pre-defined

Stewart 2019 [42]

Methods	Cluster randomised crossover clinical trial
Participants	**Sample size:** target sample size 21000 **Sex (% male):** not reported **Age (mean):** not reported **Country:** New Zealand **Setting:** patients attended by the ambulance service managed on the acute coronry syndrome pathway or with an acute coronary syndrome **Inclusion criteria:** adults of both gender of 18 years or older in New Zealand admitted to the coronary care unit and/or cardiac catheter laboratory with an acute coronary syndrome at participating hospitals; patients attended by the ambulance service with a confirmed acute coronary syndrome who die before admission to coronary care unit or catheter lab. **Exclusion criteria:** dead on ambulance arrival at the scene; presented with an out of hospital cardiac arrest; ventilated prior to admission to CCU/catheter lab; documented for end of life

	cares; on home oxygen; not admitted to coronary care unit or catheter lab because of advanced age, co-morbidity, or because a diagnosis other than acute coronary syndrome is made **Disease severity:** not reported
Interventions	**Experimental:** In the ambulance oxygen was administered by face mask at ~8L/minute. If a face mask was not tolerated, then oxygen was given by nasal prongs at ~4 l/minute. Oxygen flow rate was increased if necessary, to achieve saturation greater than or equal to 95%. Oxygen was continued until the patient was admitted to hospital or when a doctor decided that it was no longer necessary. Target oxygen saturation in hospital was 95% to 99%. **Control:** Oxygen was not administered unless the measured oxygen saturation was less than 90%. If oxygen was given, then target oxygen saturation was 90 to 94%. **Co-intervention:** Not reported **Duration:** The treating clinician decided on oxygen flow rate, method of administration, and when to discontinue oxygen when symptoms and signs of ischemia resolved, or when clinically appropriate
Outcomes	**Primary outcome:** ▪ Mortality rate, 30 days following episode of acute coronary syndrome **Secondary outcomes:** ▪ Mortality rate, 1 year following episode of acute coronary syndrome ▪ Hospital readmission for cardiovascular cause, 1 year following episode of acute coronary syndrome **Timing of outcome measurements:** see specific outcomes No relevant outcomes reported
Notes	The publication (conference abstract) reported on the design and conduct of the trial. No results were reported. Email sent to Dr Stewart 11 October 2019. Reply was received.

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described
Allocation concealment (selection bias)	High risk	Not concealed
Blinding of participants and personnel (performance bias)	High risk	Open label
Blinding of outcome assessment (detection bias)	High risk	The outcomes were obtained from ICD-10 codes for all patients in the register for the defined study periods. Patients were blinded in the sense that it was all done by computer with no 'assessor' involvement, and all patients in the registers were included (clarified by e-mail). The person who assessed whether a patient had an outcome was not blinded. Unspecified for mortality.
Incomplete outcome data (attrition bias)	Unclear risk	Can't be assessed as the results have not been published
Selective reporting (reporting bias)	Unclear risk	Can't be assessed as the results have not been published
Other bias	Unclear risk	The trial is funded by public funds. Other bias can't be assessed as the results have not been published

Stub 2015 [43]

Methods	Randomised clinical trial

Participants	**Sample size:** 638 randomised (experimental 318, control 320). Analysed: experimental 218, control 223 **Sex (men):** experimental 80%, control 78% **Age (mean):** experimental 63, control 63 **Country:** Australia **Setting:** Patients with acute ST-elevation myocardial infarction randomised in ambulance **Inclusion criteria:** adults ≥18 years of age, who describe chest pain commencing less than 12 hours prior to assessment, ST elevation Myocardial Infarction on pre-hospital 12-lead ECG (characterized by, ST-segment elevation of ≥1 mm in two contiguous limb lead or ST-segment elevation of ≥2 mm in two contiguous chest leads or new left bundle branch block pattern) **Exclusion criteria:** hypoxic with oxygen saturation measured on pulse oximeter below 94% with the patient breathing air, have bronchospasm on examination requiring nebulised salbutamol therapy using oxygen, receive oxygen prior to randomisation, or have any condition associated with altered conscious state
Interventions	**Experimental:** patients in the oxygen group were administered supplemental oxygen via face mask at 8 L/min by paramedics. This therapy continued until transfer from the cardiac catheterization laboratory to the cardiac care ward **Control:** patients randomised to the no oxygen arm received no oxygen unless oxygen saturation fell below 94%, in which case oxygen was administered via nasal cannula (4 L/min) or face mask (8 L/min) to achieve an oxygen saturation of 94% **Co-intervention:** all patients received aspirin 300 mg orally by paramedics. Additional antiplatelet therapy and choice of anticoagulation and percutaneous intervention strategy were at the discretion of the treating interventional cardiologist, according to hospital protocol **Duration:** therapy continued until transfer from the cardiac catheterization laboratory to the cardiac (mean experimental 79 minutes, control 52 minutes)
Outcomes	▪ Myocardial Infarct Size at 72 hours post infarct ▪ ST segment resolution at 1-day post reperfusion ▪ TIMI (Thrombolysis in Myocardial infarction score) flow at completion of coronary intervention procedure ▪ Survival to Hospital Discharge ▪ Major Adverse Cardiac Events (MACE: Death, recurrent myocardial infarction, and re-hospitalisation measured at 6 months) Myocardial Salvage at 4 days and 6 months magnetic resonance imaging (MRI) measurement of infarct size as percent of area at risk determined with T2-weighted MRI (in small sub set of patients) at day 4 and repeated at 6 months
Notes	Email sent to Dr Stub 16 August 2019. Reminder sent 26 August 2019. No reply was received

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer generated random numbers
Allocation concealment (selection bias)	Low risk	Opaque envelopes
Blinding of participants and personnel (performance bias)	High risk	Open label
Blinding of outcome assessment (detection bias)	Low risk	All primary efficacy and safety outcome measures, including mortality, cardiac arrest, and unplanned intubations, were assessed by an independent Data Safety Monitoring Committee. Unspecified for mortality.
Incomplete outcome data (attrition bias)	High risk	Intention to treat analysis not reported
Selective reporting (reporting bias)	Low risk	The protocol was published prior to randomisation and all outcomes were reported on
Other bias	Low risk	The trial was funded by public grants

Taher 2016 [44]		
Methods	Randomised clinical trial	
Participants	**Sample size**: 68 (experimental 34, control 34) **Sex (men)**: experimental 74%, control 68% **Age**: experimental 40 , control 46 **Country**: Iran **Setting**: adults with traumatic brain injury initially referred to the emergency department, but who were admitted to the ICU **Disease severity score:** GCS score mean 7.4 **Inclusion criteria:** ▪ Age between 18 and 65 years ▪ Less than 6 hours passed since the accident; haemodynamic stability; and GCS between 3 and 8 **Exclusion criteria:** ▪ Pregnancy ▪ People under 18 or older than 65 years ▪ GCS under 3 or more than 8 ▪ People with chronic disease such as diabetes mellitus, ischaemic heart disease, renal failure, acute pulmonary oedema, history of massive myocardial infarction, and heart failure ▪ People with a baseline blood pressure of less than 90/60 ▪ People with successful cardiopulmonary resuscitation (CPR) ▪ Death or loss to follow-up Participants in the control group in which oxygen therapy was inevitable were also excluded from this study	
Interventions	**Experimental**: FiO_2 of 80% oxygen by mechanical ventilator in the first 6 hours after the traumatic accident. **Control**: FiO_2 of 0.5 using mechanical ventilator in the first 6 hours after the traumatic accident. Categorized by us as using a low target in the control group. **Co-intervention**: not specified **Duration**: 6 hours	
Outcomes	▪ Gloscow coma scale ▪ Barthel Index ▪ mRS neurologic disability scoring system at the time of discharge from hospital and at 6-month follow-up **Timing of outcome measurements:** not specified No relevant outcomes reported	
Notes	Participants who had died were excluded (from analyses). Email sent 6 December 2018 to Dr Pilehvari. Reminder sent 16 August 2019. No reply was received	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomized, but sequence generation was not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	Unclear risk	The trial was described as double-blind; however, it was unclear who was blinded and how blinding was maintained

Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	High risk	Participants who died or were lost to follow-up were excluded.
Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	Low risk	The trial was funded by public funds. The trial appeared to be free of other components that could put it at risk of bias.
Thomas 2019 [45]		
Methods	Cluster randomised clinical trial	
Participants	**Sample size:** 35 (experimental 17, control 18) **Sex (% male):** experimental 71%, control 72% **Age (mean):** experimental 64, control 70 **Country:** UK **Setting:** patients with a sustained return of spontaneous circulation following out of hospital cardiac arrest **Inclusion criteria:** patients were enrolled if they were 18 years or older and had an out of hospital cardiac arrest (OHCA) that was not caused by trauma (trauma included drowning and hanging) **Exclusion criteria:** less than 18 years old; cardiac arrest believed to have been caused by trauma (including hanging and drowning); entered into the study previously; detained by Her Majesty's Prison Service **Disease severity:** not reported	
Interventions	**Experimental:** FiO_2 1.00 **Control:** Target saturation 94-98%. Paramedics were advised to consider titration of oxygen every 2 min. If there was no reliable saturation trace, or oxygen saturations fell below 94% study paramedics increased oxygen delivery in a stepwise approach **Co-intervention:** "all other elements of routine care were provided as usual" **Duration:** 1 hour	
Outcomes	**Primary outcome:** ▪ Recruitment rate measured as the proportion of eligible paramedics attending training and consenting to take part at the end of the recruitment period **Secondary outcomes:** ▪ Proportion of surviving participants providing quality of life data at discharge and 90 days measured by the number of patients completing the modified Rankin Scale, EuroQol EQ-5D-5L and the SF-36 instruments at discharge and 90 days following OHCA ▪ Survival to discharge and 90-days following OHCA **Timing of outcome measurements:** discharge and 90-days	
Notes	Email sent to Dr Thomas 11 October 2019. Reminder sent 18 October 2019. Reply was not received. Unit of randomisation were the paramedics. 46 paramedics were randomised, and 35 patients received the intervention. To avoid unit of analysis issues, we conducted the analyses at the same level as the allocation equivalent to each cluster being analysed as a single individual (sample size for the analysis = number of clusters). As the number of clusters overweighs the number of included patients, then sample size = included patients.	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**

Random sequence generation (selection bias)	Low risk	Computer-generated random numbers, paramedics were the unit of randomisation
Allocation concealment (selection bias)	High risk	Cluster-randomised
Blinding of participants and personnel (performance bias)	High risk	Un-blinded
Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	Low risk	No patients were lost to follow-up
Selective reporting (reporting bias)	High risk	The trial was retrospectively registered
Other bias	Low risk	The trial was supported by public funds

Ukholkina 2005 [46]	
Methods	Randomised clinical trial
Participants	**Sample size:** 65 (number of randomised to each group were not reported) **Sex:** not reported **Age:** not reported **Country:** Russia **Setting:** patients with uncomplicated myocardial infarction **Inclusion criteria:** not described **Exclusion criteria:** not described
Interventions	**Experimental:** 30 patients, 40-60% oxygen inhalation was performed for 30 min before reperfusion and 3 h after **Control:** not described **Co-intervention:** included aspirin, P-blockers, ACE inhibitors and nitrates **Duration:** 3,5 hours (30 minutes before reperfusion and 3 hours after it)
Outcomes	▪ Complications (rhythm and conduction disturbances - postinfarction angina, pericarditis, circulatory insufficiency) ▪ CKK activity ▪ Left ventricular function **Timing of outcome assessment:** not described No relevant outcomes were reported
Notes	Contact details were not identified; thus, email was not sent.

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but sequence generation was not described
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	High risk	Unblinded
Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	Unclear risk	Not reported. Stated that 65 patients were included, and 30 patients received the experimental intervention. No results were reported for the two groups, only whether authors found difference between groups.

Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	Unclear risk	It was unclear how the trial was funded
Wijesinghe 2012 [47]		
Methods	Randomised clinical trial	
Participants	**Sample size:** 150 (experimental 75, control 75) **Sex (% male):** experimental 45%, control 37% **Age (mean years):** experimental 45, control 46 **Country:** New Zealand **Setting:** Patients presenting to the emergency department with suspected community-acquired pneumonia **Inclusion criteria:** presence of cough, a respiratory rate >18 breaths per minute, and at least one systemic feature of sweating, rigors or fever >37.8°C, representing the criteria for the clinical diagnosis of pneumonia. **Exclusion criteria:** ▪ Patients with a diagnosis of COPD ▪ Patients with disorders associated with hypercapnic respiratory failure such as neuromuscular disease, or obesity hypoventilation syndrome ▪ Patients presenting with respiratory failure requiring mechanical ventilation, acute ECG changes suggesting ischaemia or suspected ▪ neutropenic sepsis **Disease severity:** CRB-65 score (pneumonia severity score): experimental 7, control 7	
Interventions	**Experimental:** 8 L/min via a medium concentration mask (Hudson RCI, Durham NC, USA) which delivers a FiO_2 of between 0.4 and 0.78 **Control:** received oxygen only if their saturation was ≤92% on room air, with oxygen titrated as required at 5 minute intervals, to achieve an oxygen saturation of 93 to 95% **Co-intervention:** antibiotics, analgesia and intravenous fluids **Duration:** 1 hour	
Outcomes	**Outcomes were not pre-defined; the following were reported:** ▪ Change in $PtCO_2$ ▪ Respiratory rate ▪ Reduction in heart rate ▪ Hospital admission **Timing of outcome measurements:** 20 minutes, 40 minutes and 1 hour No relevant outcomes reported	
Notes		
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Computer-generated random numbers
Allocation concealment (selection bias)	Low risk	Allocation concealment was achieved using a secure database which contained the randomization sequence. Allocation was revealed to the researchers only when the subjects name was entered
Blinding of participants and personnel (performance bias)	High risk	Un-blinded
Blinding of outcome assessment (detection bias)	High risk	Un-blinded

Incomplete outcome data (attrition bias)	Low risk	Less than 5% was lost to follow-up. Three patients were withdrawn from the high concentration oxygen group prior to the administration of oxygen due to the inability to obtain $PtCO_2$ recordings (n = 1) and two protocol violations which included the inadvertent enrolment of a patient with COPD and another with obesity hypoventilation syndrome
Selective reporting (reporting bias)	Low risk	The protocol was pre-published, and all outcomes were reported
Other bias	Low risk	The trial was funded by public funds
Wilson 1997 [48]		
Methods	Randomised clinical trial	
Participants	**Sample size:** 50 (experimental 25, control 25) **Sex (% male):** not reported **Age (mean):** not reported **Country:** UK **Setting:** patients with myocardial infarction admitted to the coronary care unit **Inclusion criteria:** patients with myocardial infarction confirmed by chest pain and the presence of new electrocardiographic changes of ST elevation, pathological Q waves or new left bundle branch block **Exclusion criteria:** patients with central cyanosis, pulmonary disease requiring oxygen independent of the cardiac status or those in whom blood gas estimation showed pCO_2 >5.5 kPa, and patients with left ventricular failure requirering inotrope support **Disease severity:** not specified	
Interventions	**Experimental:** 4L/min by face mask **Control:** No supplemental oxygen **Co-intervention:** oxygen could be given by the personnel for clinical cyanosis or respiratory distress, which was documented in all cases. Thus, all patients were given oxygen. **Duration:** 24 hours	
Outcomes	**Outcomes were not pre-defined; the following were reported:** ▪ Arrhythmia ▪ ST segment changes ▪ Opiate use ▪ Hypoxaemia present, defined as SpO_2 < 90% ▪ Severe hypoxaemia present, defined as SpO_2 < 80% ▪ Lowest oxygen saturation **Timing of outcome measurements:** not specified No relevant outcomes reported	
Notes	Contact information was not identified; thus, email was not sent	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	Stated that the trial was randomised, but the method of sequence generation was not described
Allocation concealment (selection bias)	Low risk	Sealed envelopes
Blinding of participants and personnel (performance bias)	Unclear risk	Not adequately described. The control group was not given supplemental oxygen and mask with room air was not used/not described. It is stated that the medical and nursing staff looking after the patients were unaware of the pulse oximeter recordings, however as one group did not receive oxygen, this cannot be justified as a blinded trial

Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	High risk	Eight patients did not complete the the 24 hour trial period (1 died, 1 had a cerebrovascular accident, 4 withdrew consent, 2 had incomplete data collection) and were excluded from the analysis. In addition, it was not stated in which group the patients were allocated
Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	Unclear risk	It was unclear how the trial was funded

Wu 2014 [49]	
Methods	Randomised clinical trial
Participants	**Sample size:** 9 **Sex:** 6 men and 3 women (not specified by treatment group) **Age:** 61 (not specified by treatment group) **Country:** China **Setting:** patients with acute exacerbation chonic obstructive pilmonary disease **Inclusion criteria:** patients diagnosed with acute exacerbation chonic obstructive pilmonary disease **Exclusion criteria:** patients with serious heart diseases and other diseases of respiratory system
Interventions	**Experimental:** group B, 6–7 L/min and group C, 8–9 L/min **Control:** group A, 4–5 L/min **Co-intervention:** normal saline (0.9%, 5 mL), gentamicin (8 3 104 IU), chymotrypsin (0.4 3 104 U) and ipratropium bromide (500 mg) were nebulized for 15 minutes by means of a breath-enhanced nebulizer (PARILCD, Bonn, Germany) driven by oxygen **Duration:** 15 minutes
Outcomes	▪ Heart rate ▪ Respiratory rate ▪ SpO_2 ▪ PaO_2 ▪ $PaCO_2$ ▪ SaO_2 ▪ pH No relevant outcomes reported
Notes	Email sent to Dr Gen-di Lu 16 August 2019. Reminder sent 26 August 2019. No reply was received

Risk of bias assessment

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Random number table
Allocation concealment (selection bias)	Unclear risk	Not described
Blinding of participants and personnel (performance bias)	High risk	Unblinded
Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	Unclear risk	It was unclear whether all patients completed the trial
Selective reporting (reporting bias)	Unclear risk	No protocol could be found
Other bias	Unclear risk	It was unclear how the trial was funded

Young 2014 [50]		
Methods	Randomised clinical trial	
Participants	**Sample size:** 18 (experimental 10 (1 lost to follow-up), control 8) **Sex (% males):** experimental 100%, control 87.5% **Age (mean years):** experimental 61, control 72 **Country:** New Zealand **Setting:** adults with return of spontaneous circulation (ROSC) following out-of-hospital cardiac arrest (OHCA) caused by ventricular fibrillation (VF) or ventricular tachycardia (VT) **Inclusion criteria:** patients who were ventilated via a laryngeal mask airway or endotracheal tube were potentially eligible for study inclusion if they had an estimated age of 16–90 years and had ROSC following an OHCA due to a suspected primary cardiac cause with an initial rhythm of VF or VT **Exclusion criteria:** if patients were obviously pregnant, living in supported care or a nursing home, were known to have a terminal disease, or if more than 20 min had elapsed since ROSC	
Interventions	**Experimental:** SoO_2 < 95% **Control:** SpO_2 90-94% **Co-intervention:** not described **Duration:** 72 hours or until extubation (whichever was sooner)	
Outcomes	**Primary:** ▪ Median SpO_2 in the pre-hospital period **Secondary:** ▪ Assessments of oxygen exposure in the emergency department and the ICU SpO_2 on arrival and every 30 min thereafter while in the emergency department ▪ SpO_2 and PaO_2 every 6 hours up until extubation or 72 hours in the ICU ▪ Number of patients with hypoxia episodes (SpO_2 < 88%) in the ICU ▪ Arterial partial pressure of carbon dioxide ($PaCO_2$) every 6 h in the ICU up until extubation or 72 hours (whichever was first). **Tertiary:** ▪ Recruitment rate (based on the number of patients recruited into the study as a proportion of the total number of eligible patients) ▪ Proportion of patients with sufficiently good neurological function to be discharged home or to a rehabilitation facility ▪ ICU and hospital length of stay ▪ Quality of life at six months assessing using the EQ5D	
Notes	Email sent to Dr Young 16 August 2019 and reply was received. Additional data on mortality and health related quality of life was received)	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomisation schedule generated by a statistician
Allocation concealment (selection bias)	Low risk	Numbered sealed envelopes
Blinding of participants and personnel (performance bias)	High risk	Only patients were blinded to allocation group

Blinding of outcome assessment (detection bias)	Unclear risk	Not described
Incomplete outcome data (attrition bias)	High risk	Dropout 5.5% post randomisation
Selective reporting (reporting bias)	High risk	Protocol was registered prior to randomisation; however, the following outcomes were not reported: proportion of participants who survived with good neurological outcomes, duration of ICU and hospital stay, quality of life at 6 months
Other bias	High risk	The trial was funded by public funds. Early stopping: planned sample size was 42 patients; however, the trial was terminated early after the recruitment of 18 patients after consultation between the Data Safety Monitoring Board and Study Management Committee
Zughaft 2013 [51]		
Methods	Randomised clinical trial	
Participants	**Sample size:** 305 (300 analysed: experimental 154, control 146) **Sex (male):** experimental 79, control 80 **Age (mean years):** experimental 66, control 66 **Country:** Sweden **Setting:** patients undergoing PCI **Inclusion criteria:** clinical evidence of stable angina or acute coronary syndrome, 18 years of age, angiographic significant stenosis eligible for PCI according to ESC guidelines (8), an oxygen saturation 95% and signed informed consent **Exclusion criteria:** patients presenting with STEMI, hypoxia defined as oxygen saturation 95%, confusion and/or inability to comprehend the study information	
Interventions	**Experimental:** 3L/min nasal oxygen **Control:** 3L/min nasal air **Co-intervention:** all patients received an intravenous injection of 2.5 mg diazepam upon arrival at cathlab. Following insertion of an arterial sheath, 70 U/kg heparin and 0.2 mg nitroglycerin was injected. During the procedure 2.5 mg of morphine was administered. If pain did not diminish, the morphine dose was repeated **Duration:** the treatment was terminated after the PCI, defined as 5 minutes after removal of the guiding catheter	
Outcomes	▪ Analgesic effects of oxygen using VAS (chest pain) ▪ Troponin levels after PCI Quantity of analgesic agents administered	
Notes	Email sent to Dr Erlinge 16 August 2019 and reply was received	
Risk of bias assessment		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	The random generation was achieved manually by a system of concealed envelops in a closed box (confirmed by email).
Allocation concealment (selection bias)	Low risk	The allocation was concealed by the system of envelopes. An independent person (nurse) performed the randomization and provided the patient with either oxygen or air. This person was not allowed to address the patient, nor the staff involved in the PCI. The oxygen and air tubes were exactly of the same fabric and colour, and the outlet from the wall was concealed by a non-transparent screen (confirmed by email).
Blinding of participants and personnel (performance bias)	Low risk	A non-transparent screen was present during the entire procedure. If the patient developed hypoxia, the blinding was aborted by a nurse not involved in the trial. After the procedure, a nurse blinded to all involvement also performed the VAS measurement (confirmed by email).

Blinding of outcome assessment (detection bias)	Low risk	Only investigator blinded to treatment allocation pain-scored the patients. Not specified for mortality.
Incomplete outcome data (attrition bias)	High risk	Participants who developed SpO_2 < 95% were excluded. 5 patients were excluded, and groups of allocation were not reported for these patients.
Selective reporting (reporting bias)	High risk	The trial was registered retrospectively
Other bias	Low risk	The trial was funded by local budget (confirmed by email).

In total 42 emails were sent (excluding reminders): 24 replied and 18 did not reply or provided no clarification. Contact details could not be identified in 6 trial reports.

RESULTS OF META-ANALYSES ON PRIMARY AND SECONDARY OUTCOMES, SUBGROUP ANALYSES AND SENSITIVITY ANALYSES

Outcome or Subgroup	Trials	Participants	Statistical Method	Effect Estimate
All-cause mortality	**34**	**19439**	**RR (M-H, Fixed, 95% CI)**	**1.04 [0.96, 1.13]**
All-cause mortality - risk of bias Test-of-interaction P=0.02	34	19439	RR (M-H, Fixed, 95% CI)	1.04 [0.96, 1.11]
▪ Overall low risk of bias except for blinding	8	16156	RR (M-H, Fixed, 95% CI)	0.98 [0.89, 1.09]
▪ Overall high risk of bias	26	3283	RR (M-H, Fixed, 95% CI)	1.21 [1.05, 1.38]
All-cause mortality - used oxygenation/target Test-of-interaction P=0.47	34	19439	RR (M-H, Fixed, 95% CI)	1.04 [0.96, 1.13]
▪ FiO_2	15	15957	RR (M-H, Fixed, 95% CI)	1.02 [0.90, 1.15]
▪ PaO_2 or SaO_2/SpO_2	9	1838	RR (M-H, Fixed, 95% CI)	1.03 [0.90, 1.18]
▪ FiO_2 in higher group and SaO_2/SpO_2 in lower group	10	1644	RR (M-H, Fixed, 95% CI)	1.16 [0.97, 1.38]
All-cause mortality - oxygen saturation/target in control group Test-of-interaction P=0.75	34	19439	RR (M-H, Fixed, 95% CI)	1.04 [0.96, 1.13]
▪ Low	18	16868	RR (M-H, Fixed, 95% CI)	1.05 [0.95, 1.17]
▪ High	16	2571	RR (M-H, Fixed, 95% CI)	1.03 [0.90, 1.17]
All-cause mortality - subpopulation - patients randomised prior to hospital admission Test-of-interaction P=0.36	34	19439	RR (M-H, Fixed, 95% CI)	1.04 [0.96, 1.13]
▪ Yes	8	7710	RR (M-H, Fixed, 95% CI)	1.15 [0.92, 1.44]
▪ No	26	11729	RR (M-H, Fixed, 95% CI)	1.03 [0.94, 1.12]
All-cause mortality - subpopulation - patients admitted to the ICU Test-of-interaction P=0.71	34	19439	RR (M-H, Fixed, 95% CI)	1.04 [0.96, 1.13]
▪ Yes	8	2244	RR (M-H, Fixed, 95% CI)	1.06 [0.95, 1.19]
▪ No	26	17195	RR (M-H, Fixed, 95% CI)	1.03 [0.92, 1.15]
All-cause mortality - subpopulation - patients with any cerebral disease Test-of-interaction P=0.81	34	19439	RR (M-H, Fixed, 95% CI)	1.04 [0.96, 1.13]
▪ Yes	15	8561	RR (M-H, Fixed, 95% CI)	1.03 [0.91, 1.17]
▪ No	19	10878	RR (M-H, Fixed, 95% CI)	1.05 [0.95, 1.17]
All-cause mortality - subpopulation - patients with any cardiac disease Test-of-interaction P=0.67	34	19439	RR (M-H, Fixed, 95% CI)	1.04 [0.96, 1.13]
▪ Yes	14	8222	RR (M-H, Fixed, 95% CI)	1.09 [0.88, 1.34]
▪ No	20	11217	RR (M-H, Fixed, 95% CI)	1.04 [0.95, 1.13]
All-cause mortality - subpopulation - patients with any trauma Test-of-interaction P=0.59	34	19439	RR (M-H, Fixed, 95% CI)	1.04 [0.96, 1.13]
▪ Yes	2	103	RR (M-H, Fixed, 95% CI)	0.85 [0.40, 1.80]
▪ No	32	19336	RR (M-H, Fixed, 95% CI)	1.05 [0.96, 1.14]
All-cause mortality - subpopulation - patients with out-of-hospital-cardiac arrest Test-of-interaction P=0.38	34	19439	RR (M-H, Fixed, 95% CI)	1.04 [0.96, 1.13]
▪ Yes	5	261	RR (M-H, Fixed, 95% CI)	1.19 [0.89, 1.59]
▪ No	29	19178	RR (M-H, Fixed, 95% CI)	1.04 [0.95, 1.13]

All-cause mortality - subpopulation - patients with any lung disease Test-of-interaction P=0.09	34	19439	RR (M-H, Fixed, 95% CI)	1.04 [0.96, 1.13]
▪ Yes	2	439	RR (M-H, Fixed, 95% CI)	2.05 [0.94, 4.46]
▪ No	32	19000	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.12]
All-cause mortality - subpopulation - patients with COPD Test-of-interaction P=0.09	34	19439	RR (M-H, Fixed, 95% CI)	1.04 [0.96, 1.13]
▪ Yes	2	439	RR (M-H, Fixed, 95% CI)	2.05 [0.94, 4.46]
▪ No	32	19000	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.12]
All-cause mortality - duration of oxygen administration Test-of-interaction P=0.53	33	19391	RR (M-H, Fixed, 95% CI)	1.04 [0.96, 1.13]
▪ Above median duration ▪ (12 hours)	18	10653	RR (M-H, Fixed, 95% CI)	1.03 [0.94, 1.13]
▪ Below median duration ▪ (12 hours)	15	8738	RR (M-H, Fixed, 95% CI)	1.11 [0.90, 1.36]
All-cause mortality - post hoc analysis - default administration of supplemental oxygen in control group Test-of-interaction P=0.92	34	19439	RR (M-H, Fixed, 95% CI)	1.04 [0.96, 1.13]
▪ Yes	21	3926	RR (M-H, Fixed, 95% CI)	1.05 [0.94, 1.17]
▪ No	13	15513	RR (M-H, Fixed, 95% CI)	1.04 [0.92, 1.18]
All-cause mortality - sensitivity analysis – best-worst case scenario	34	20139	RR (M-H, Random, 95% CI)	0.80 [0.65, 0.99]
All-cause mortality - sensitivity analysis – worst-best case scenario	34	20119	RR (M-H, Fixed, 95% CI)	1.40 [1.30, 1.52]
Proportion of participants with at least one SAE as reported by trialists	**6**	**8874**	**RR (M-H, Fixed, 95% CI)**	**1.03 [0.95, 1.13]**
Proportion of participants with at least one SAE as reported by trialists - risk of bias Test-of-interaction P=0.08	6	8874	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.13]
▪ Overall low risk of bias except for blinding	3	8056	RR (M-H, Fixed, 95% CI)	0.99 [0.89, 1.12]
▪ Overall high risk of bias	3	818	RR (M-H, Fixed, 95% CI)	1.14 [1.03, 1.26]
Proportion of participants with at least one SAE as reported by trialists - used oxygenation/target Test-of-interaction P=0.15	6	8874	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.13]
▪ FiO_2	5	8440	RR (M-H, Fixed, 95% CI)	1.01 [0.90, 1.12]
▪ PaO_2 or SaO_2/SpO_2	0	0	RR M-H, Fixed, 95% CI)	Not estimable
▪ FiO_2 in higher group and SaO_2/SpO_2 in lower group	1	434	RR (M-H, Fixed, 95% CI)	1.12 [1.02, 1.23]
Proportion of participants with at least one SAE as reported by trialists - oxygen saturation/target in control group Test-of-interaction: not applicaple	6	8874	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.13]
▪ Low	6	8874	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.13]
▪ High	0	0	RR (M-H, Fixed, 95% CI)	Not estimable
Proportion of participants with at least one SAE as reported by trialists - subpopulation - patients randomised prior to hospital admission Test-of-interaction: not applicaple	6	8874	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.13]
▪ Yes	0	0	RR (M-H, Fixed, 95% CI)	Not estimable
▪ No	6	8874	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.13]
Proportion of participants with at least one SAE as reported by trialists -	6	8874	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.13]

subpopulation - patients admitted to the ICU Test-of-interaction P=0.15				
▪ Yes	1	434	RR (M-H, Fixed, 95% CI)	1.12 [1.02, 1.23]
▪ No	5	8440	RR (M-H, Fixed, 95% CI)	1.01 [0.90, 1.12]
Proportion of participants with at least one SAE as reported by trialists - subpopulation - patients with any cerebral disease Test-of-interaction P=0.21	6	8874	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.13]
▪ Yes	3	7761	RR (M-H, Fixed, 95% CI)	1.00 [0.88, 1.13]
▪ No	3	1113	RR (M-H, Fixed, 95% CI)	1.10 [1.00, 1.23]
Proportion of particioants with at least one SAE as reported by trialists - subpopulation - patients with any cardiac disease Test-of-interaction P=0.40	6	8874	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.13]
▪ Yes	1	300	RR (M-H, Fixed, 95% CI)	1.58 [0.59, 4.24]
▪ No	5	8574	RR (M-H, Fixed, 95% CI)	1.03 [0.94, 1.12]
Proportion of participants with at least one SAE as reported by trialists - subpopulation - patients with any trauma Test-of-interaction: not applicable	6	8874	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.13]
▪ Yes	0	0	RR (M-H, Fixed, 95% CI)	Not estimable
▪ No	6	8874	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.13]
Proportion of participants with at least one SAE as reported by trialists - subpopulation - patients with out-of-hospital-cardiac arrest Test-of-interaction: not applicaple	6	8874	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.13]
▪ Yes	0	0	RR (M-H, Fixed, 95% CI)	Not estimable
▪ No	6	8874	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.13]
Proportion of participants with at least one SAE as reported by trialists - subpopulation - patients with any lung disease Test-of-interaction: not applicaple	6	8874	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.13]
▪ Yes	0	0	RR (M-H, Fixed, 95% CI)	Not estimable
▪ No	6	8874	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.13]
Proportion of participants with at least one SAE as reported by trialists - subpopulation - patients with COPD Test-of-interaction: not applicaple	6	8874	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.13]
▪ Yes	0	0	RR (M-H, Fixed, 95% CI)	Not estimable
▪ No	6	8874	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.13]
Proportion of participants with at least one SAE as reported by trialists - duration of oxygen administration Test-of-interaction P=0.67	6	8874	RR (M-H, Fixed, 95% CI)	1.03 [0.95, 1.13]
▪ Above median duration (12 hours)	3	8111	RR (M-H, Fixed, 95% CI)	1.03 [0.94, 1.13]
▪ Below median duration (12 hours)	3	763	RR (M-H, Fixed, 95% CI)	1.08 [0.86, 1.36]
Proportion of participants with at least one SAE as reported by trialists - sensitivity analysis - best-worst case scenario	6	9215	RR (M-H, Fixed, 95% CI)	0.84 [0.77, 0.91]
Proportion of participants with at least one SAE as reported by trialists -	6	9215	RR (M-H, Fixed, 95% CI)	1.25 [1.15, 1.37]

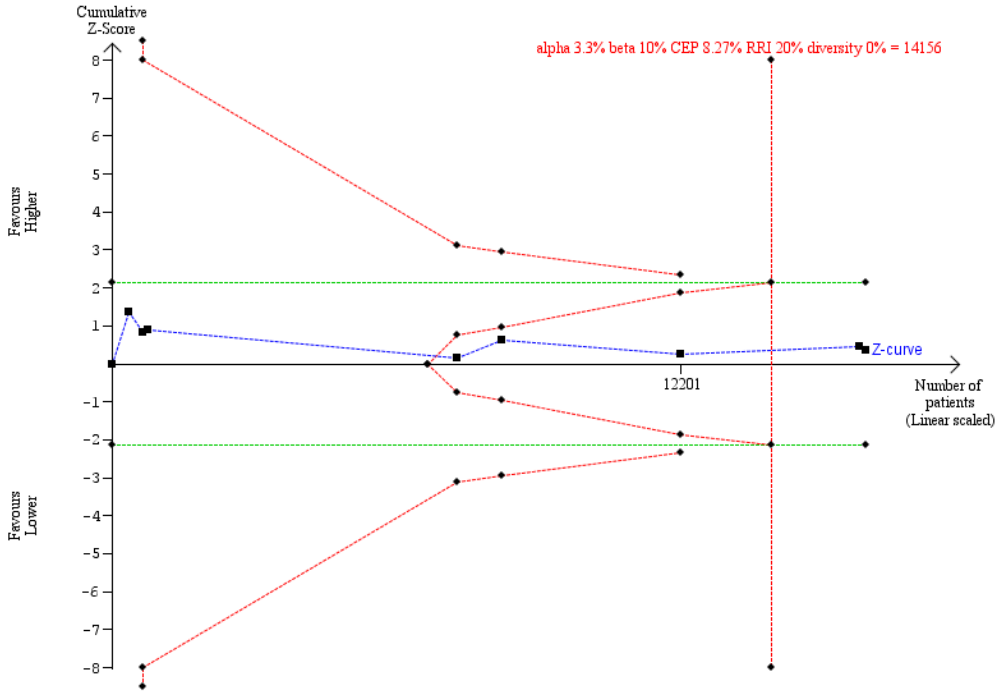
sensitivity analysis - worst-best case scenario				
SAEs - highest proportion	**35**	**19454**	**RR (M-H, Fixed, 95% CI)**	**1.05 [0.98, 1.11]**
SAEs - cumulated	35	19502	RR (M-H, Fixed, 95% CI)	1.03 [0.98, 1.09]
SAEs - highest proportion - risk of bias Test-of-interaction P=0.01	35	19454	RR (M-H, Fixed, 95% CI)	1.05 [0.98, 1.11]
▪ Overall low risk of bias except for blinding	8	16156	RR (M-H, Fixed, 95% CI)	1.00 [0.93, 1.08]
▪ Overall high risk of bias	27	3298	RR (M-H, Fixed, 95% CI)	1.17 [1.06, 1.29]
SAEs - highest proportion - used oxygenation/target Test-of-interaction P=0.05	35	19454	RR (M-H, Fixed, 95% CI)	1.05 [0.98, 1.11]
▪ FiO_2	16	15972	RR (M-H, Fixed, 95% CI)	1.01 [0.93, 1.10]
▪ PaO_2 or SaO_2/SpO_2	9	1838	RR (M-H, Fixed, 95% CI)	1.04 [0.91, 1.19]
▪ FiO_2 in higher group and SaO_2/SpO_2 in lower group	10	1644	RR (M-H, Fixed, 95% CI)	1.19 [1.08, 1.33]
SAEs - highest proportion - oxygen saturation/target in control group Test-of-interaction P=0.81	35	19454	RR (M-H, Fixed, 95% CI)	1.05 [0.98, 1.11]
▪ Low	18	16827	RR (M-H, Fixed, 95% CI)	1.04 [0.97, 1.12]
▪ High	17	2627	RR (M-H, Fixed, 95% CI)	1.06 [0.94, 1.20]
SAEs - highest proportion - subpopulation - patients randomised prior to hospital admission Test-of-interaction P=0.37	35	19454	RR (M-H, Fixed, 95% CI)	1.05 [0.98, 1.11]
▪ Yes	8	7711	RR (M-H, Fixed, 95% CI)	1.10 [0.97, 1.24]
▪ No	27	11743	RR (M-H, Fixed, 95% CI)	1.03 [0.96, 1.10]
SAEs - highest proportion - subpopulation - patients admitted to the ICU Test-of-interaction P=0.93	35	19454	RR (M-H, Fixed, 95% CI)	1.05 [0.98, 1.11]
▪ Yes	9	2304	RR (M-H, Fixed, 95% CI)	1.05 [0.96, 1.15]
▪ No	26	17150	RR (M-H, Fixed, 95% CI)	1.05 [0.97, 1.13]
SAEs - highest proportion - subpopulation - patients with any cerebral disease Test-of-interaction P=0.38	35	19454	RR (M-H, Fixed, 95% CI)	1.05 [0.98, 1.11]
▪ Yes	15	8560	RR (M-H, Fixed, 95% CI)	1.01 [0.91, 1.12]
▪ No	20	10894	RR (M-H, Fixed, 95% CI)	1.07 [0.99, 1.15]
SAEs - highest proportion - subpopulation - patients with any cardiac disease Test-of-interaction P=0.63	35	19454	RR (M-H, Fixed, 95% CI)	1.05 [0.98, 1.11]
▪ Yes	15	8238	RR (M-H, Fixed, 95% CI)	1.07 [0.95, 1.20]
▪ No	20	11216	RR (M-H, Fixed, 95% CI)	1.04 [0.96, 1.11]
SAEs - highest proportion - subpopulation - patients with any trauma Test-of-interaction P=0.81	35	19454	RR (M-H, Fixed, 95% CI)	1.05 [0.98, 1.11]
▪ Yes	2	101	RR (M-H, Fixed, 95% CI)	0.97 [0.51, 1.85]
▪ No	33	19353	RR (M-H, Fixed, 95% CI)	1.05 [0.98, 1.11]
SAEs - highest proportion - subpopulation - patients with out-of-hospital-cardiac arrest Test-of-interaction P=0.41	35	19454	RR (M-H, Fixed, 95% CI)	1.05 [0.98, 1.11]
▪ Yes	5	261	RR (M-H, Fixed, 95% CI)	1.18 [0.89, 1.55]
▪ No	30	19193	RR (M-H, Fixed, 95% CI)	1.04 [0.98, 1.11]

SAEs - highest proportion - subpopulation - patients with any lung disease Test-of-interaction P=0.68	35	19454	RR (M-H, Fixed, 95% CI)	1.05 [0.98, 1.11]
▪ Yes	3	499	RR (M-H, Fixed, 95% CI)	1.19 [0.65, 2.16]
▪ No	32	18955	RR (M-H, Fixed, 95% CI)	1.05 [0.98, 1.11]
SAEs - highest proportion - subpopulation - patients with COPD Test-of-interaction P=0.68	35	19454	RR (M-H, Fixed, 95% CI)	1.05 [0.98, 1.11]
▪ Yes	3	499	RR (M-H, Fixed, 95% CI)	1.19 [0.65, 2.16]
▪ No	32	18955	RR (M-H, Fixed, 95% CI)	1.05 [0.98, 1.11]
SAEs - highest proportion - duration of oxygen administration Test-of-interaction P=0.52	35	19454	RR (M-H, Fixed, 95% CI)	1.05 [0.98, 1.11]
▪ Above median duration ▪ (12 hours)	18	10610	RR (M-H, Fixed, 95% CI)	1.03 [0.96, 1.11]
▪ Below median duration ▪ (12 hours)	17	8844	RR (M-H, Fixed, 95% CI)	1.08 [0.97, 1.20]
SAEs - sensitivity analysis – best-worst case scenario	34	20138	RR (M-H, Random, 95% CI)	0.82 [0.69, 0.97]
SAEs - sensitivity analysis – worst-best case scenario	34	20118	RR (M-H, Fixed, 95% CI)	1.35 [1.27, 1.44]
Quality of life	**6**	**7445**	**MD (IV, Random, 95% CI)**	**0.37 [-1.61, 2.35]**
Lung injury - highest proportion	**10**	**9227**	**RR (M-H, Fixed, 95% CI)**	**0.93 [0.76, 1.12]**
Lung injury - cumulated	10	9279	RR (M-H, Fixed, 95% CI)	0.92 [0.78, 1.10]
Pulmonary fibrosis	0	0	RR (M-H, Fixed, 95% CI)	Not estimable
ARDS	4	324	RR (M-H, Fixed, 95% CI)	0.85 [0.45, 1.60]
Pneumonia	7	9039	RR (M-H, Fixed, 95% CI)	0.92 [0.76, 1.12]
Sepsis	**4**	**1307**	**RR (M-H, Fixed, 95% CI)**	**1.64 [0.96, 2.80]**
Cardiovascular events - highest proportion	**16**	**16607**	**RR (M-H, Random, 95% CI)**	**1.06 [0.86, 1.31]**
Cardiovascular events - cumulated	16	16615	RR (M-H, Fixed, 95% CI)	1.10 [0.98, 1.23]
Myocardial infarction	7	7971	RR (M-H, Random, 95% CI)	1.29 [0.69, 2.39]
Stroke	5	8797	RR (M-H, Fixed, 95% CI)	1.00 [0.70, 1.43]
Peripheral arterial thrombosis	1	434	RR (M-H, Fixed, 95% CI)	0.92 [0.43, 1.98]
Deep vein thrombosis	2	7676	RR (M-H, Fixed, 95% CI)	1.16 [0.30, 4.48]
Pulmonary embolism	3	8056	RR (M-H, Random, 95% CI)	0.98 [0.35, 2.73]

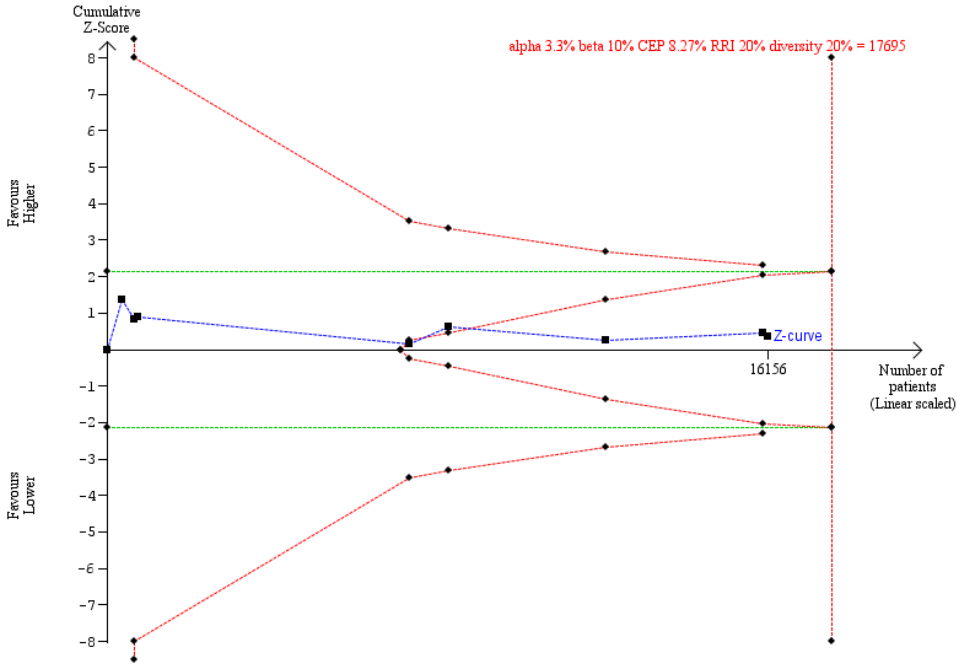
ARDS: Acute respiratory distress syndrome; **CI**: Confidence interval; **COPD**: Chronic obstructive pulmonary disease; **FiO_2**: Fractions of inspired oxygen; **M-H**: Mantel-Haenszel; **PaO_2**: Partial pressure of arterial oxygen; **RR**: Risk ratio; **SAE:** Serious adverse events; **SaO_2**: Arterial oxygen saturation; **SpO_2**: Peripheral oxygen saturation

ADDITIONAL INFORMATION AND ANALYSES ON THE OUTCOMES

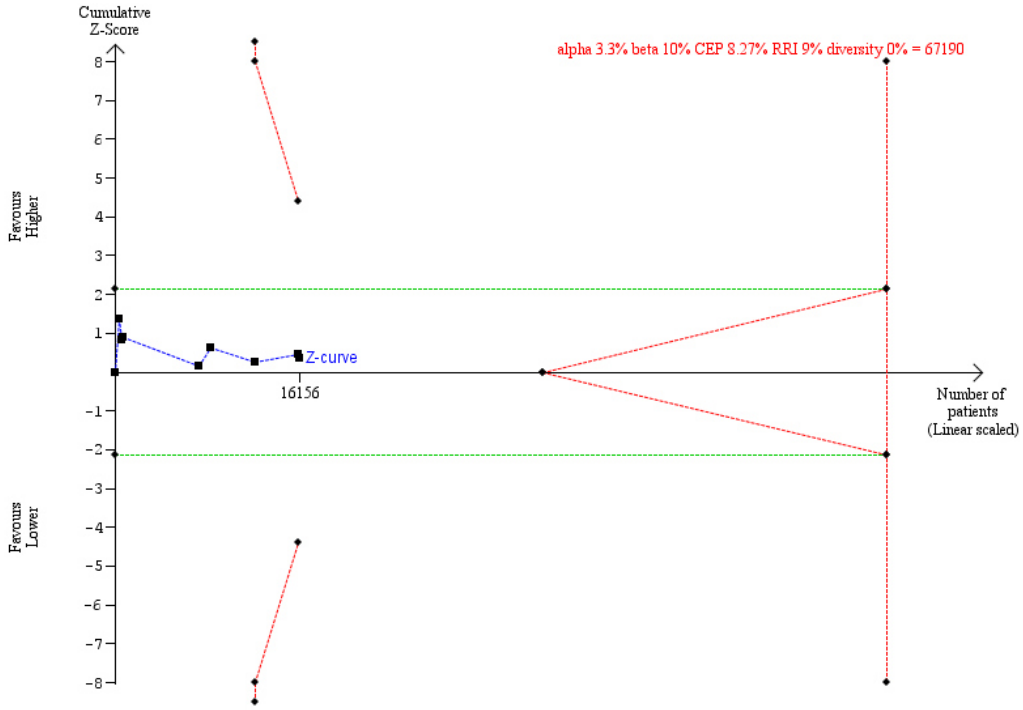
ALL CAUSE MORTALITY – OVERALL LOW RISK OF BIAS EXCEPT FOR BLINDING



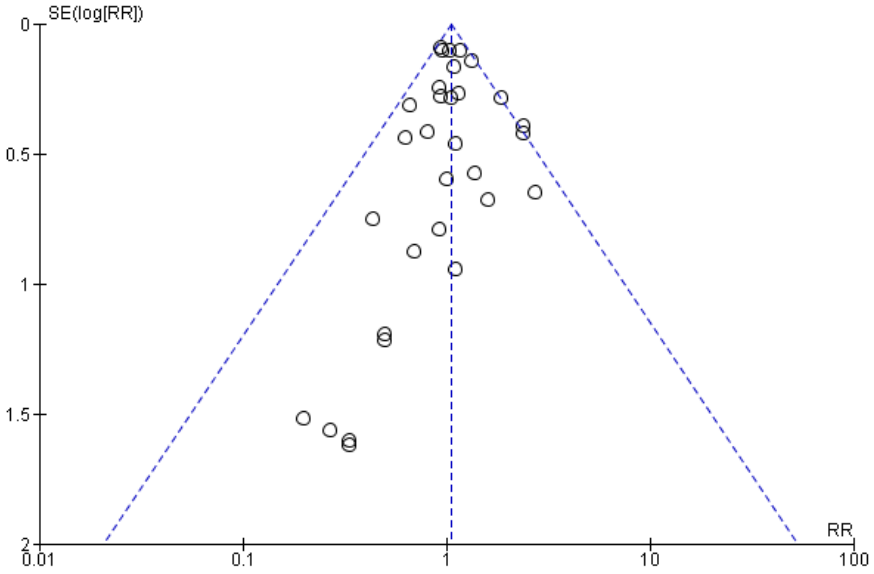
TSA on mortality on trials with overall low risk of bias except for blinding for a 20% RRI.



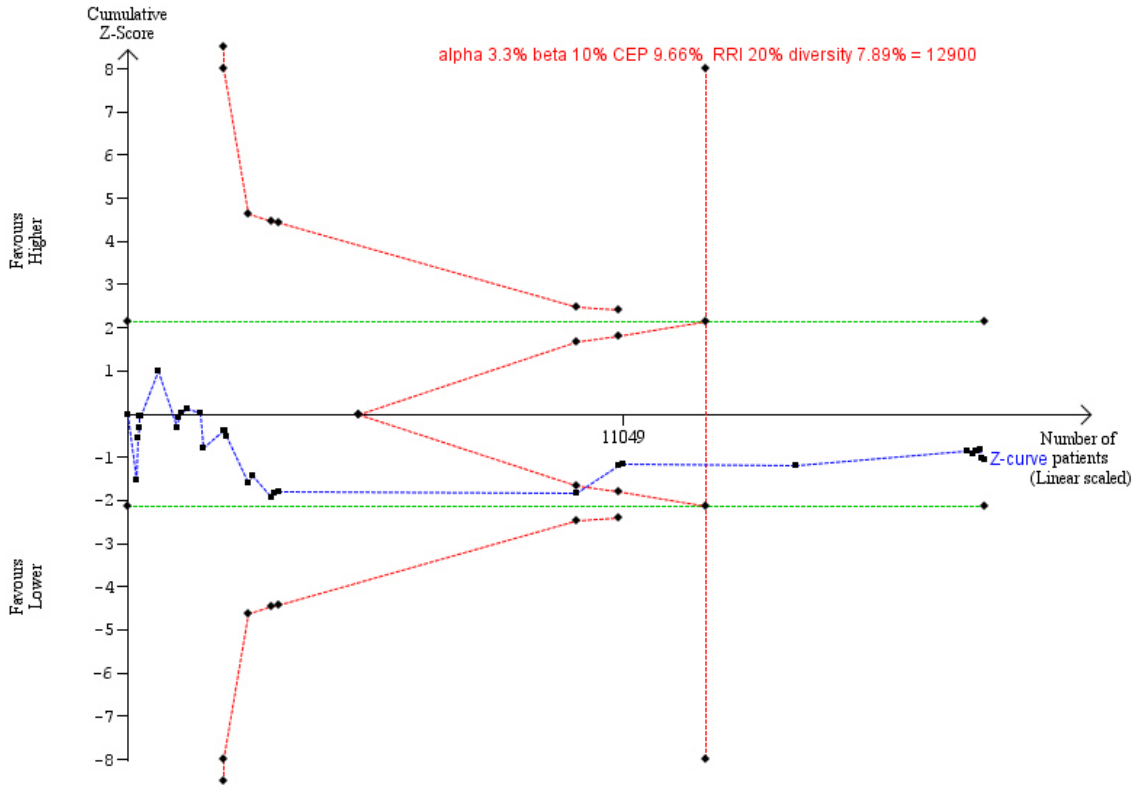
TSA on mortality on trials with overall low risk of bias except for blinding for a 20% RRI and diversity of 20%.



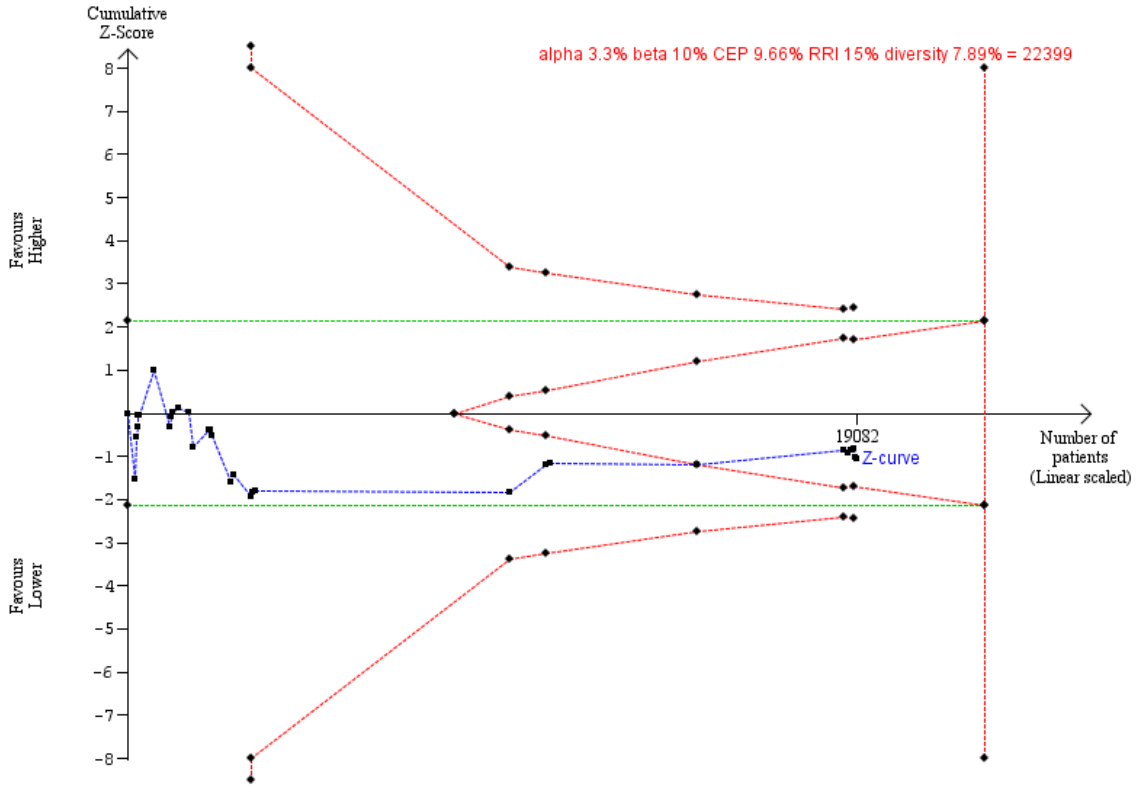
TSA on mortality on trials with overall low risk of bias except of blinding for a 9% RRI.



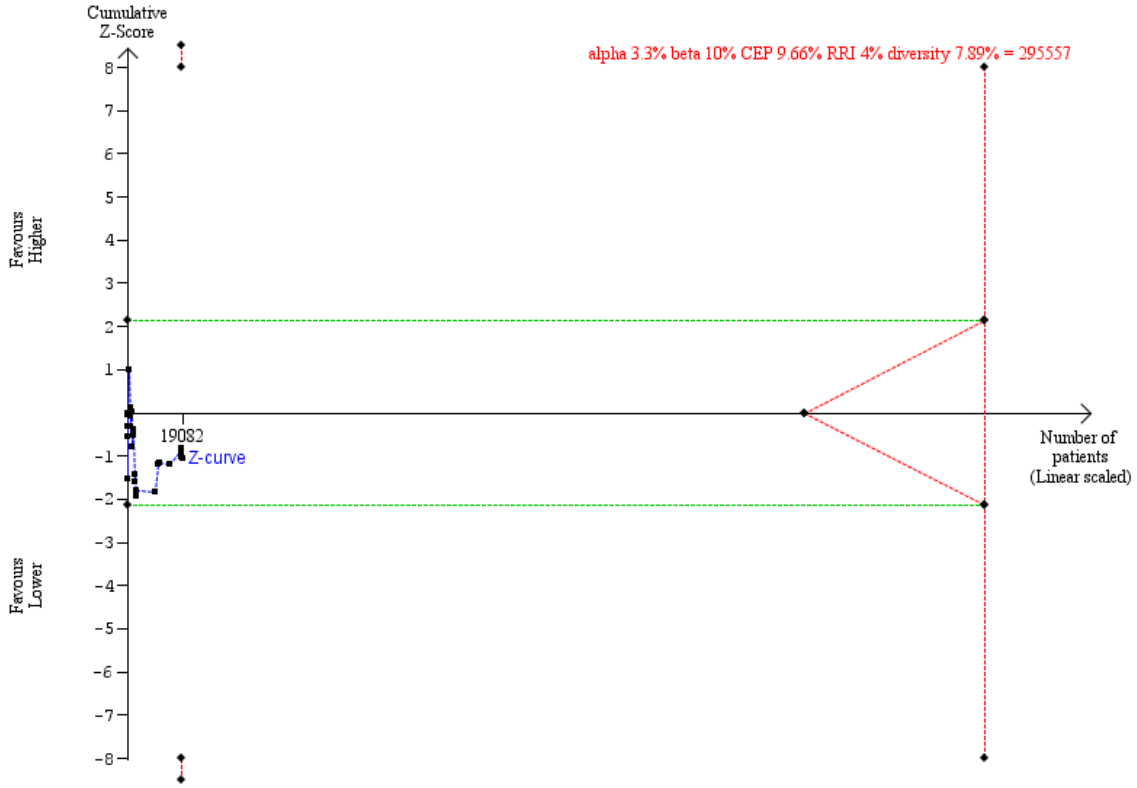
Funnel plot on mortality. Harbord test P=0.693.



TSA on mortality on all trials for a 20% RRI.

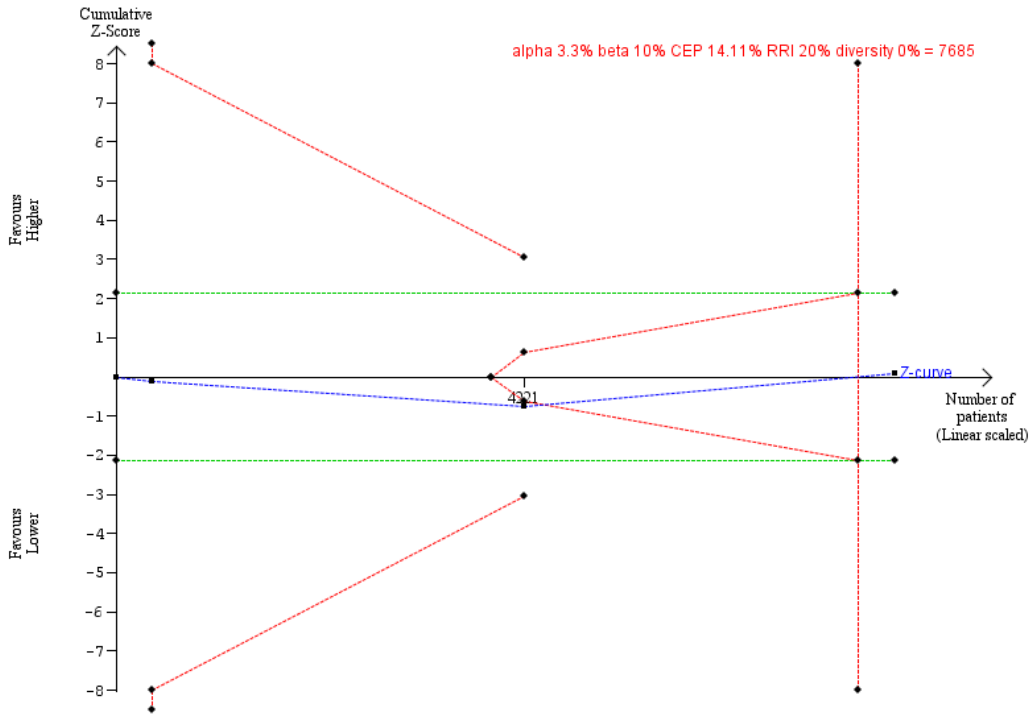


TSA on mortality on all trials for a 15% RRI.

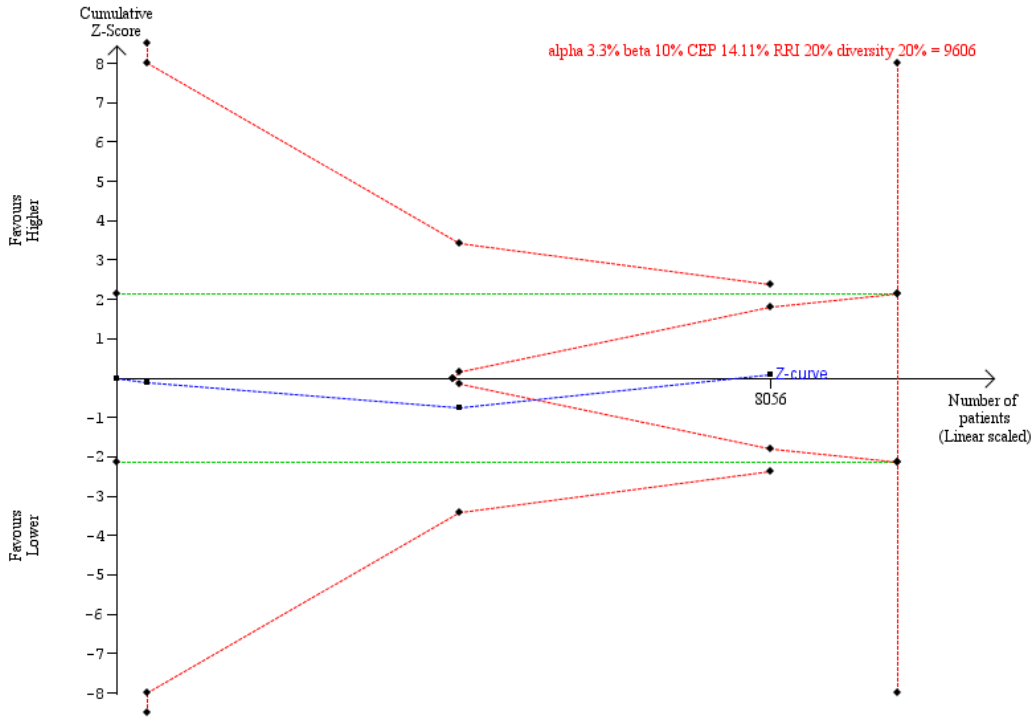


TSA on mortality on all trials for a 4% RRR.

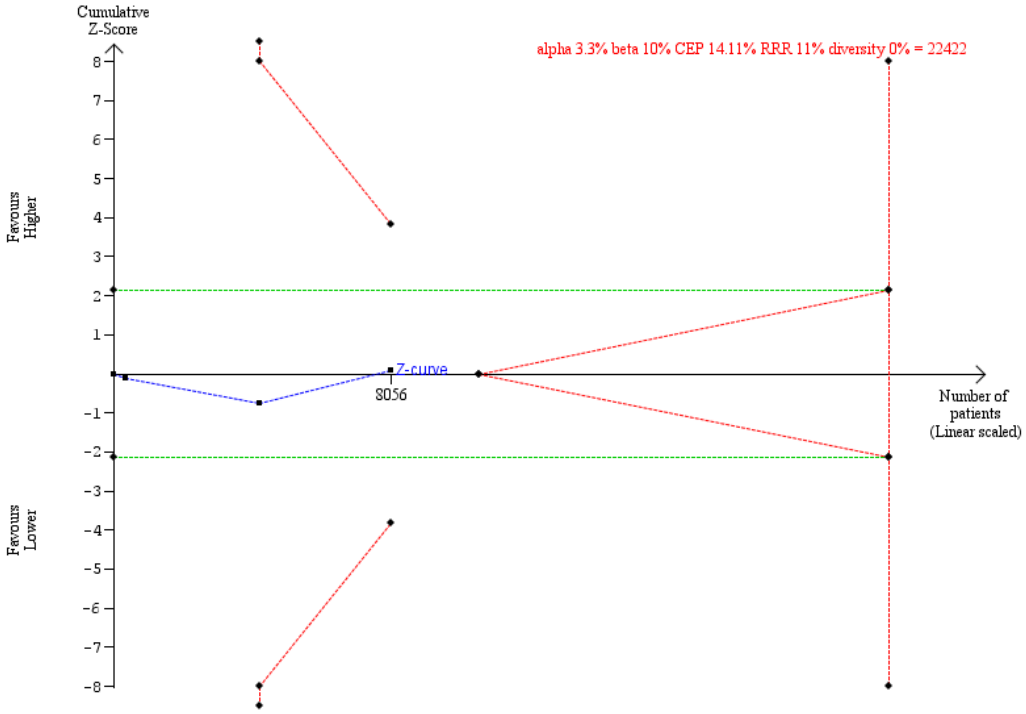
SERIOUS ADVERSE EVENTS – OVERALL LOW RISK OF BIAS



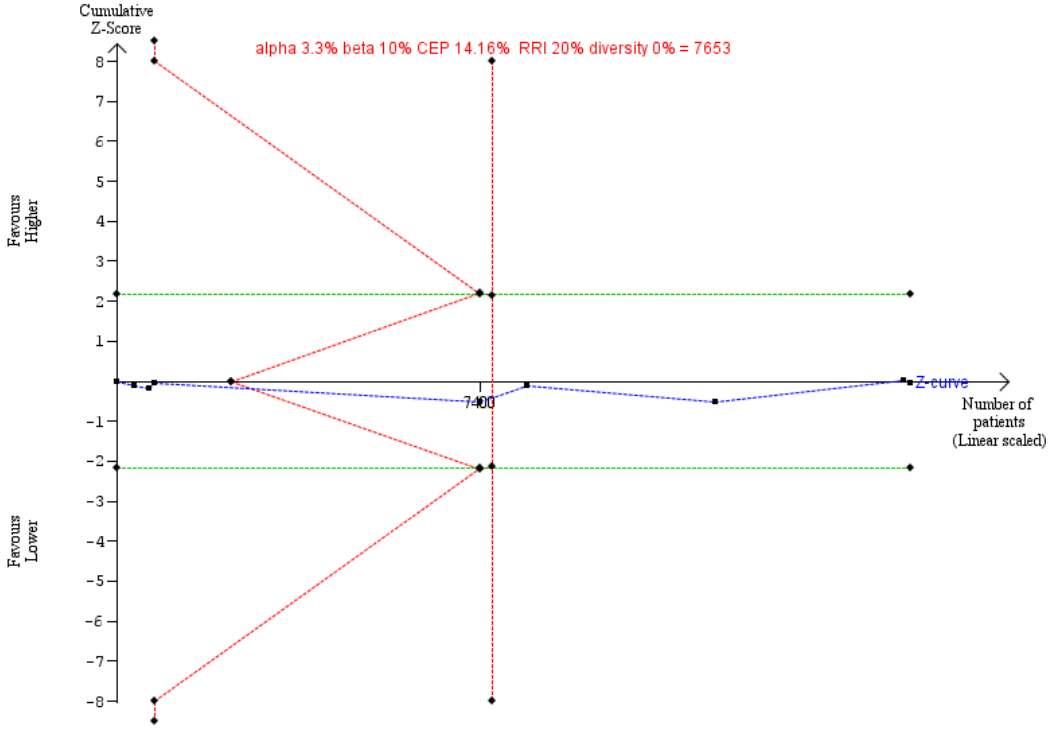
TSA on trials with overall low risk of bias except for blinding on SAEs reported by trialists on the proportion of participants with at least one SAE for a 20% RRI.



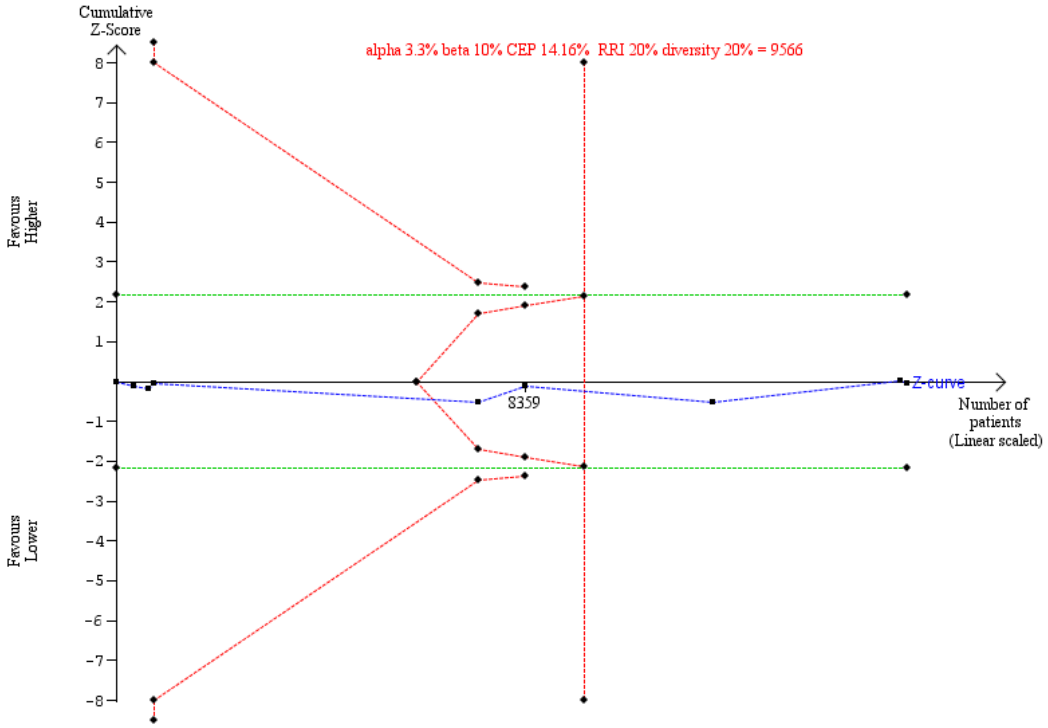
TSA on trials with overall low risk of bias except for blinding on SAEs reported by trialists as the proportion of participants with at least one SAE for a 20% RRI and a diversity of 20%.



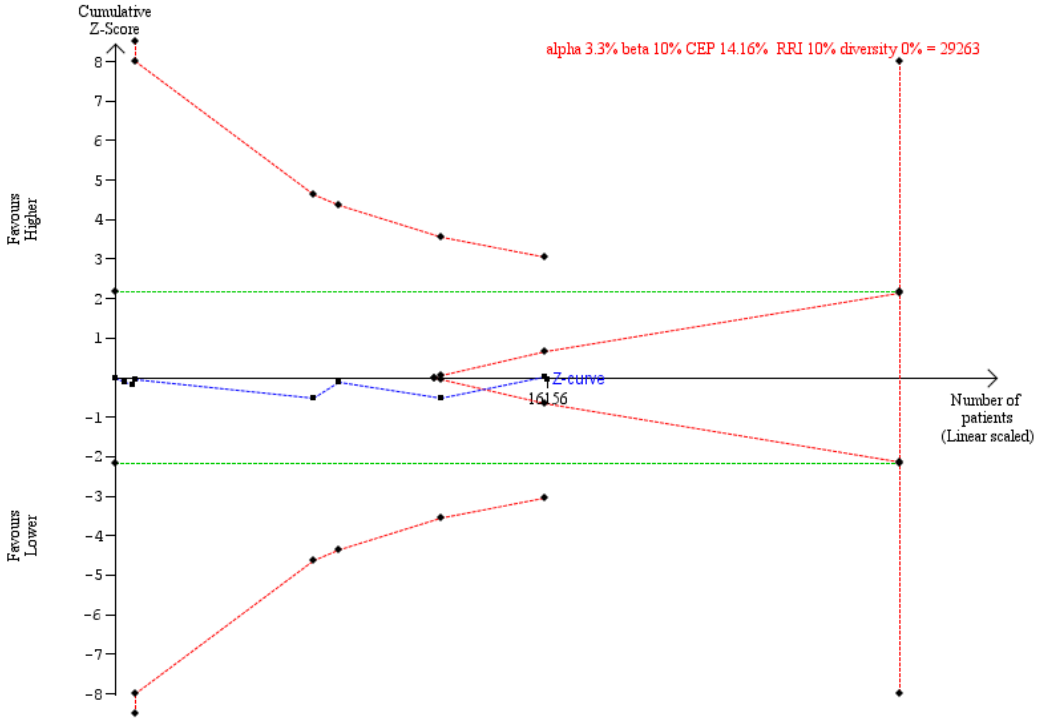
TSA on trials with overall low risk of bias except for blinding on SAEs reported by trialists as the proportion of participants with at least one SAE for a 11% RRR.



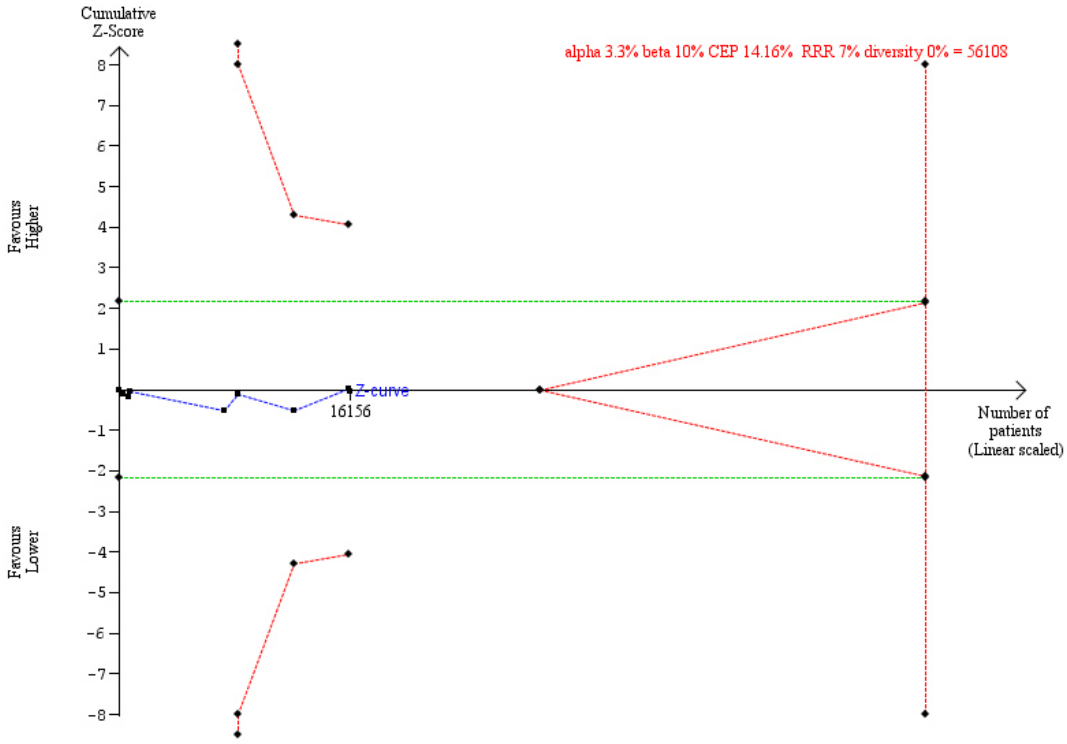
TSA on trials with overall low risk of bias except for blinding on the estimated highest reported proportion of specific SAEs in each trial for a 20% RRI.



TSA on trials with overall low risk of bias except for blinding on the estimated highest reported proportion of specific SAEs in each trial for a 20% RRI and a diversity of 20%.



TSA on trials with overall low risk of bias except for blinding on the estimated highest reported proportion of specific SAEs in each trial for a 10%.



TSA on trials with overall low risk of bias except for blinding on the estimated highest reported proportion of specific SAEs in each trial for a 7% RRR.

Results of subgroup analyses and sensitivity analyses

*P*roportion of participants with at least one SAE*, as reported by trialits*

Incomplete outcome data alone had the potential to influence the results (best-worst case scenario: RR 0.82; 95% CI 0.69-0.97 and worst-best case scenario: RR 1.35; 95% CI 1.27-1.44). The following tests of interaction showed evidence of a difference: 1) trials with overall low risk of bias except for blinding versus trials with overall high risk of bias (P=0.08). When analysing each subgroup separately, meta-analysis of trials with overall low risk of bias except for blinding showed no evidence of a difference in SAEs (RR 0.99, 95% CI 0.89-1.12), whilst trials of overall high risk of bias showed an increase in SAEs with higher oxygen supplementation (RR 1.14, 95% 1.03-1.26). Additional subgroup analyses were consistent with the primary analysis or could not be performed due du limited data.

Estimated highest reported proportion of specific SAEs in each trial

Incomplete outcome data alone had the potential to influence the results (best-worst case scenario: RR 0.82; 95% CI 0.69-0.97 and worst-best case scenario: RR 1.35; 95% CI 1.27-1.44). The following tests of interaction showed evidence of a difference: 1) trials with overall low risk of bias except for blinding versus trials with overall high risk of bias (P=0.01). When analysing each subgroup separately, meta-analysis of trials with overall low risk of bias except for blinding showed no evidence of a difference in SAEs (RR 1.00, 95% CI 0.93-1.08), whilst trials of overall high risk of bias showed an increase in SAEs with higher oxygen supplementation (RR 1.17, 95% 1.06-1.29). 2) trials using FiO_2 in the higher group and PaO_2 or SaO_2/SpO_2 in the lower group versus trials using either FiO_2 or $PaO_2/SaO_2/SpO_2$ in both groups (P=0.05). When analysing each subgroup separately, meta-analyses of trials using only FiO_2 or $PaO_2/SaO_2/SpO_2$ showed no evidence of a difference in SAEs respectively (RR 1.01, 95% CI 0.93-1.10) (RR 1.04, 95% CI 0.91-1.19), whilst trials using FiO_2 in the higher group and $PaO_2/SaO_2/SpO_2$ in the lower group showed an increase in SAEs with higher oxygen supplementation (RR 1.19, 95% CI 1.08-1.33). Additional subgroup analyses were consistent with the primary analysis.

Types of serious adverse events, from each trial, included in meta-analysis on the estimated highest reported proportion of specific serious adverse events in each trial

TRIAL ID	OUTCOME	Higher: events	Higher: analysed	Lower: events	Higher: analysed
Ali 2013	· Mortality	22	148	20	141
Asfar 2017	· Prop. of pts with one or more SAEs	185	217	165	217
Austin 2010	· Mortality	21	226	7	179
Baekgaard 2019	· 30-day mortality and/or major pulmonary complications	6	19	4	17
Bray 2018	· Mortality	11	24	18	37
Butler 1987B	· Mortality	2	17	6	22
Girardis 2016	· Mortality	80	243	58	235
Gomersall 2002	· Mechanical ventilation, invasive + non-invasive	2	19	2	17
Heidari2017	· Mortality	0	36	1	36
Hofmann 2017	· Composite of all-cause mortality or hosp for heart failure or MI	363	3311	350	3318
Huynh Ky 2017	· Mortality	0	19	0	20
ICU-ROX investigators 2019	· Mortality	156	480	166	479
Jakkula 2018	· Mortality	20	59	18	61
Jun 2019	· Myocardial Infarction	2	29	9	29
Khoshnood 2018	· Cardiogenic shock	6	46	9	49
Kuisma 2006	· Need for inotropic support	7	14	7	14
Lång 2018	· Mortality	9	38	8	27
Mazdeh 2015	· Mortality	5	26	3	25
Meyhoff 2009	· Prop. of pts with one or more SAEs	63	185	65	194
NCT02378545	· Mortality	6	25	4	23
Padma 2010	· Mortality	0	20	2	20
Panwar 2016	· Mortality	19	51	21	52
Ranchord 2012	· MACE	2	68	2	68
Rawles 1976	· Ventricular tachycardia	11	80	5	77
Roffe 2010	· Mortality	2	29	3	30
Roffe 2017A	· Prop. of pts with one or more SAEs	348	2567	161	1275
Roffe 2017B	· Prop. of pts with one or more SAEs	294	2561	161	1274
Sepehrvand 2019	· Mortality	1	25	2	25
Shi 2017	· Mortality	0	9	0	9
Singhal 2005	· Mortality	0	9	1	7
Singhal 2013	· Prop. of pts with one or more SAEs	24	43	20	41
Stub 2015	· MACE	46	218	34	223
Thomas 2019	· Mortality	14	17	8	18
Young 2014	· Mortality	5	9	4	8
Zughaft 2013	· Prop. of pts with one or more SAEs	10	154	6	146

Types of serious adverse events, from each trial, included in meta-analysis on the estimated cumulated number of serious adverse events

TRIAL ID	OUTCOME	Higher: events	Higher: analysed	Lower: events	Higher: analysed
Ali 2013	· Mortality	22	148	20	141
Asfar 2017	· Mortality · Pneumonia · Peripheral Arterial Thrombosis · Pneumothorax · Ventricular arrhythmias · Patients with 1 or more nosocomial infection during ICU stay	196	217	186	217
Austin 2010	· Mortality	21	226	7	179
Baekgaard 2019	· Mortality · ARDS · Pneumonia · Sepsis · Surgical site infection	9	19	7	20
Bray 2018	· Mortality · Cardias arrest incl. re-arrest	12	24	18	37
Butler 1987B	· Mortality	2	17	6	22
Girardis 2016	· Mortality · Pneumonia · Sepsis · No of patients with new infection(s) during ICU stay (respiratory, bacteriamia, surgical site) · Coma	191	243	139	235
Gomersall 2002	· Mortality · Mechanical ventilation, invasive + non-invasive	2	17	3	17
Heidari2017	· Mortality	0	36	1	36
Hofmann 2017	· Mortality · Myocardial infarction · Atrioventricular block, second or third degree · Cardiogenic shock · Cardiac arrest incl. re-arrest · Re-hosp for heart failure	413	3311	385	3318
Huynh Ky 2017	· Mortality	0	19	0	20
ICU-ROX investigators 2019	· Mortality · Renal replacement therapy	264	480	260	479

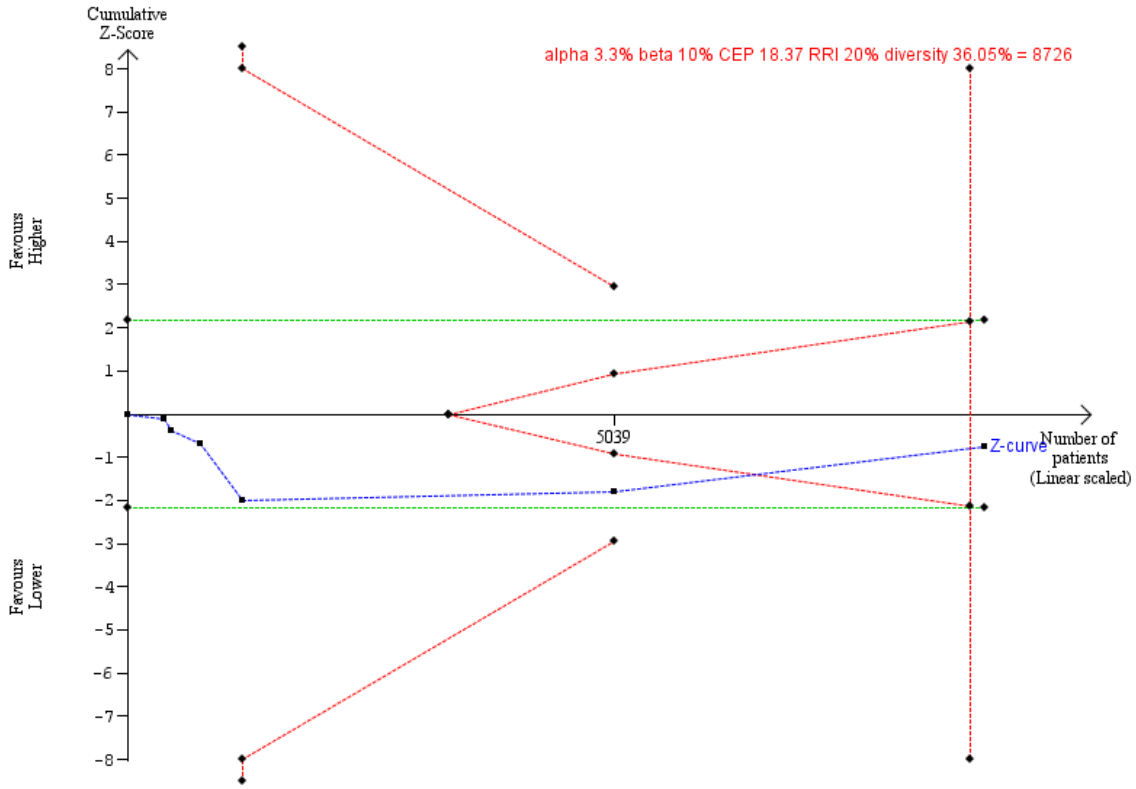
Jakkula 2018	· Mortality · ARDS · Severe hypercapnia and respiratory acidosis ($PaCO_2$ > 10 kPa and pH < 7.15)	21	59	20	61
Jun 2019	· Myocardial Infarction	2	29	9	29
Khoshnood 2018	· Mortality · Myocardial infarction · Cardiogenic shock · Cardiac arerst incl. re-arrest · Heart failure	12	49	16	45
Kuisma 2006	· Mortality · Need for inotropic support	11	14	11	14
Lång 2018	· Mortality · ARDS · Pneumonia	15	38	17	27
Mazdeh 2015	· Mortality	5	26	3	25
Meyhoff 2009	· Mortality · Pneumonia · Sepsis · MI · Stroke · Pulmonary Embolism · Re-operation	71	188	84	195
NCT02378545	· Mortality	6	25	4	23
Padma 2010	· Mortality	0	20	2	20
Panwar 2016	· Mortality · ARDS · Haemodynamic instability, defined as cardiac arrest or addition of two or more new vasopressor/inotrope agents in a day	42	51	41	52
Ranchord 2012	· Mortality · Myocardial infarction	2	68	2	68
Rawles 1976	· Mortality · Atrioventricular block, second or third degree · Ventricular tachycardia · Ventricular fibrillation	23	105	14	95
Roffe 2010	· Mortality	2	29	3	30
Roffe 2017A	· Mortality · Pneumonia · Stroke	427	2567	211	1275

	· Deep Vein Thrombosis · Pulmonary Embolism · intracranial hemorrhage · Agitation · Seizure · Transient ischemic attack				
Roffe 2017B	· Mortality · Pneumonia · Stroke · Deep Vein Thrombosis · Pulmonary Embolism · intracranial hemorrhage · Agitation · Seizure · Transient ischemic attack	381	2561	208	1274
Sepehrvand 2019	· Mortality	1	25	2	25
Shi 2017	· Mortality · Intracranial hemorrhage · Subarachnoid hemorrhage	0	9	0	9
Singhal 2005	· Mortality	0	9	1	7
Singhal 2013	· Mortality · Brain Hemorrhage	24	43	27	42
Stub 2015	· Mortality · Sepsis · Myocardial infarction · Stroke · Repeated re-vascularisation[a] · Major bleeding[b] · Cardiogenic shock[c] · Coronary artery bupass grafting	86	218	76	223
Thomas 2019	· Mortality · Cardiac arrest incl. re-arrest	22	17	11	18
Young 2014	· Mortality · Pneumothorax	5	9	5	8
Zughaft 2013	· Mortality · Myocardial infarction · Stroke	10	154	6	146

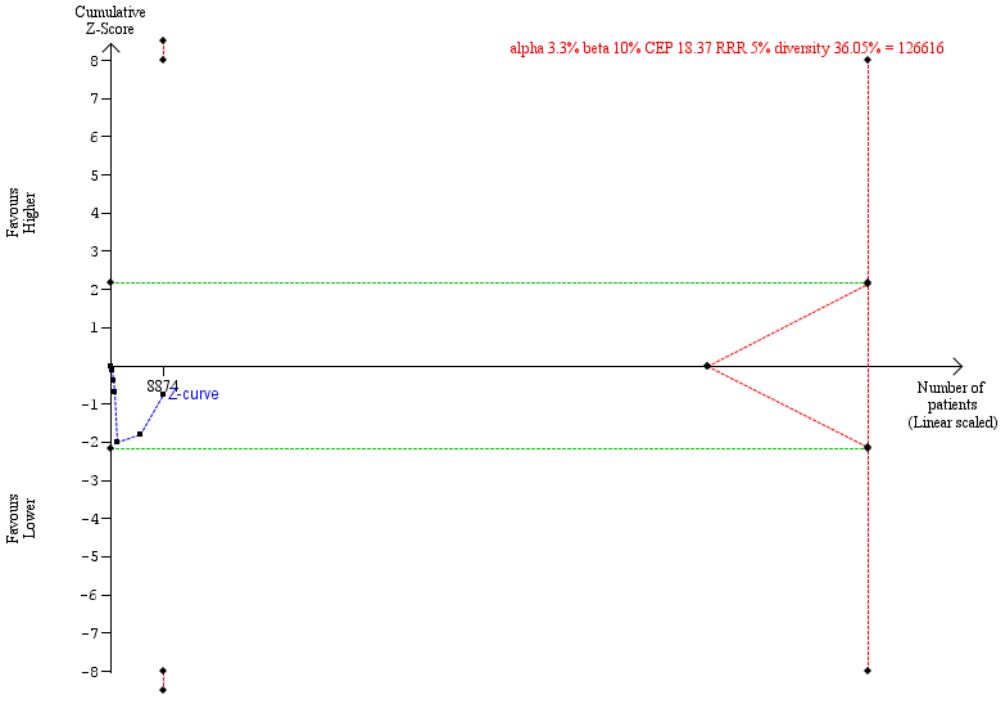
[a] defined as "subsequent revascularisation (ie percutaneous coronary intervention or coronary artery bypass grafting) of any lesion which occurs after the index admission and verified at six months follow-up"

[b] defined as bleeding occuring after the index admission when associated with death, hospital admission, blood transfusion, or intracranial hemorrhage

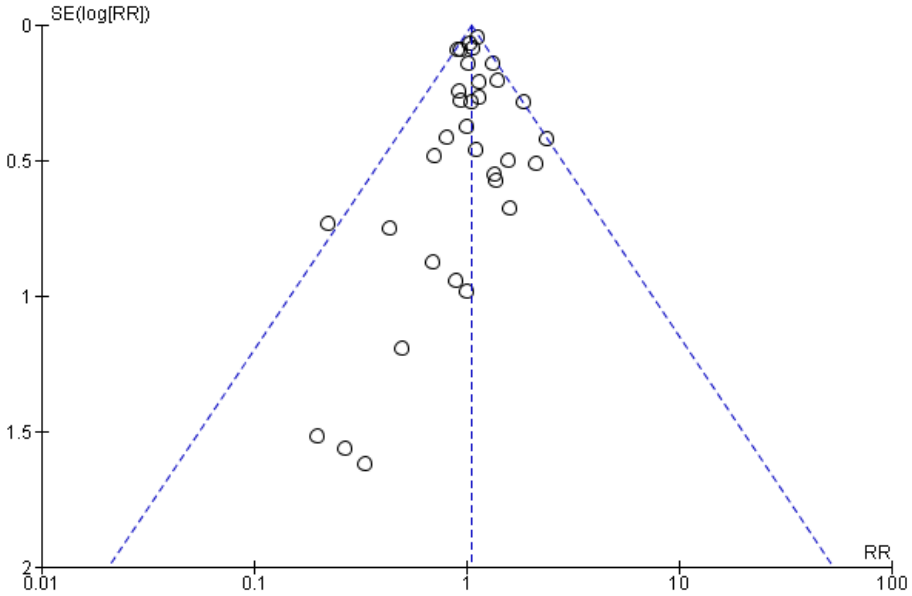
[c] defined as Evidence of inadequate tissue perfusion in the setting of adequate intravascular volume, characterised by persistent hypotension (systolic blood pressure ≤ 90 mm Hg), with or without altered mental status and peripheral hypoperfusion, requiring either pharmacologic or mechanical circulatory support



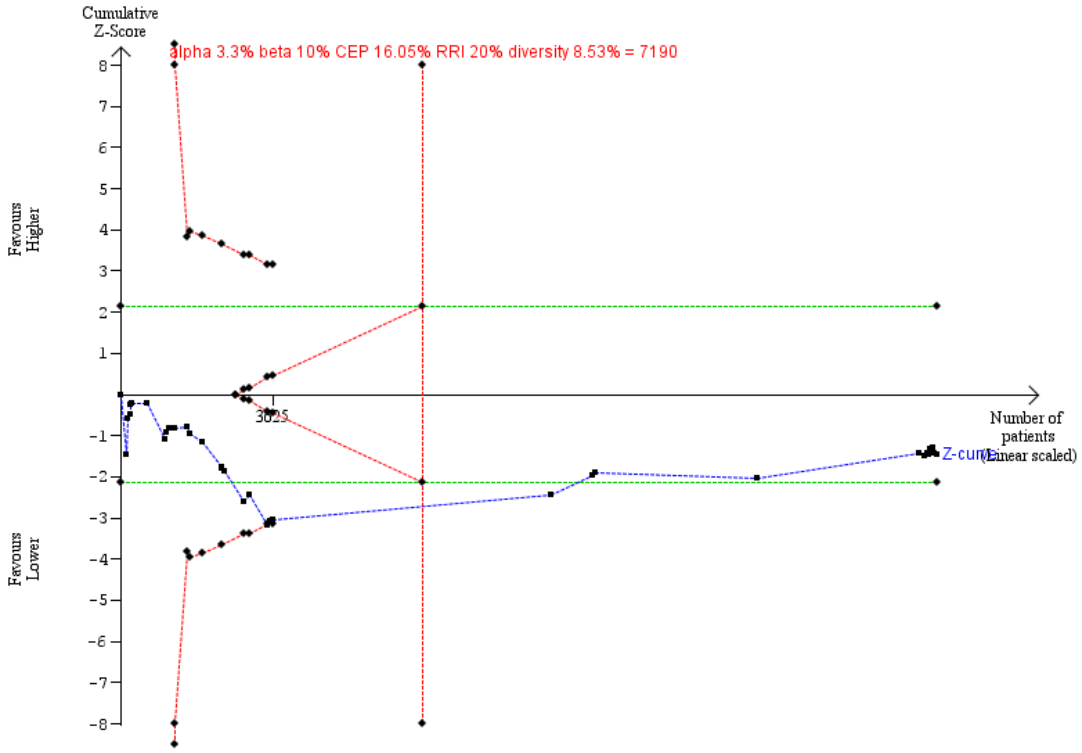
TSA on all trials on SAEs reported by trialists as the proportion of participants with at least one SAE for a 20% RRI.



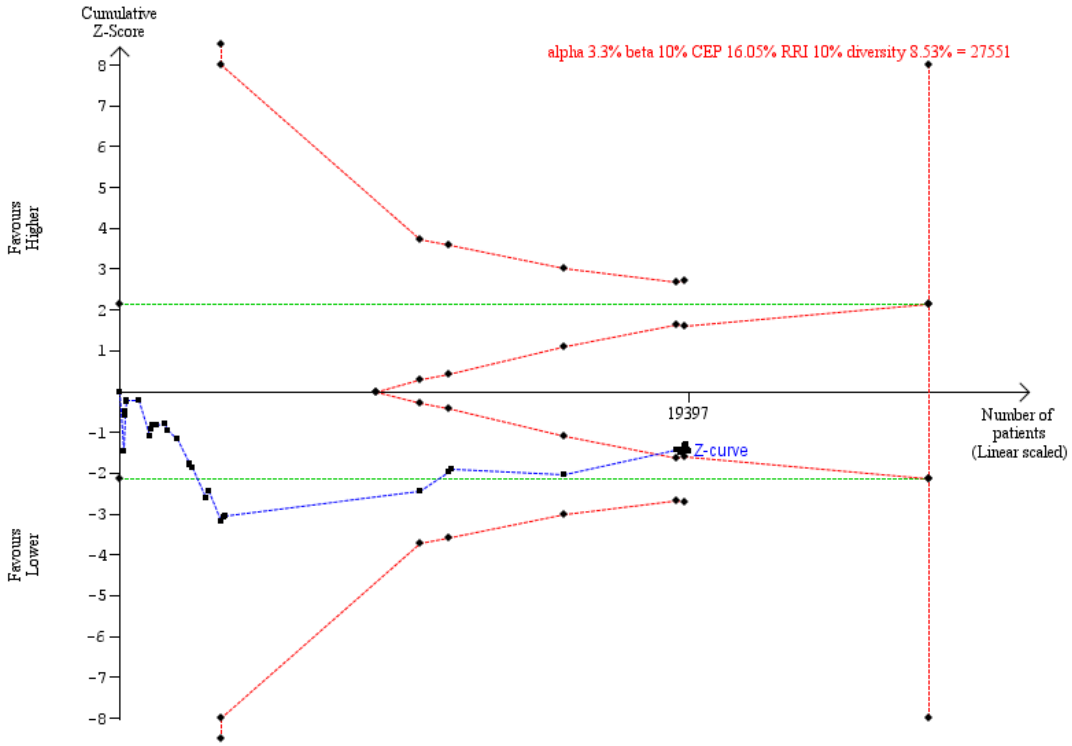
TSA on all trials on SAEs reported by trialists as the proportion of participants with at least one SAE for a 5% RRR.



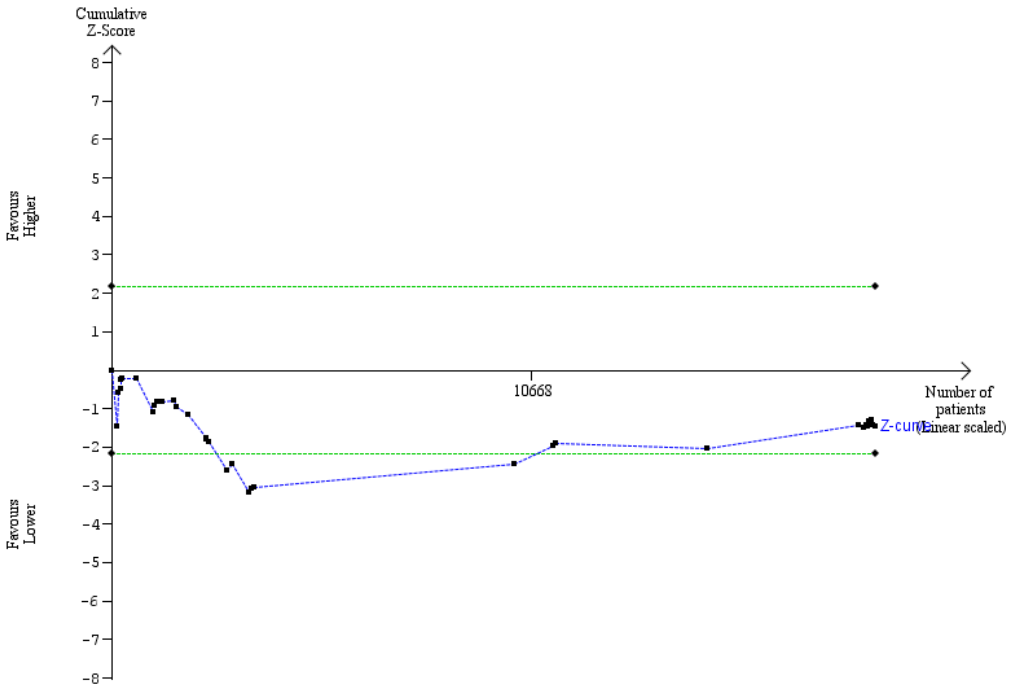
Funnel plot on the estimated highest reported proportion of specific SAEs in each trial. Harbord test P=0.986.



TSA on all trials on the estimated highest reported proportion of specific SAEs in each trial for a 20%.

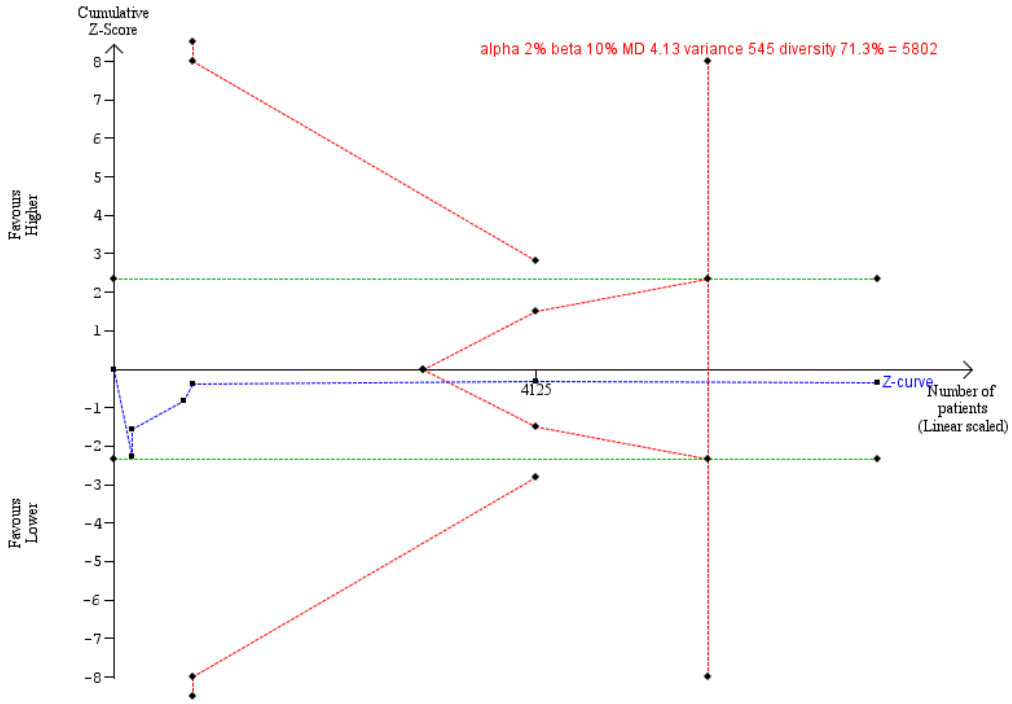


TSA on all trials on the estimated highest reported proportion of specific SAEs in each trial for a 10%.



TSA on all trials on the estimated highest reported proportion of specific SAEs in each trial for a 2% RRR. 2.93% of the required information (665,104) had been reached.

QUALITY OF LIFE



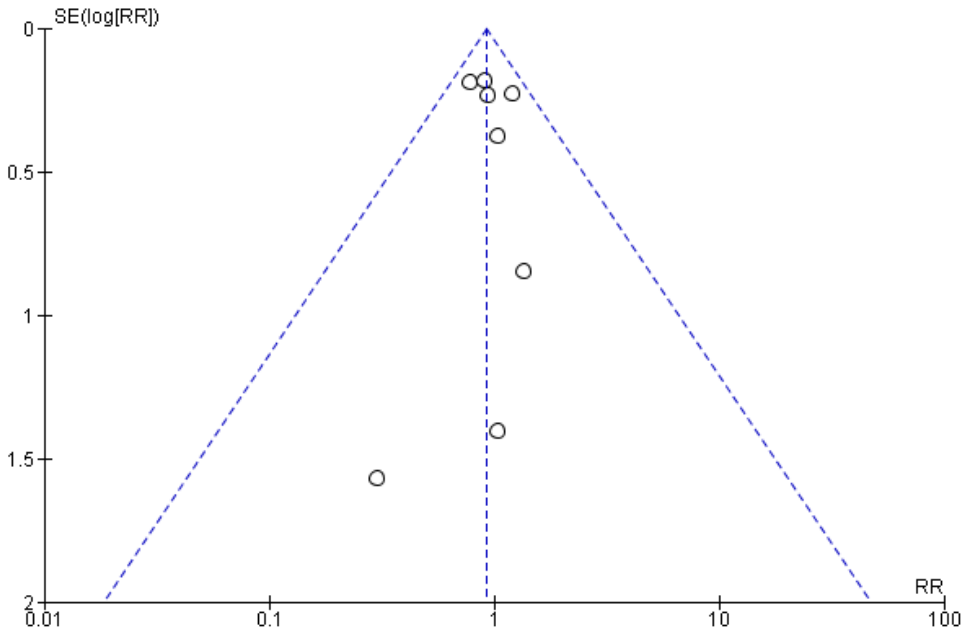
TSA on quality of life on all trials.

Types of lung injury, from each trial, included in meta-analysis on the estimated highest reported proportion of specific lung injury in each trial

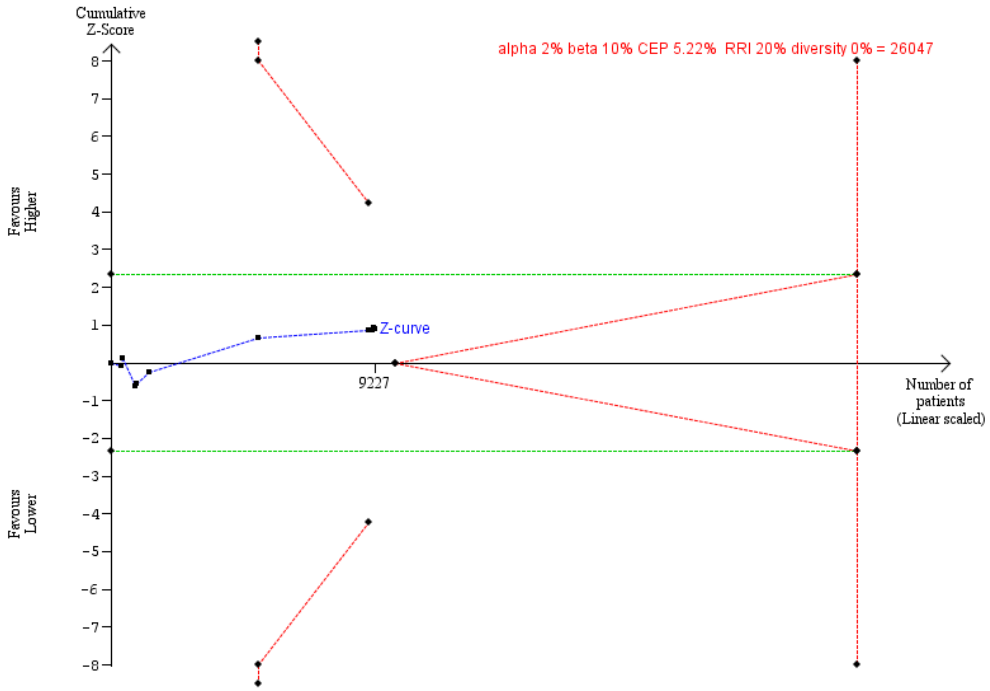
TRIAL ID	OUTCOME	Higher: events	Higher: analysed	Lower: events	Higher: analysed
Asfar 2017	Pneumonia	30	217	32	217
Baekgaard 2019	Pneumonia	3	19	2	17
Girardis 2016	Pneumonia	37	225	30	220
Jakkula 2018	ALI/ARDS	1	59	1	61
Lång 2018	Pneumonia	6	38	6	27
Meyhoff 2009	Pneumonia	13	188	13	195
Panwar 2016	ALI/ARDS	11	51	11	52
Roffe 2017A	Pneumonia	69	2567	44	1274
Roffe 2017B	Pneumonia	78	2561	43	1274
Young 2014	Pneumothorax	0	9	1	8

Types of lung injury, from each trial, included in meta-analysis on the estimated cumulated number of lung injury

TRIAL ID	OUTCOME	Higher: events	Higher: analysed	Lower: events	Higher: analysed
Asfar 2017	· Pneumonia · Pneumothorax	35	217	37	217
Baekgaard 2019	· ALI/ARDS · Pneumonia	5	19	3	17
Girardis 2016	· Pneumonia	37	225	30	220
Jakkula 2018	· ALI/ARDS · Severe hypercapnia or respiratory ascidosis	1	59	2	61
Lång 2018	· ALI/ARDS · Pneumonia	6	38	9	27
Meyhoff 2009	· Pneumonia · Pulmonary embolism	13	188	16	195
Panwar 2016	· ALI/ARDS	11	51	11	52
Roffe 2017A	· Pneumonia · Pulmonary embolism	87	2567	49	1274
Roffe 2017B	· Pneumonia · Pulmonary embolism	84	2561	47	1274
Young 2014	· Pneumothorax	0	9	1	8

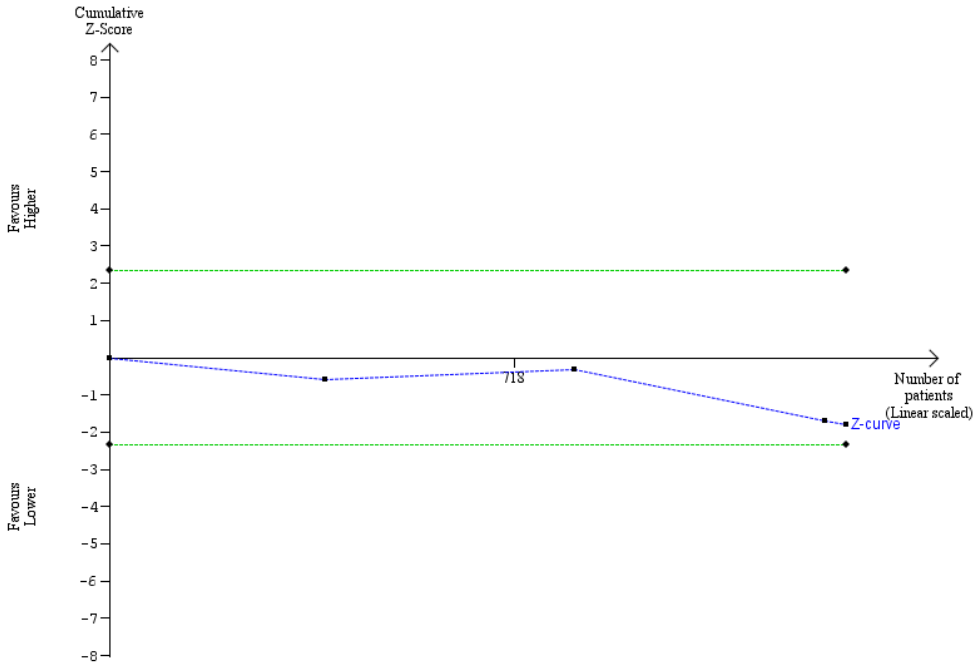


Funnel plot on the estimated highest reported proportion of specific lung injuries in each trial. Harbord test P=0.813.



TSA on all trials on the estimated highest reported proportion of specific lung injuries in each trial for a 20% RRI.

SEPSIS



TSA on sepsis on all trials for a 20% RRI. Only 2.89% of required information (45,241) size has been reached.

Types of cardiovascular events, from each trial, included in meta-analysis on the estimated highest reported proportion of cardiovascular events in each trial

TRIAL ID	OUTCOME	Higher: events	Higher: analysed	Lower: events	Higher: analysed
Asfar 2017	· Peripheral Arterial Thrombosis	12	217	13	217
Bray 2018	· Cardiac arrest incl. Re-arrest	1	24	0	37
Hofmann 2017	· Re-hosp. for heart failure	121	3311	114	3318
Jun 2019	· Myocardial infarction	2	29	9	29
Khoshnood 2018	· Cardiogenic shock	6	46	9	49
Meyhoff 2009	· Myocardial infarction	2	185	1	194
Panwar 2016	· Haemodynamic instability[a]	12	51	9	52
Ranchord 2012	· Myocardial infarction	1	68	0	68
Rawles 1976	· Ventricular tachycardia	11	80	5	77
Roffe 2017A	· Stroke	41	2567	21	1275
Roffe 2017B	· Stroke	36	2561	21	1274
Shi 2017	· Intracranial hemorrhage	0	9	0	9
Singhal 2013	· Brain Hemorrhage	4	43	0	41
Stub 2015	· Cardiogenic shock[b]	20	218	20	223
Thomas 2019	· Cardiac arrest incl. re-arrest	8	17	3	18
Zughaft 2013	· Myocardial infarction	0	154	0	146

[a] defined as cardiac arrest or addition of two or more new vasopressor/inotrope agents in a day

[b] defined as Evidence of inadequate tissue perfusion in the setting of adequate intravascular volume, characterised by persistent hypotension (systolic blood pressure ≤ 90 mm Hg), with or without altered mental status and peripheral hypoperfusion, requiring either pharmacologic or mechanical circulatory support

Types of cardiovascular events, from each trial, included in meta-analysis on the estimated cumulated number of cardiovascular events

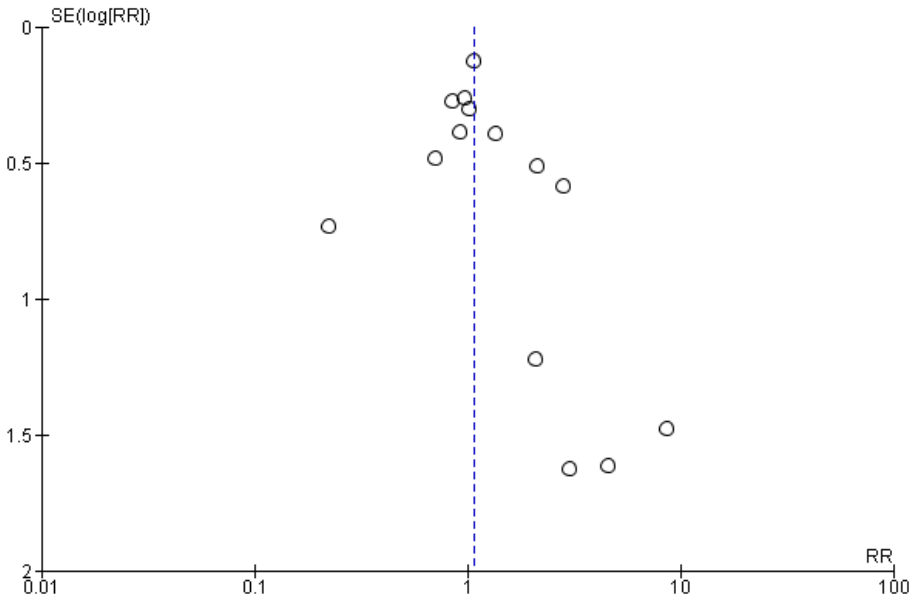
TRIAL ID	OUTCOME	Higher: events	Higher: analysed	Lower: events	Higher: analysed
Asfar 2017	· Peripheral Arterial Thrombosis · Ventricular arrhythmias	12	217	14	217
Bray 2018	· Cardiac arrest incl. re-arrest	1	24	0	37
Hofmann 2017	· Myocardial Infarction · Atrioventricular block, second or third degree · Cardiogenic shock · Cardiac arrest incl. re-arrest · Re-hosp for heart failre	340	3311	318	3318
Jun 2019	· Myocardial Infarction	2	29	9	29
Khoshnood 2018	· Myocardial Infarction · Cardiogenic shock · Cardiac arrest incl. re-arrest · Heart failure	9	46	13	49
Meyhoff 2009	· Myocardial Infarction · Stroke · Pulmonary Embolism	4	185	5	194
Panwar 2016	· Haemodynamic instability[a]	12	51	9	52
Ranchord 2012	· Myocardial Infarction	1	68	0	68
Rawles 1976	· Atrioventricular block, second or third degree · Ventricular tachycardia · Ventricular fibrillation	14	80	11	77
Roffe 2017A	· Myocardial Infarction · Deep Vein Thrombosis · Pulmonary Embolism · intracranial hemorrhage · Transient ischemic attack	88	2567	35	1275

Roffe 2017B	· Myocardial Infarction · Deep Vein Thrombosis · Pulmonary Embolism · intracranial hemorrhage · Transient ischemic attack	65	2567	33	1274
Shi 2017	· Intracranial hemorrhage · Subarachnoid hemorrhage	0	9	0	9
Singhal 2013	· Brain Hemorrhage	4	43	0	43
Stub 2015	· Myocardial Infarction · Stroke · Repeated re-vascularisation[b] · Cardiogenic shock[c] · Coronary artery bypass grafting	69	218	56	223
Thomas 2019	· Cardiac arrest incl. re-arrest	8	17	3	18
Zughaft 2013	· Myocardial Infarction · Stroke	0	154	0	146

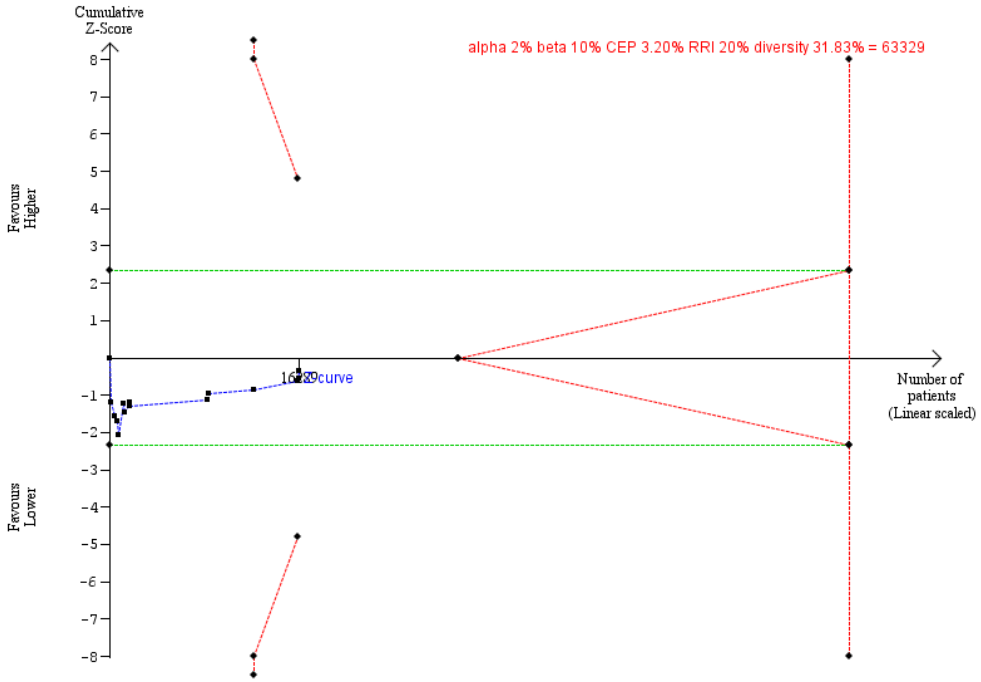
[a] defined as cardiac arrest or addition of two or more new vasopressor/inotrope agents in a day

[b] defined as "subsequent revascularisation (ie percutaneous coronary intervention or coronary artery bypass grafting) of any lesion which occurs after the index admission and verified at six months follow-up"

[c] defined as Evidence of inadequate tissue perfusion in the setting of adequate intravascular volume, characterised by persistent hypotension (systolic blood pressure ≤ 90 mm Hg), with or without altered mental status and peripheral hypoperfusion, requiring either pharmacologic or mechanical circulatory support



Funnel plot on the estimated highest reported proportion of specific cardiovascular events in each trial. Harbord test P=0.190.



TSA on all trials on the estimated highest reported proportion of specific cardiovascular events in each trial for a 20% RRI

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

1. In the TSA, we used a power of 90%, not 80% as reported in the protocol [1], as a meta-analysis should use higher (or same) power as its included trials in order to communicate the best available evidence.
2. We conducted additional TSA's for a 15% and 10% RRR/RRI on mortality and SAE.
3. In our protocol we stated that we would search the Allied and Complementary Medicine Database (AMED) for eligible trials. We had no access to AMED, and so this search was not conducted.
4. We performed two analyses on the effects of higher versus lower levels of oxygen supplementation on the composite outcomes SAEs, lung injuries and cardiovascular events.
5. We conducted a post-hoc analysis on the effect of supplemental oxygen versus no supplemental oxygen on mortality.
6. We post-hoc decided to accommodate the possible challenges of blinding outcome assessors and presented trials as overall low risk of bias when blinding was not maintained or not reported adequately.
7. Title was changed from *'Oxygen supplementation for critically ill patients'* to '*Higher versus lower levels of oxygen supplementation in critically ill adults*'

REFERENCES

1. Barbateskovic M, Schjørring OL, Jakobsen JC, Meyhoff CS, Rasmussen BS, Perner A, et al (2018) Oxygen supplementation for critically ill patients - a protocol for a systematic review. Acta Anaesthesiologica Scandinavica 62 (7):1020-1030
2. Ali K, Warusevitane A, Lally F, Sim J, Sills S, Pountain S, et al (2013) The stroke oxygen pilot study: a randomized controlled trial of the effects of routine oxygen supplementation early after acute stroke - effect on key outcomes at six months. PLOS ONE 8 (6):e59274-e59274
3. Asfar P, Schortgen F, Boisramé-Helms J, Charpentier J, Guérot E, Megarbane B, et al (2017) Hyperoxia and hypertonic saline in patients with septic shock (HYPERS2S): a two-by-two factorial, multicentre, randomised, clinical trial. Lancet Respiratory Medicine 5 (3):180-190
4. Austin MA, Wills KE, Blizzard L, Walters EH, Wood-Baker R (2010) Effect of high flow oxygen on mortality in chronic obstructive pulmonary disease patients in prehospital setting: randomised controlled trial. BMJ 341:c5462-c5462
5. Baekgaard JS, Isbye D, Ottosen CI, Larsen MH, Andersen JH, Rasmussen LS, Steinmetz J (2019) Restrictive vs liberal oxygen for trauma patients-the TRAUMOX1 pilot randomised clinical trial. Acta Anaesthesiol Scand 63 (7):947-955. doi:10.1111/aas.13362
6. Bardsley G, Pilcher J, McKinstry S, Shirtcliffe P, Berry J, Fingleton J, Weatherall M, Beasley R (2018) Oxygen versus air-driven nebulisers for exacerbations of chronic obstructive pulmonary disease: a randomised controlled trial. BMC Pulm Med 18 (1):157. doi:10.1186/s12890-018-0720-7
7. Bickel A, Gurevits M, Vamos R, Ivry S, Eitan A (2011) Perioperative hyperoxygenation and wound site infection following surgery for acute appendicitis: a randomized, prospective, controlled trial. Archives of Surgery 146 (4):464-470
8. Bray JE, Hein C, Smith K, Stephenson M, Grantham H, Finn J (2018) Oxygen titration after resuscitation from out-of-hospital cardiac arrest: a multi-centre, randomised controlled pilot study (the EXACT pilot trial). Resuscitation 128:211-215
9. Butler CM, Ham RO, Lafferty K, Cotton LT, Roberts VC (1987) The effect of adjuvant oxygen therapy on transcutaneous pO2 and healing in the below-knee amputee. Prosthet Orthot Int 11 (1):10-16. doi:10.3109/03093648709079373
10. Girardis M, Busani S, Damiani E, Donati A, Rinaldi L, Marudi A, et al (2016) Effect of conservative vs conventional oxygen therapy on mortality among patients in an intensive care unit: the oxygen-ICU randomized clinical trial. JAMA 316 (15):1583-1589
11. Gomersall CD, Joynt GM, Freebairn RC, Lai CK, Oh TE (2002) Oxygen therapy for hypercapnic patients with chronic obstructive pulmonary disease and acute respiratory failure: a randomized, controlled pilot study. Critical Care Medicine 30 (1):113-116
12. Heidari F, Rahzani K, Iranpoor D, Rezaeed K (2017) The effect of oxygen on the outcomes of non-ST-segment elevation acute coronary syndromes. IJC Metabolic & Endocrine 14:67-71
13. Hofmann R, James SK, Jernberg T, Lindahl B, Erlinge D, Witt N, et al (2017) Oxygen therapy in suspected acute myocardial infarction. New England Journal of Medicine 377 (13):1240-1249
14. Jernberg T, Lindahl B, Alfredsson J, Berglund E, Bergstrom O, Engstrom A, Erlinge D, Herlitz J, Jumatate R, Kellerth T, Lauermann J, Lindmark K, Lingman M, Ljung L, Nilsson C, Omerovic E, Pernow J, Ravn-Fischer A, Sparv D, Yndigegn T, Ostlund O, James SK, Hofmann R (2018) Long-Term Effects of Oxygen Therapy on Death or Hospitalization for Heart Failure in Patients With Suspected Acute Myocardial Infarction. Circulation 138 (24):2754-2762. doi:10.1161/circulationaha.118.036220
15. Huynh Ky M, Bouchard PA, Morin J, L'Her E, Sarrazin JF, Lellouche F (2017) Closed-loop adjustment of oxygen flowrate with FreeO2 in patients with acute coronary syndrome: comparison of automated titration with FreeO2 (set at two SpO2 target) and of manual titration. A randomized controlled study. Annals of Intensive Care 7 (Suppl 1):O59-O59
16. Mackle D, Bellomo R, Bailey M, Beasley R, Deane A, Eastwood G, Finfer S, Freebairn R, King V, Linke N, Litton E, McArthur C, McGuinness S, Panwar R, Young P (2019) Conservative Oxygen Therapy during Mechanical Ventilation in the ICU. N Engl J Med. doi:10.1056/NEJMoa1903297
17. Ishii K, Morimatsu H, Hyodo T, Ono K, Hidaka H, Koyama Y, et al (2018) Relationship between inspired oxygen concentration and atelectasis formation after extubation. Critical Care Medicine 46 (1 Suppl 1):533-533

18. Jakkula P, Reinikainen M, Hästbacka J, Loisa P, Tiainen M, Pettilä V, et al (2018) Targeting two different levels of both arterial carbon dioxide and arterial oxygen after cardiac arrest and resuscitation: a randomised pilot trial. Intensive Care Medicine 44 (12):2112-2121
19. Jun J, Sun L, Wang Y, Liu F, Yang G, Han B (2019) Invasive mechanical ventilation with high concentration oxygen therapy for AECOPD patients with acute myocardial infarction. Chest 156 (4):A958
20. Khoshnood A, Akbarzadeh M, Roijer A, Meurling C, Carlsson M, Bhiladvala P, et al (2017) Effects of oxygen therapy on wall-motion score index in patients with ST elevation myocardial infarction - the randomized SOCCER trial. Echocardiography 34 (8):1130-1137
21. Khoshnood A, Carlsson M, Akbarzadeh M, Bhiladvala P, Roijer A, Nordlund D, et al (2018) Effect of oxygen therapy on myocardial salvage in ST elevation myocardial infarction: the randomized SOCCER trial. European Journal of Emergency Medicine 25 (2):78-84
22. Kuisma M, Boyd J, Voipio V, Alaspää A, Roine RO, Rosenberg P (2006) Comparison of 30 and the 100% inspired oxygen concentrations during early post-resuscitation period: a randomised controlled pilot study. Resuscitation 69 (2):199-206
23. Lång M, Skrifvars MB, Siironen J, Tanskanen P, Ala-Peijari M, Koivisto T, et al (2018) A pilot study of hyperoxemia on neurological injury, inflammation and oxidative stress. Acta Anaesthesiologica Scandinavica 62 (6):801-810
24. Mazdeh M, Taher A, Torabian S, Seifirad S (2015) Effects of normobaric hyperoxia in severe acute stroke: a randomized controlled clinical trial study. Acta Medica Iranica 53 (11):676-680
25. Meyhoff CS, Wetterslev J, Jorgensen LN, Henneberg SW, Høgdall C, Lundvall L, et al (2009) Effect of high perioperative oxygen fraction on surgical site infection and pulmonary complications after abdominal surgery: the PROXI randomized clinical trial. JAMA 302 (14):1543-1550
26. NCT02378545. Trial of Hyperoxic O2 Therapy vs. Normoxic O2 Therapy in Sepsis (HO2T or NO2T) 2015. https://clinicaltrials.gov/ct2/show/NCT02378545 Accessed 4 Devember 2019.
27. NCT02687217. Effect of Peri-operative Supplemental Oxygen in Wound Infection After Appendectomy 2016. https://clinicaltrials.gov/ct2/show/NCT02687217 Accessed 4 December 2019. .
28. Padma MV, Bhasin A, Bhatia R, Garg A, Singh MB, Tripathi M, et al (2010) Normobaric oxygen therapy in acute ischemic stroke: a pilot study in Indian patients. Annals of Indian Academy of Neurology 13 (4):284-288
29. Panwar R, Hardie M, Bellomo R, Barrot L, Eastwood GM, Young PJ, et al (2016) Conservative versus liberal oxygenation targets for mechanically ventilated patients. A pilot multicenter randomized controlled trial. American Journal of Respiratory and Critical Care Medicine 193 (1):43-51
30. Perrin K, Wijesinghe M, Healy B, Wadsworth K, Bowditch R, Bibby S, et al (2011) Randomised controlled trial of high concentration versus titrated oxygen therapy in severe exacerbations of asthma. Thorax 66 (11):937-941
31. Ranchord AM, Argyle R, Beynon R, Perrin K, Sharma V, Weatherall M, et al (2012) High-concentration versus titrated oxygen therapy in ST elevation myocardial infarction: a pilot randomized controlled trial. American Heart Journal 163 (2):168-175
32. Rawles JM, Kenmure AC (1976) Controlled trial of oxygen in uncomplicated myocardial infarction. BMJ 1 (6018):1121-1123
33. Rodrigo GJ, Rodriquez Verde M, Peregalli V, Rodrigo C (2003) Effects of short-term 28% and 100% oxygen on PaCO2 and peak expiratory flow rate in acute asthma: a randomized trial. Chest 124 (4):1312-1317
34. Vianna JR, Pires Di Lorenzo VA, Simoes MM, Jamami M (2017) Comparing the Effects of Two Different Levels of Hyperoxygenation on Gas Exchange During Open Endotracheal Suctioning: A Randomized Crossover Study. Respir Care 62 (1):92-101. doi:10.4187/respcare.04665
35. Roffe C, Sills S, Pountain SJ, Allen M (2010) A randomized controlled trial of the effect of fixed-dose routine nocturnal oxygen supplementation on oxygen saturation in patients with acute stroke. Journal of Stroke and Cerebrovascular Diseases 19 (1):29-35
36. Roffe C, Nevatte T, Sim J, Bishop J, Ives N, Ferdinand P, et al (2017) Effect of routine low-dose oxygen supplementation on death and disability in adults with acute stroke: the stroke oxygen study randomized clinical trial. JAMA 318 (12):1125-1135
37. Sepehrvand N, Alemayehu W, Rowe BH, McAlister FA, van Diepen S, Stickland M, Ezekowitz JA (2019) High vs. low oxygen therapy in patients with acute heart failure: HiLo-HF pilot trial. ESC Heart Fail 6 (4):667-677. doi:10.1002/ehf2.12448
38. Shi S, Qi Z, Ma Q, Pan R, Timmins GS, Zhao Y, Shi W, Zhang Y, Ji X, Liu KJ (2017) Normobaric Hyperoxia Reduces Blood Occludin Fragments in Rats and Patients With Acute Ischemic Stroke. Stroke 48 (10):2848-2854. doi:10.1161/strokeaha.117.017713

39. Sills S, Halim M, Roffe C (2003) A pilot study of routine nocturnal oxygen supplementation in patients with acute stroke. Age and Ageing 32 (Suppl 2):ii41-ii41
40. Singhal AB, Benner T, Roccatagliata L, Koroshetz WJ, Schaefer PW, Lo EH, et al (2005) A pilot study of normobaric oxygen therapy in acute ischemic stroke. Stroke 36 (4):797-802
41. Singhal A (2013) A phase IIb clinical trial of normobaric oxygen therapy (NBO) in acute ischemic stroke (AIS). Neurology 80 (Suppl 7):S02.001
42. Stewart R (2019) Design and conduct of the New Zealand oxygen therapy in acute coronary syndromes trial. Heart Lung and Circulation 28:S16
43. Stub D, Smith K, Bernard S, Nehme Z, Stephenson M, Bray JE, et al (2015) Air versus oxygen in S-segment-elevation myocardial infarction. Circulation 131 (24):2143-2150
44. Taher A, Pilehvari Z, Poorolajal J, Aghajanloo M (2016) Effects of normobaric hyperoxia in traumatic brain injury: a randomized controlled clinical trial. Trauma Monthly 21 (1):e26772-e26772
45. Thomas M, Voss S, Benger J, Kirby K, Nolan JP (2019) Cluster randomised comparison of the effectiveness of 100% oxygen versus titrated oxygen in patients with a sustained return of spontaneous circulation following out of hospital cardiac arrest: a feasibility study. PROXY: post ROSC OXYgenation study. BMC Emerg Med 19 (1):16. doi:10.1186/s12873-018-0214-1
46. Ukholkina GB, Kostianov IIu, Kuchkina NV, Grendo EP, Gofman IaB (2005) Effect of oxygenotherapy used in combination with reperfusion in patients with acute myocardial infarction. Kardiologiia 45 (5):59-59
47. Wijesinghe M, Perrin K, Healy B, Weatherall M, Beasley R (2012) Randomized controlled trial of high concentration oxygen in suspected community-acquired pneumonia. J R Soc Med 105 (5):208-216. doi:10.1258/jrsm.2012.110084
48. Wilson AT, Channer KS (1997) Hypoxaemia and supplemental oxygen therapy in the first 24 hours after myocardial infarction: the role of pulse oximetry. J R Coll Physicians Lond 31 (6):657-661
49. Wu WW, Hong HH, Shao XP, Rui L, Jing M, Lu GD (2014) Effect of oxygen-driven nebulization at different oxygen flows in acute exacerbation of chronic obstructive pulmonary disease patients. Am J Med Sci 347 (5):343-346. doi:10.1097/MAJ.0b013e3182937766
50. Young P, Bailey M, Bellomo R, Bernard S, Dicker B, Freebairn R, et al (2014) HyperOxic Therapy OR NormOxic Therapy after out-of-hospital cardiac arrest (HOT OR NOT): a randomised controlled feasibility trial. Resuscitation 85 (12):1686-1691
51. Zughaft D, Bhiladvala P, Van Dijkman A, Harnek J, Madsen Hardig B, Bjork J (2013) The analgesic effect of oxygen during percutaneous coronary intervention (the OXYPAIN Trial). Acute Cardiac Care 15 (3):63-68

Clinical Heterogeneity In Meta-analysis Score (CHIMS) guidance manual

This supplementary has been provided by the authors to give readers additional information about their work.

Supplement to: Barbateskovic M, Koster TM, Eck R, et al. CHIMS: Clinical Heterogeneity In Meta-Analysis Score – a new tool to assess and quantify clinical heterogeneity in meta-analyses of interventions.

Clinical Heterogeneity In Meta-analysis Score (CHIMS) guidance manual

Explanation for the use of CHIMS

To guide evaluators of Clinical Heterogeneity In Meta-analysis Score (CHIMS) we have provided somewhat arbitrary thresholds for the scores 0, 1, and 2 and this guidance manual for using the CHIMS should help getting higher agreements between independent evaluators. We hope everyone can agree that e.g. more than 30% relative difference between different trials dose of a drug intervention or e.g. that more than 30% relative difference in risk of severe disease (or severity score) are substantial differences. However, differences between 20% and 30% are probably not substantial but may influence results, and less than 20% is probably less important.

If difference between trials for a specific item is impossible to detect or reject, we suggest that the meta-analysis score 1 on the given item corresponding to unclear.

Assessment of clinical heterogeneity with CHIMS

CHIMS is measured on a scale including the following 4 domains with overall 11 selected items covering the most important domains and items describing clinical heterogeneity:

1. **Domain: population heterogeneity (4 items)**: disease severity, comorbidities, age, and sex.
2. **Domain: setting heterogeneity (1 item)**: period where the trial was performed/reported whether it was performed in a developed (D) or in a developing country (DC), or e.g. performed in a high-dependency unit compared to an intensive care unit (ICU) or a high-dependency unit compared to standard unit, etc.
3. **Domain: intervention heterogeneity (4 items)**: Intervention intensity (dose, frequency, duration, device, cut-off values), timing of intervention (number of times per time unit, continuous), heterogeneity of control-interventions in included trials, distribution of co-interventions in randomised groups.
4. **Domain: outcome heterogeneity (2 items)**: Definition of outcome, timing of outcome assessment.

Overall CHIMS score: The 11 items in an unweighted CHIMS are scored 0, 1, or 2 corresponding to low=0, moderate or unclear=1, or high clinical heterogeneity=2, with a total score of 0 to 22 with equal weight assigned to each item.

The Clinical Heterogeneity In Meta-analysis Score (CHIMS) tool

Domains	Items	Score	Explanation of score for extreme differences between trials in a meta-analysis
Setting heterogeneity	**1. Years reported (A), performed in developed vs developing country (B), unit type (C)**	0 1 2	No differences: A) years reported differ < 15, B) No developed vs developing countries, AND C) slight variations in the unit or facility type and there is low risk of affecting other fields of heterogeneity Slight variation (at least one of A-C involved): A) years reported differ ≥ 15, OR B) developed vs developing countries, OR treating units not similar, OR C) if there are slight variations in the unit or facility type but there is risk of affecting other fields of heterogeneity Considerable variation (all of A-C involved): A) years reported differ ≥ 15, AND B) developed vs developing countries, AND C) treating units not similar (all of A-C involved), OR if the trials in the opinion of the assessor differs markedly in setting heterogeneity
Population heterogeneity	**2. Age**	0 1 2	Mean/median age ≤ 10 years difference 11 to 20 years difference in mean/median age Mean/median age > 20 years difference
	3. Sex	0 1 2	% women ≤ 20 % absolute difference between trials 21% to 30% absolute difference of % women between trials More than > 30% absolute difference between trials
	4. Participant inclusion criteria and baseline disease severity	0 1 2	Different trials include patients that are equally ill or the difference in risk or score for disease severity of patients ≤ 20% Condition/patient population differs slightly with 50% or more overlap of types of participants and/or the difference in risk or score for disease severity of patients is 21% to 30% Condition/patient population differs considerable and/or the difference in risk or score for disease severity of patients > 30% Use relative difference when inclusion criteria are assessed (disease severity scores).

	5. Co-morbidities	0 1 2 Use absolute difference when comparing important comorbidities.	Difference in frequency of important comorbidities ≤ 20% or no co-morbidities are reported in the included trials and differences in co-morbidities are assumed absent Slight differences in important co-morbidities, between 21% and 30%, or no co-morbidities are reported in the trials, but differences in co-morbidities are assumed Differences in frequency of important comorbidities > 30% or highly likely variations in co-morbidities
Intervention heterogeneity	**6. Intensity, strengths, or duration of intervention**	0 1 2 Use relative differences when assessing intensity, strengths, duration.	Little variation: differences in dose, strengths, devices, cut-offs, or duration of interventions ≤20% Slight variation: 21% to 30% differences in dose, strengths, devices, cut-offs, or duration intervention, or if dose, strength, cut-offs or duration of intervention cannot be assessed from the information in the included trials Considerable variation: if different types of interventions are used, or different doses, strengths, devices, cut-offs, or duration of intervention > 30%
	7. Timing	0 1 2	Criteria for starting the intervention are similar, or relative differences of timing of intervention differs ≤ 20% Criteria for starting the intervention differ slightly, or the relative timing difference is 21% to 30% Criteria for starting the intervention differ, or relative timing difference exceeds > 30%
	8. Control intervention	0 1 2	All control interventions are the same Control interventions include placebo AND no intervention, assess as item 6 if an active intervention is used Including trials with different active control interventions OR trials with active and placebo/no intervention

	9. Co-interventions	0 1 2	No apparent differences in co-interventions, OR standard care is not described or assumed to be the same, OR equally applied in groups, OR different co-interventions are used but the effects of the co-interventions are assumed to be small Slight variation in co-interventions or the same cointerventions are used with slight variation (< 30% difference in e.g. doses or numbers of participants using the co-intervention) Considerable differences if it is assumed that the cointervention is not usual care, or differences in use or e.g. doses of cointerventions > 30% Use absolute difference when assessing co-interventions.
Outcome heterogeneity	**10. Definition of the outcome in the meta-analysis**	0 1 2	Same definition of outcome Slight variations in definition of outcome Considerable variations in definition of outcome
	11. Timing of outcome measurement	0 1 2	Less than one month between follow-up of outcome More than one but less than or equal to 3 months between follow-ups More than 3 months between follow-up of outcome

Guidance manual for assessing clinical heterogeneity using CHIMS

CHIMS has been developed to detect and quantify clinical heterogeneity in meta-analyses. When a difference has been identified between two trials for an item resulting in the score 2, it is not necessary to investigate the remaining trials, but one may move on to the next item.

In general, if differences between trials for a specific item are impossible to detect or reject, we suggest that the meta-analysis score 1 on the given item.

SETTING HETEROGENEITY

This domain assesses the difference in setting of the included trials.

Signalling question 1: Is there a difference in setting between the trials, such as the years the trials were reported?

Score 0: If differences in time period of conduct is less than 15 years.

Score 1: If the conduct of the studies differs more than 15 years.

If the conduct of the studies is not reported, use the publication year.

Signalling question 2: Was the study conducted in developed vs developing countries? Or in other words, is it assumable the level of 'standard care' provided to the patients is the same in the trials?

- Score 0: If it is assumable that standard care in the included studies is the same.
- Score 1: If it is not assumable that the standard care is the same.

Signalling question 3: Were the studies conducted in the same type of unit/facility?

- Score 0: If there are slight variations in the unit or facility type and there is low risk of affecting other fields of heterogeneity.
- Score 1: If there are slight variations in the unit or facility type, but there is risk of affecting other fields of heterogeneity.

Overall score of the domain Setting heterogeneity:

Score 0: If no signalling questions have scored points and the overall setting heterogeneity is assumedly low.

Score 1: If at least one of the three signalling questions have scored points and the overall setting heterogeneity is assumedly slight.

Score 2: If all three signalling questions have scored points and the overall setting heterogeneity is assumedly high.

Example 1. Meta-analysis on the use of dopamine for blood pressure regulation in septic shock patients.

Trial 1 is conducted in 1990 in Denmark.

Trial 2 is conducted in 2017 in Denmark.

The conduct of the trials differs more than 15 years. The standard of care has changed in these years. This meta-analysis scores 2 points.

Example 2. Meta-analysis on the use of dopamine for blood pressure regulation in septic shock patients.

Trial 1 is conducted in 2016 in Denmark.

Trial 2 is conducted in 2017 in Uganda.

The standard care will probably differ. This Meta-analysis will score 1 point. However, if we change Uganda to a large city in India, the standard care may probably be the same.

Example 3. Meta-analysis on the use of dopamine for blood pressure regulation in septic shock patients.

Trial 1 is conducted in 2016 in the mixed ICU of university hospital with 30 ICU beds in Denmark.

Trial 2 is conducted in 2017 in a medical ICU of a small-town hospital with 3 ICU beds in the Netherlands.

The unit type is the same. However, probably the patients admitted to the ICU and the standard care may differ between the two groups, therefore this item scores 1 points.

POPULATION HETEROGENEITY

This domain is defined by multiple subdomains namely; 1. age, 2. sex, 3. inclusion criteria or baseline disease severity, 4. and co-morbidities.

1. AGE

Assess the difference in mean age between trials. If mean age is only given for each of the groups in a single trial, calculate mean age of the included populations.

- Score 0: Mean/median age ≤ 10 years difference
- Score 1: 10-20 years difference in mean/median age
- Score 2: More than 20 difference in mean/median age

2. SEX

Assess the difference in sex between trials. If sex is only reported for each of the groups in a single trial, then calculate percentage of total males/females.

- Score 0: % women ≤ 20 % absolute difference between trials
- Score 1: 20% <% women <30 % absolute difference between trials
- Score 2: % women ≥ 30% absolute difference between trials

Example 1. Trial 1 included 20% females and trial 2 included 45% females.

The absolute difference is 25%, therefore score a 1.

3. PARTICIPANT INCLUSION CRITERIA AND BASELINE DISEASE SEVERITY

This subdomain assesses possible differences in patient diseases and the severity of these diseases.

Signalling question 1: Do the trials include the same type of participants or do the trials have similar inclusion criteria?

- Score 0: If the inclusion criteria of the trials describe the same types of participants.
- Score 1: If inclusion criteria differ slightly with 50% or more overlap of types of participants
- Score 2: If the trials include different types of patients.

Signalling question 2: How do the patients with the same inclusion criteria in the trials compare to each other? Are the patients' conditions similar? Are the patients in the trials equally 'ill'?

- Score 0: If the trials include patients that are equally ill.
- Score 1: If the trials include slightly different patients based on their illness, between 20 and 30% difference in score for disease severity.
- Score 2: If the difference exceeds 30%.

Example 1. Meta-analysis on the use of desmopressin for nocturia.

Trial 1 includes men with voiding > 2 episodes/night.

Trial 2 includes men aged 40 to 65 years with lower urinary tract symptoms, International Prostate Symptom Score ≥ 13, persistent nocturia (≥ 2 episodes/night), nocturia index score ≥ 1 despite use of alpha-blocker treatment for ≥ 8 weeks, and nocturnal polyuria defined as nocturnal polyuria index > 33%.

The inclusion criteria differ between the two studies. The criteria of trial 1 are less strict than in trial 2. The participants included in the studies will assumedly be different. This item scores 1 point.

Example 2. Meta-analysis on the use of dopamine for blood pressure regulation in ICU patients.

Trial 1 includes all patient admitted to the ICU with septic shock. Mean APACHE II (ICU mortality score) score of 17 points.

Trial 2 includes all patient admitted to the ICU with septic shock. Mean APACHE II score of 15 points.

The patient inclusion criteria are the same. However, the disease severity differs. The difference in severity in this meta-analysis is 12%, thus this meta-analysis will score 1 point.

Example 3. Meta-analysis on the use of honey as intervention in wound treatment.

Trial 1 includes patients with burn injury.

Trial 2 includes patients post caesarean section or hysterectomy.

These patient categories differ. Thus, this meta-analysis scores 2 points on this item.

4. CO-MORBIDITIES

Co-morbidities are defined as the characteristics on diseases of the patients besides the inclusion criteria of the trials that is included in the meta-analysis.

Signalling question 1: Are co-morbidities reported in the trials?

- Score 0: If no co-morbidities are reported in the trials and it is not assumable there are differences in co-morbidities.
- Score 1: If no co-morbidities are reported in the trials, but it is assumable that there are differences in co-morbidities.
- Score 2: If differences in important co-morbidities exceeds 30% or if differences are highly likely.

Signalling question 2: If co-morbidities are reported, are the co-morbidities equally presented in the trials?

- Score 0: If there are little differences in clinical important comorbidities, less than 20%.
- Score 1: If there are slight differences, between 20 and 30%.
- Score 2: If there are important differences, more than 30%.

Example 1. Meta-analysis on the use of dopamine for blood pressure regulation in ICU patients.

Trial 1 includes trauma patients and no reporting of co-morbidities

Trial 2 includes post cardiac surgery patients and no reporting of co-morbidities.

It is assumable that there are differences in co-morbidity between the trials, for example renal function pre-trial admission probably differs between the two groups. However, it is not stated, thus score a 1.

Example 2. Meta-analysis on the use of antibiotic prophylaxis in mechanical ventilated patients.
Trial 1 includes all ICU admitted patients in need for ventilation and reports the number of immune-compromised patients a 10%
Trial 2 includes all ICU admitted patients in need for ventilation but does not report the number of immune-compromised patients.

The number of immune compromised patients is high however, trial 2 does not report the number of included immune compromised patients. This item scores 1 point.

Example 3. Meta-analysis on the use of antibiotic prophylaxis in mechanical ventilated patients.
Trial 1 includes all ICU admitted patients in need for ventilation and reports the number of immune-compromised patients a 10%.
Trial 2 includes all ICU admitted patients in need for ventilation and report the number of immune-compromised patients a 6%.
In this example, the absolute difference is 4%, therefore score 0 points.

INTERVENTION HETEROGENEOTY

This domain is defined by multiple subdomains namely; 1. Intensity, strengths, or duration of intervention, 2. timing of the intervention, 3. control intervention, and 4. co-interventions.

1. INTENSITY, STRENGHTS, OR DURATION OF INTERVENTION

This subdomain assesses the intervention used in the different trials.

Signalling question 1: Do all trials use the same intervention?

- Score 2: If different types of interventions are used, there is clearly a heterogeneity.

Signalling question 2: If one intervention is used in all trials, is the intervention similar in each study? Is the dose, strength, cut-offs or duration of intervention similar?

- Score 0: When little variations, <20%, are present.
- Score 1: When slight variations, 20-30%, are present.
- Score 2: When considerable variations, > 30%, are presents.

If signalling question 2 cannot be answered by the information in the trials, this item should score a 1.

Example 1. Meta-analysis on the use of dopamine for blood pressure regulation in septic shock patients.

Trial 1 uses a dopamine dosage of 1 mcg/kg/min

Trial 2 uses a dopamine dosage of 3 mcg/kg/min

In this case the dosage differs > 30% relatively and this item scores with 2 points.

Example 2. Meta-analysis on the use of honey for wound treatment.

Trial 1 uses honey, undefined dosage.

Trial 2 uses honey 5 g/20 cm2.

The dosage is not defined, thus this item scores 1 point.

2. TIMING OF THE INTERVENTION

This subdomain assesses the timing of the intervention and should be scored whether the intervention is started at the same time.

Signalling question 1: Are the criteria of the start of the therapy well defined?

- Score 0: Well defined
- Score 1: If there is no information.
- Score 2: If different patient groups are included, because we expect that the criteria for starting cannot be similar if the patient groups are dissimilar.

Signalling question 2: Is the definition of the intervention stated in the different trials similar?

- Score 0: If the criteria on starting the therapy are similar, or differences of timing of intervention differs ≤20%.
- Score 1: If the criteria slightly differ or the timing difference is 20-30%.
- Score 2: If other criteria are used or the timing difference exceeds 30%.

Example 1. Meta-analysis on the use of dopamine for blood pressure regulation in septic shock patients.

Trial 1 starts the therapy when the systolic blood pressure is < 90mmHg.

Trial 2 starts the therapy when the systolic blood pressure is < 60mmHg.

In this example there is clearly a difference between the start of the therapy. Therefore, this item scores 2 points.

Example 2. Meta-analysis on the start of barbiturates in traumatic brain injury patients.

Trial 1 includes patients within 2 hours after admission to the hospital.

Trial 2 includes patients within 48 hours after admission to the hospital.

There is a substantial difference in timing, therefore this item scores 2.

3. CONTROL INTERVENTION

This subdomain assesses the use of comparable control interventions in the trials.

Signalling question 1: Do all trials use the same control intervention such as placebo, 'active' intervention, or no control?

- Score 0: If all trials use same control intervention or little variations, < 10%, are present.
- Score 1: If all trials use placebo or no intervention, or all trials use the same control intervention with slight variations, 10-20%.
- Score 2: If trials included in the meta-analysis use different control interventions or the same control intervention with considerable variations > 20%.

The focus is on the type of control intervention and to lesser extent the dosage or timing of the control intervention, however, assess as item 6 if an active control intervention is used.

Example 1. Meta-analysis on the use of desmopressin for nocturia.

Trial 1 uses alfuzosin 10 mg as a control intervention
Trial 2 uses a placebo.
The control interventions are different, therefore this meta-analysis scores 2 points on this subdomain.

4. CO-INTERVENTONS

This subdomain assesses the use of different co-interventions in the trials.

Signalling question 1: Do the trials give information on standard care?

Standard care is a widely used term that should be defined by the trials as the standard care often differs between hospitals (even within one country).

- Score 0: If it is assumable that all trials will have the same standard care or if no information is given on standard care but it is assumable that the trials use the same standard care, this item should score a 0.

- Score 1: If no information is given on standard care and it is not assumable that the trials use the same standard care, this item should score a 1.
- Score 2: Definition of standard care is different between trials.

Signalling question 2: Do the trials state a specific co-intervention?

- Score 0: If it is assumable that the other trials also used this specific co-intervention as standard care or if other trials do not use the same co-intervention, but the effect of the co-intervention will assumedly be little.
- Score 1: If other trials use the same co-intervention, but with slight variation (< 30 %).
- Score 2: If it is not assumable that the co-intervention is usual care, or differences in use of co-interventions ≥ 30 %.

Example 1. Meta-analysis on the use of dopamine for blood pressure regulation in septic shock patients.
Trial 1 states patients received the usual care.

Trial 2 states the use of fluid resuscitation and oxygen therapy.

In this example standard care of septic shock includes fluid resuscitation and oxygen therapy. Therefore, this example scores 0 points if there is no indication that standard care differed substantially for other interventions

Example 2. Meta-analysis on the use of dopamine for blood pressure regulation in septic shock patients.
Trial 1 states patients received the usual care.

Trial 2 states all patients were administered hydrocortisone.

The use of hydrocortisone in septic shock patients is not standard care, therefore this item scores 2 points.

Example 3. Meta-analysis on prophylactic antibiotics in ventilated patients.

Trial 1 includes post-operative liver transplantation patients.

Trial 2 includes post-operative cardiac surgery patients.

The patients included in trial 1 will also receive immunosuppressive medication, therefore the co-interventions will differ between trial 1 and 2. This item scores 2 points.

OUTCOME HETEROGENEITY

This domain is defined by two categories: 1. definition of the outcome and 2. timing of outcome measurement.

1. DEFINITION OF THE OUTCOME

Signalling question 1: Is the definition of the outcome in the meta-analysis and the trials similar?

- Score 0: If the same definition or *criteria* used in the trials and meta-analysis are the same.
- Score 1: If there are slight variation in the definition of the outcome.
- Score 2: If there are considerable variation in definition of outcome.

Example 1. Meta-analysis on the use of dopamine for blood pressure regulation in septic shock patients has the outcome all-cause mortality.
In this case mortality is not disputable; a patient is alive or deceased. This meta-analysis scores 0 points on this item. If the outcome differed among the trials and included both disease-specific mortality, e.g. mortality in organ-confined bladder cancer, and non-organ-confined mortality the meta-analysis would score 1.

Example 2. Meta-analysis on the use of antibiotic prophylaxis in mechanical ventilated patients has the outcome pneumonia.
Trial 1 defines pneumonia as positive sputum cultures.

Trial 2 defines pneumonia as diagnosed by a radiologist on an x-ray.

The definitions between the trials variate much, thus the meta-analysis scores 2 points on this item.

Example 3. Meta-analysis on the use of antibiotic prophylaxis in mechanical ventilated patients has the outcome pneumonia.

Trial 1 defines pneumonia as one positive sputum culture.

Trial 2 defines pneumonia as at least two positive sputum cultures.

The definitions slightly variate between trials, the meta-analysis scores 1 point.

2. TIMING OF OUTCOME MEASUREMENT

Signalling question 1: Is the time of the outcome measurement the same in all trials?

- Score 0: If the difference in the follow up of the outcome is less than one month.
- Score 1: If the difference in follow up is more than 1 month, but less than or equal to 3 months, or if timing of outcome measurement is not reported
- Score 2: If the difference in follow up exceeds 3 months.

Example 1. Meta-analysis on the use of dopamine for blood pressure regulation in septic shock patients has the outcome mortality.
Trial 1 scores mortality 14 days after start of the therapy.

Trial 2 score mortality 6 months after start of the therapy.

The difference between the follow up exceeds 3 months, thus this meta-analysis scores 2 points.

Example 2. Meta-analysis on the use of dopamine for blood pressure regulation in septic shock patients has the outcome mortality.
Trial 1 scores mortality 14 days after start of the therapy.

Trial 2 scores mortality 28 days after start of the therapy

The difference is less than 1 month, therefore score 0 points.

SUPPLEMENTARY APPENDIX 3

This supplementary has been provided by the authors to give readers additional information about their work.

Supplement to: Barbateskovic M, Koster TM, Eck R, et al. CHIMS: Clinical Heterogeneity I Meta-Analysis Score – a new tool to assess and quantify clinical heterogeneity in meta-analyses of interventions.

Contents

Main characteristics of 60 CHIMS assessed meta-analyses

	Review*	Meta-analysis outcome	Number of trials in meta-analysis	Intervention	Control
1	Bunn et al[1]	Mortality	31	Volume replacement with colloid solutions (albumin solutions or plasma protein fraction) in critically ill patients or surgical patients	Hydroxyethyl starch
2	Den Hertog et al[2]	Death or dependency	9	Temperature lowering therapy in patients with acute stroke	Placebo or open label
3	Mutter et al[3]	Renal replacement therapy	19	Fluid resuscitation with hydroxyethyl starch in patients with effective intravascular volume depletion	Other fluid resuscitation therapies
4	Bunn et al[4]	Death-trauma	6	Hypertonic crystalloid for fluid Resuscitation in patients with trauma, burns or who were undergoing surgery	Near isotonic crystalloid
5	Arrich et al[5]	Conventional cooling neurologic outcome conv	5	Therapeutic hypothermia (body temperature of 34°C or lower) in patients after cardiac arrest	Standard treatment (at the time of the trial)
6	Sydenham et al[6]	Death	21	Therapeutic cooling to maximum 35°C, either locally or systemically in patients with traumatic brain injury	No cooling
7	Alejandria et al[7]	Mortality	17	Intravenous immunoglobulin (IVIG) for treating sepsis, severe sepsis and septic shock in adults	Placebo or no intervention
8	Szakmany et al[8]	Mortality	11	N-acetylcysteine for sepsis and systemic inflammatory response in adults	Placebo or standard treatment
9	Afshari et al[9]	Mortality	13	Inhaled nitric oxide for acute respiratory distress syndrome	Placebo or no intervention

				(ARDS) and acute lung injury in children and adults	
10	Adhikari et al[10]	Prostaglandin - Mortality	7	Any pharmacologic therapy for the treatment of established ALI or ARDS	Placebo or no therapy
11	Marti-Carvajal et al[11]	Mortality	6	Human recombinant protein C plus conventional care for severe sepsis and septic shock in adult and paediatric patients	Placebo or no intervention plus conventional care
12	Berton et al[12]	Mortality	3	Quantitative cultures of respiratory secretions and invasive strategies in immunocompetent patients with ventilator-associated pneumonia	Qualitative cultures and non-invasive strategies
13	Brass et al[13]	Mortality - procedure	4	Percutaneous techniques for tracheostomy	Surgical techniques
14	Annane et al[14]	Mortality - 28 days	27	Corticosteroids for the treatment of sepsis in children and adults	Standard therapy (antibiotics, fluid replacement, inotropic or vasopressor therapy, mechanical ventilation, renal replacement therapy) or placebo
15	Subirana et al[15]	Ventilator associated pneumonia	11	Closed tracheal suction systems for mechanically ventilated adult patients	Open tracheal suction systems
16	Wikkelsø et al[16]	Mortality	8	Thromboelastography or thromboelastometry to monitor haemostatic treatment in adults or children with bleeding	Usual care
17	Sud et al[17]	Hospital or 30-day mortality	8	High-frequency oscillatory ventilation for acute respiratory distress syndrome	Conventional ventilation
18	Tian et al[18]	Mortality	3	Bicarbonate-buffered solutions for acute continuous haemodiafiltration or haemofiltration	Lactate-buffered solutions
19	Kelly et al[19]	Artificial airway occlusion	15	Heated humidification for ventilated adults and children	Heat and moisture exchangers

20	Allingstrup et al[20]	Mortality	16	Antithrombin III for critically ill patients	Placebo or no intervention
21	Lockhart et al[21]	Hernia recurrence	21	Mesh for inguinal and femoral hernia repair	Non-mesh
22	Brand et al[22]	All-cause mortality at 30 days	4	Surgical portosystemic shunts for variceal haemorrhage in people with cirrhosis	Transjugular intrahepatic portosystemic shunt
23	Saab et al[23]	24-months mortality	5	Transjugular intrahepatic portosystemic stent-shunts for cirrhotic patients with refractory ascites	Paracentesis
24	Cipriani et al[24]	Failure to respond at endpoint	6	Escitalopram for depression	Citalopram
25	Theologou et al[25]	In hospital death	5	Preoperative intra-aortic balloon pumps in patients undergoing coronary artery bypass grafting	No preoperative intra-aortic balloon pump
26	Shantsila et al[26]	Mortality	4	Antiplatelet treatment for patients with heart failure in sinus rhythm	Anticoagulation treatment
27	Andrade-Castelanos et al[27]	Incidence of death	7	Heparin for non-ST elevation acute coronary syndromes	Placebo
28	Aboumarzouk et al[28]	Stone free rate	7	Extracorporeal shock wave lithotripsy for ureteric calculi	Ureteroscopic management
29	Webster et al[29]	Surgical site infection	4	Preoperative bathing or showering with skin antiseptics to prevent surgical site infection	Placebo
30	Holme et al[30]	Colorectal cancer mortality, flex sigmoidoscopy	5	Flexible sigmoidoscopy for colorectal cancer screening in asymptomatic individuals	Faecal occult blood testing
31	Pouwer et al[31]	Live birth rate	5	Long-acting FSH for women undergoing assisted reproduction	Daily FSH
32	Bhutia et al[32]	Recurrence thromboembolic events	3	Once daily lowmolecular weight heparin for the initial treatment of venous thromboembolism	Twice daily lowmolecular weight heparin
33	Derry et al[33]	At least 50% pain intensity reduction	4	Pregabalin for neuropathic pain in adults	Placebo
34	Dasari et al[34]	Mortality	8	Surgical treatment of bile duct stones	Endoscopic treatment

35	Andras et al[35]	Incidence of recurrent VTE	16	Vitamin K antagonists for the long term treatment of symptomatic venous thromboembolism	Low-molecular-weight heparin
36	Wardlaw et al[36]	Mortality within 7-10 days	13	Thrombolysis for acute ischaemic stroke	Placebo
37	Sandercock et al[37]	Mortality	8	Corticosteroids for acute ischaemic stroke	Placebo
38	Hurley et al[38]	Survival to hospital discharge	5	Aminophylline for bradyasystolic cardiac arrest in adults	Placebo
39	Iheozor-Ejiofor et al[39]	Wound infection (short to medium term follow-up)	4	Negative pressure wound therapy for open traumatic wounds	Standard care
40	Wilson et al[40]	Perioperative complications	5	Laparoscopic for live kidney donors	Open nephrectomy
41	Rygard et al[41]	Mortality, time point closest to day 90.	5	Shorter storage time of transfused red blood cells in adult ICU patients	Longer storage time
42	Gebistorf et al[42]	Mortality	13	Inhaled nitric oxide for acute respiratory distress syndrome in children and adults	Placebo or no intervention
43	Lundstrøm et al[43]	Difficult tracheal intubation	4	Avoidance of neuromuscular blocking agents for improving conditions during tracheal intubation or direct laryngoscopy in adults and adolescents	Use of neuromuscular blocking agents
44	Fabritius et al[44]	24-hour postoperative opioid consumption and the incidence of Serious Adverse Events	9	Gabapentin for post-operative pain management	Placebo or an active placebo group mimicking the sedative effect of gabapentin
45	Sethi et al[45]	Mortality	18	Rhythm control strategies for atrial fibrillation and atrial flutter	Rate control strategies
46	Barbateskovic et al[46]	Mortality	25	Proton pump inhibitors or histamin-2 receptor antagonists for stress ulcer prophylaxis in adult intensive care patients	Placebo or no prophylaxis
47	Feinberg et al[47]	Mortality	114	Nutrition support in hospitalised adults at nutritional risk	Placebo, no intervention, treatment as usual
48	Meyhoff et al[48]	Mortality	8	Lower fluid volumes in patients with sepsis	Higher fluid volumes
49	Sethi et al[49]	Mortality	6	Digoxin for atrial fibrillation and atrial flutter	Any control intervention

50	Buggeskov et al[50]	Mortality	4	Pulmonary artery perfusion during cardiopulmonary bypass for open heart surgery in adults	No perfusion during cardiopulmonary bypass
51	Liang et al[51]	Prop. of participants with one or more SAE	10	Radix Sophorae flavescentis for chronic hepatitis B	Placebo or no intervention
52	Petersen et al[52]	Mortality ≤ 90 days	4	Untargeted antifungal therapy in adult patients with complicated intra-abdominal infection	Placebo or no treatment
53	Gareb et al[53]	Malocclusion	3	Biodegradable osteosyntheses (i.e., biodegradable plates and screws to fixate bone segments after maxillofacial trauma)	Titanium osteosyntheses
54	Blokzijl et al[54]	Mortality	6	Cardiac rehabilitation for patients after cardiac surgery	Usual care
55	French et al[55]	Mortality	8	Erythropoiesis stimulating agents in critically ill trauma patients	Placebo or no erythropoiesis stimulating agents
56	Lauridsen et al[56]	Number of patients with postoperative complications requiring treatment within 30 days	3	Robot-assisted cystectomy	Open radical cystectomy
57	Hiemstra et al[57]	Mortality (maximum follow-up)	17	Dopamine for cardiac dysfunction in critically ill patients	Placebo, no intervention, or any potentially active Comparator
58	Volbeda et al[58]	Mortality	31	Steroids for sepsis in patients with systemic inflammatory response syndrome, sepsis, severe sepsis or septic shock	Placebo, no intervention, or any other control intervention
59	Lai et al [59]	Clinical cure rate at day 3 after treatment	10	Tui Na for acute diarrhea in children under 5 years old	Conventional treatment or no treatment
60	Koster et al [60]	Mortality	44	Levosimendan for low cardiac output syndrome in critically ill patients	Any type of control

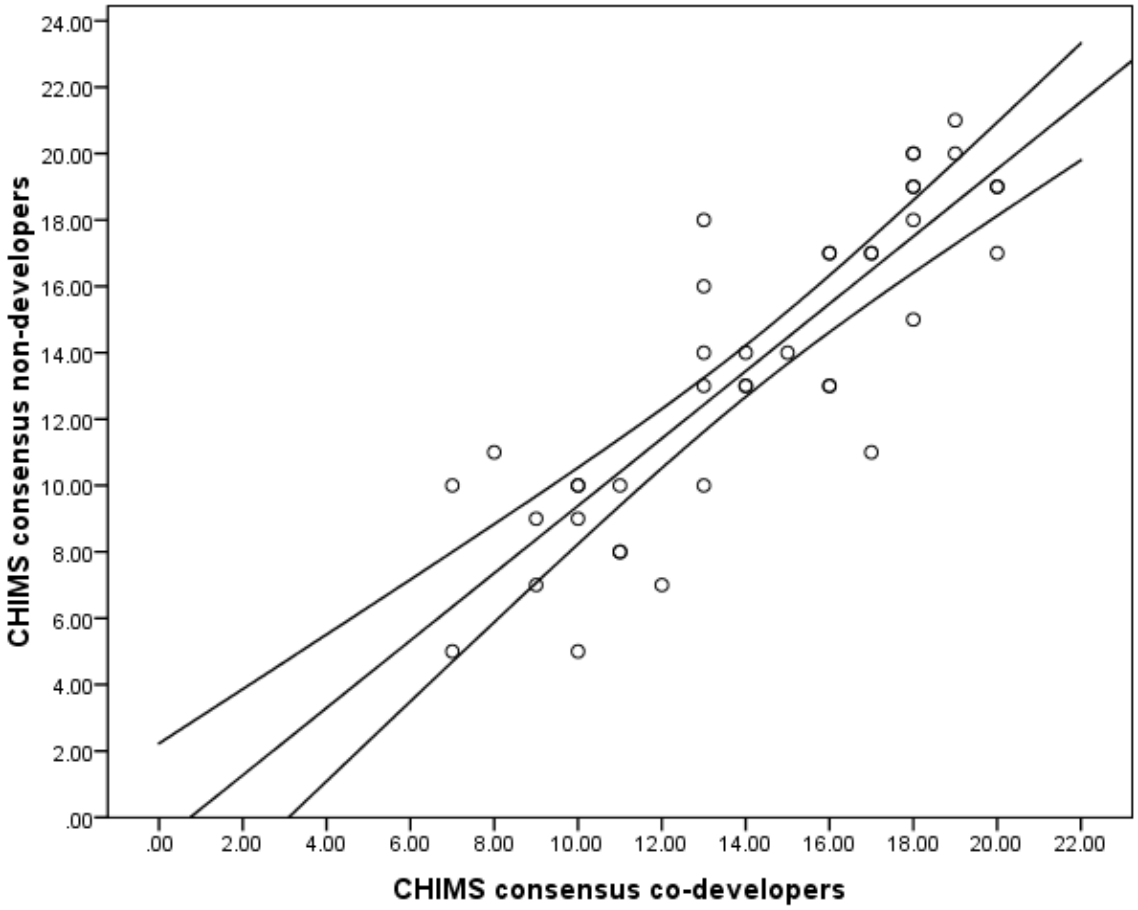
*1-20: ICU meta-analyses, 21-40: non-ICU meta-analyses; 41-60: reviewers of own meta-analyses

Interrater scale reliability between original review authors within domains (setting, population, intervention, and outcome) of CHIMS

Coefficients → / **Domains analysed** ↓	**Scale reliability: Intraclass correlation coefficient on average measures (95% CI)**	**Intraclass correlation coefficient on single measures (95% CI)**	**Pearson's correlation coefficient (interclass correlation) (95% CI)**
Summarised score of *setting* heterogeneity domain in CHIMS*	0.86 (0.67 to 0.95)	0.76 (0.49 to 0.90)	0.78 (48 to 1.13)
Summarised score of *population* heterogeneity domain in CHIMS*	0.93 (0.82 to 0.97)	0.87 (0.70 to 0.95)	0.86 (0.61 to 1.11)
Summarised score of *intervention* heterogeneity domain in CHIMS*	0.90 (0.75 to 0.96)	0.82 (0.60 to 0.93)	0.83 (0.55 to 1.09)
Summarised score of *outcome* heterogeneity domain in CHIMS*	0.68 (0.17 to 0.87)	0.51 (0.09 to 0.76)	0.50 (0.07 to 0.89)
Overall scale reliability of the total summarised CHIMS by pairs of original review authors*	0.94 (0.85 to 0.98)	0.89 (0.75 to 0.96)	0.90 (0.69 to 1.12)

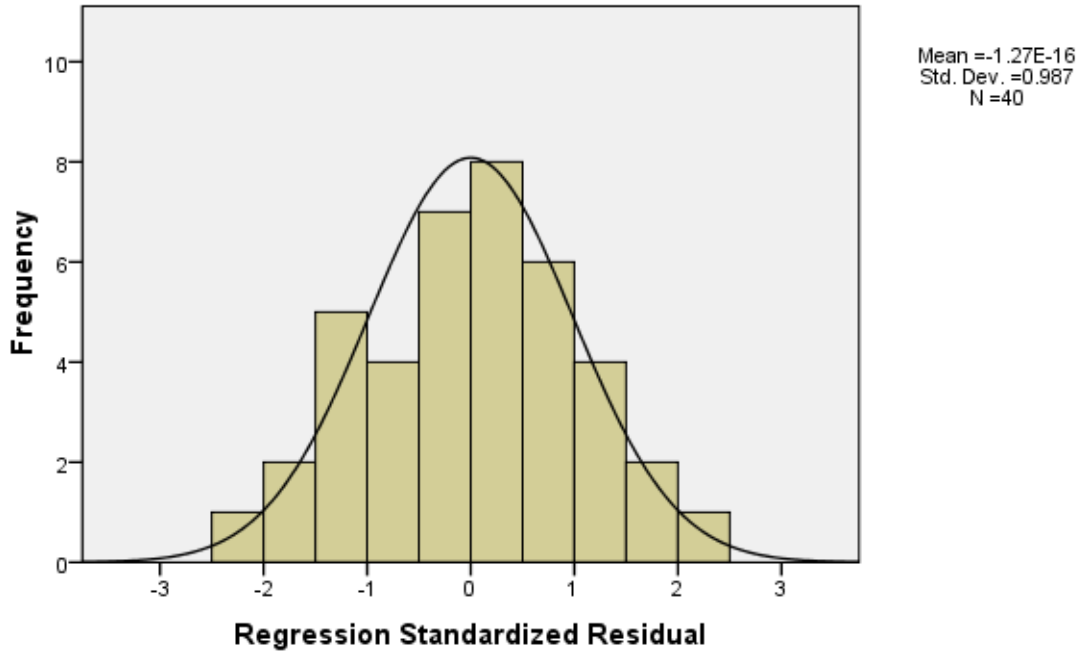
*Two-way random reliability of exact agreement analysis of 20 pairs of original review authors rating 20 meta-analyses with CHIMS not involved in the development of CHIMS. CI is confidence interval. SPSS ver. 19 was used.

Fitted regression line of consensus CHIMS from non-developers on consensus CHIMS from co-developers

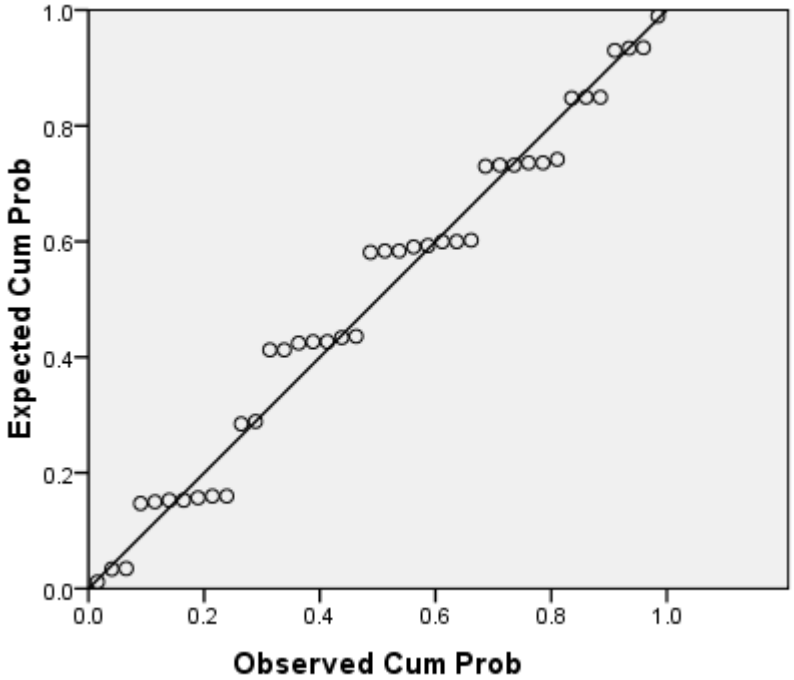


Fitted regression line (Y = 0.85 · X - 0.76) of consensus CHIMS from non-developers on consensus CHIMS from co-developers in 40 meta-analyses from 20 ICU and 20 non-ICU meta-analyses. Hyperbolic lines around fitted line represents 95% CI for the regression line. R^2=0.73.

Residual histogram

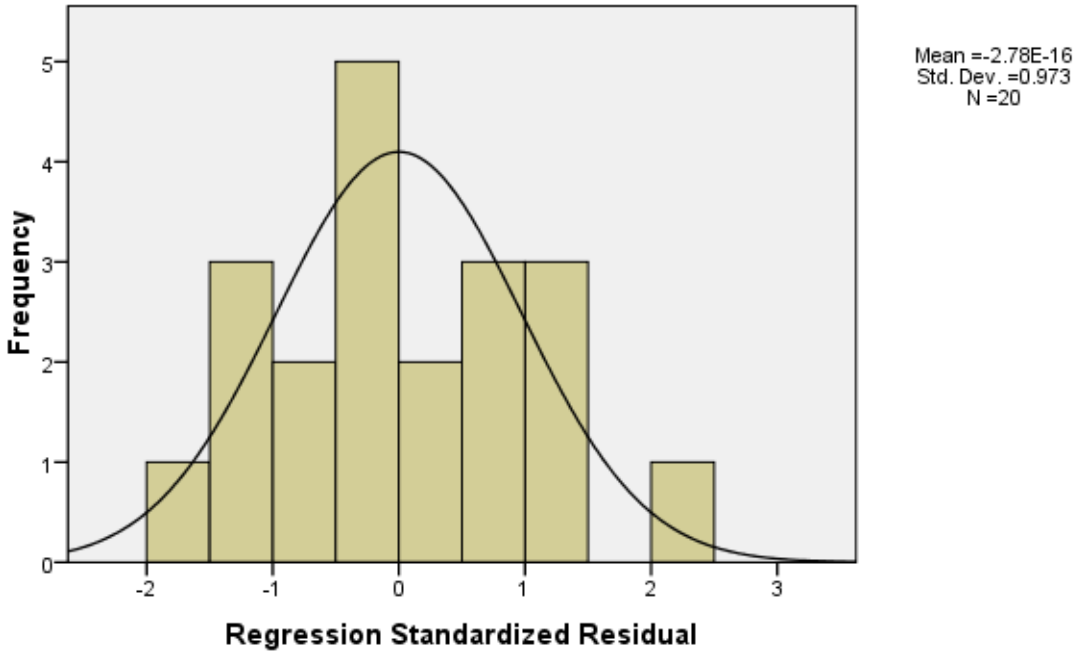


Normal P-P plot of regression standardised residuals

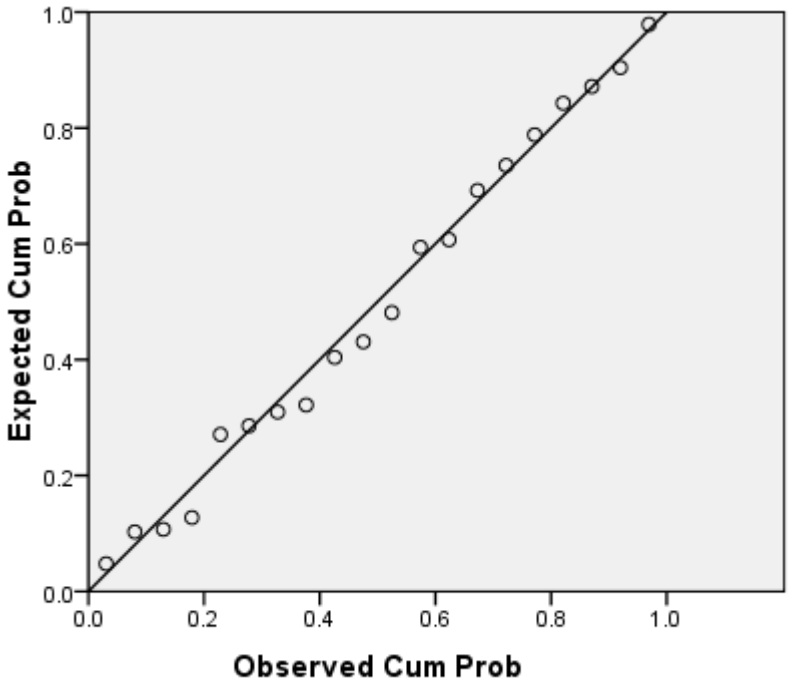


Supplemental figures for figure 3

Residual histogram

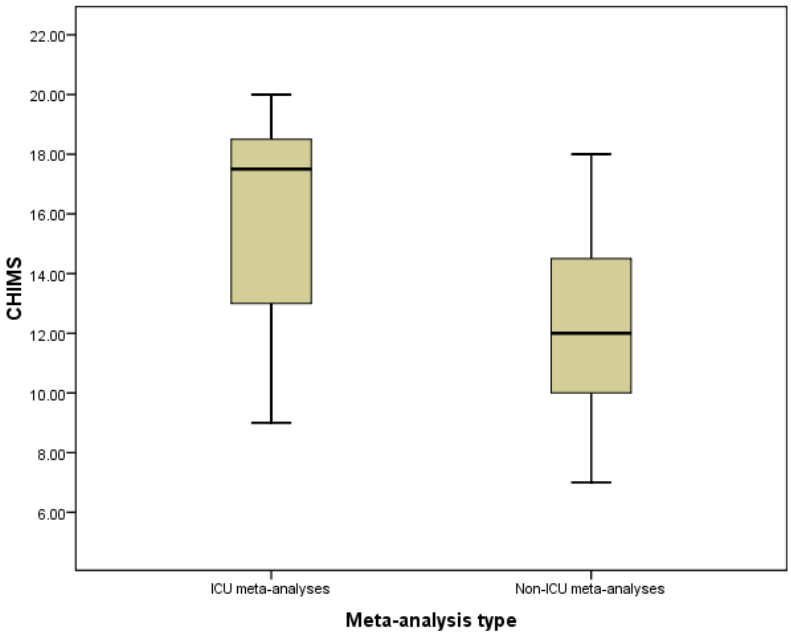


Normal P-P plot of regression standardised residuals

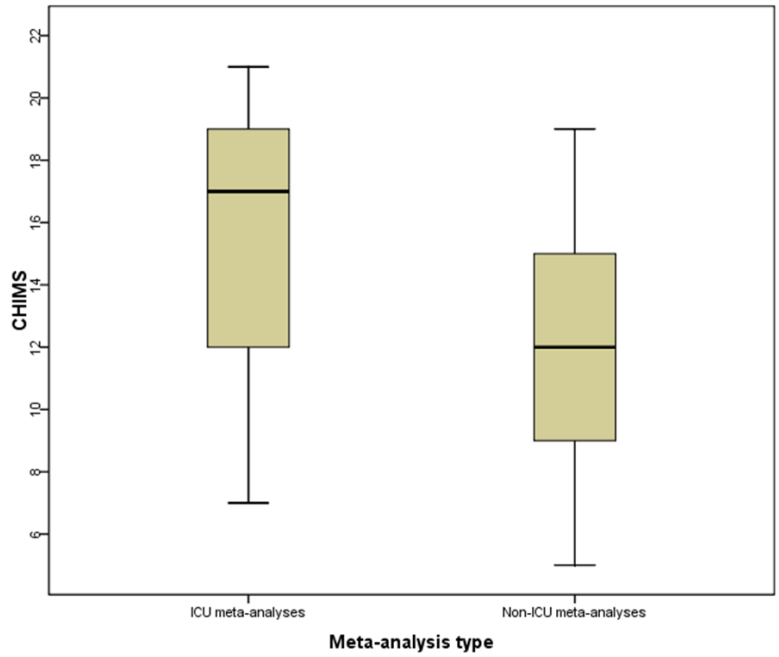


Supplemental figures for Table 3

Box and whiskers plots



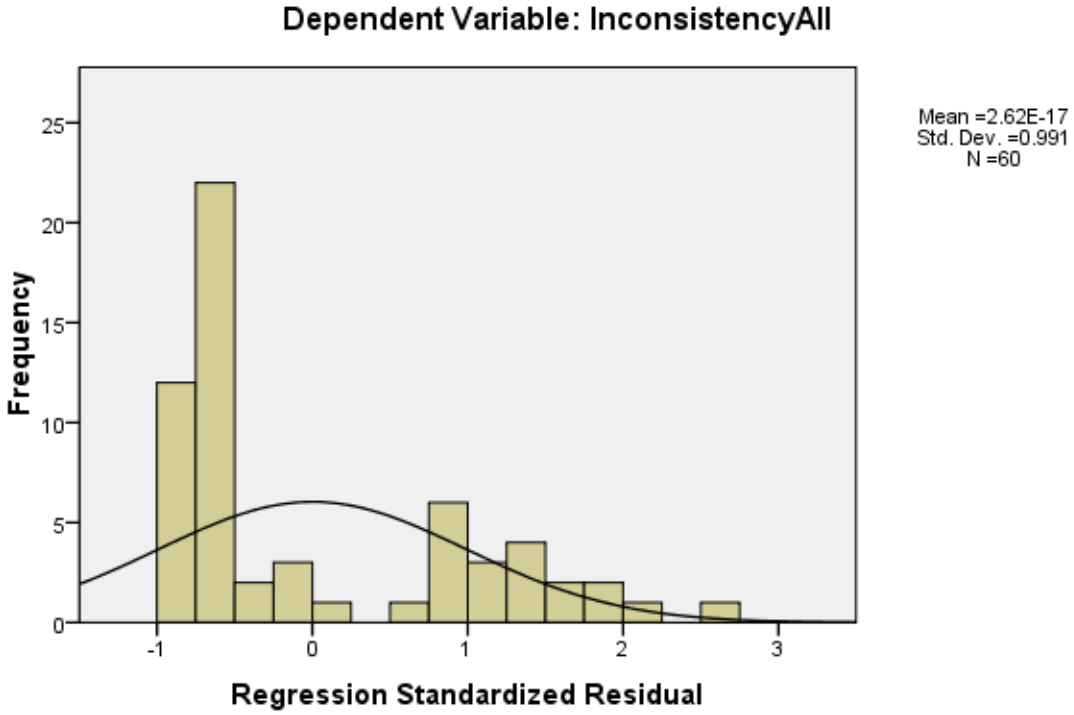
Box and whiskers plot (95% CI) of consensus CHIMS score between developers of CHIMS stratified for ICU and non-ICU meta-analyses. Mann-Whitney U test for different distributions of CHIMS in ICU (median 18; range 9-20) and non-ICU (median 12; range 7-18) meta-analyses, P=0.001.



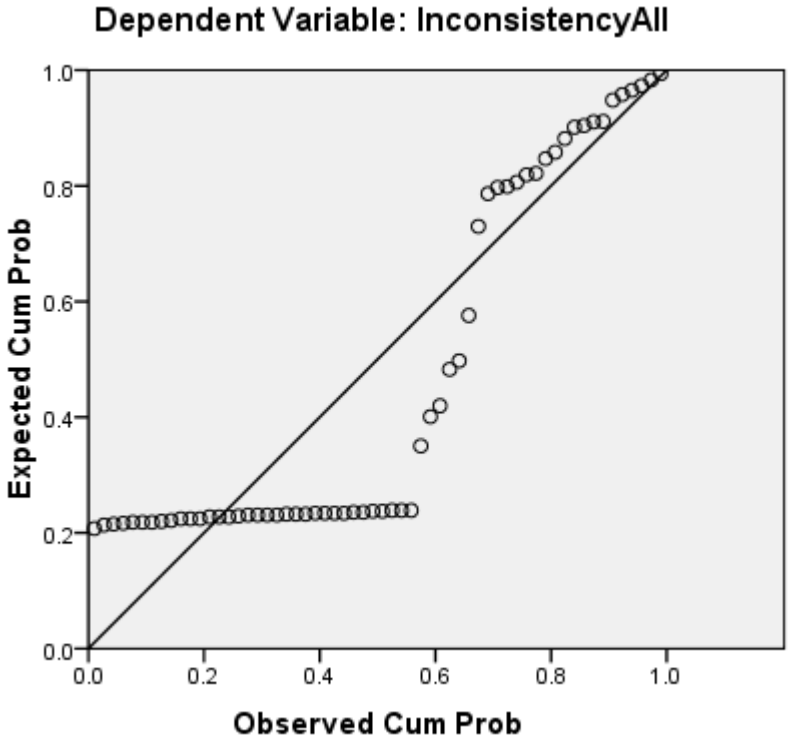
Box and whiskers plot (95% CI) of consensus CHIMS score between non-developers of CHIMS stratified for ICU and non-ICU meta-analyses. Mann-Whitney test for different distributions of CHIMS in ICU (median 17; range 7-21) and non-ICU meta-analyses (median 12; range 5-19) P=0.016.

Supplemental figures for Table 4

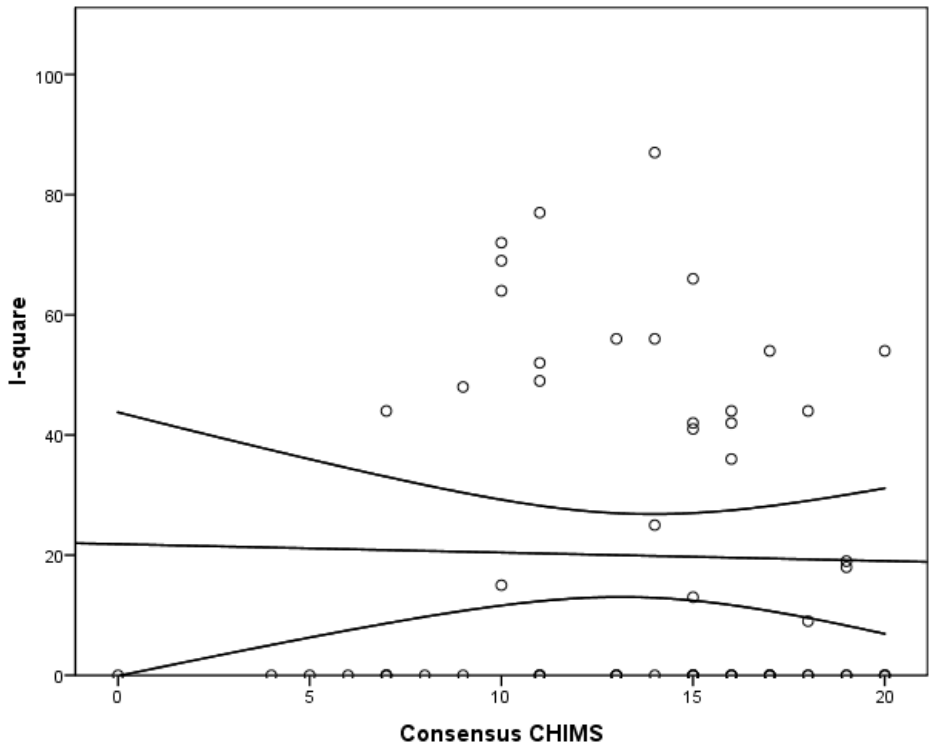
Histogram, co-developers



Normal P-P plot of regression standardised residuals, co-developers



Fitted regression line for regression of I^2 on consensus CHIMS



Fitted regression line for regression of I^2 on consensus CHIMS between pairs of co-developers supplemented with consensus CHIMS of pairs of reviewers of own meta-analyses. The hyperbolic curves around the regression line (Y= -0.02 · X + 21.5) represent 95% CI for the fitted line. R^2=0.001.

References

1. Bunn F, Trivedi D. Colloid solutions for fluid resuscitation. Cochrane Database Syst Rev. 2012(7):Cd001319.
2. Den Hertog HM, van der Worp HB, Tseng MC, Dippel DW. Cooling therapy for acute stroke. Cochrane Database Syst Rev. 2009(1):Cd001247.
3. Mutter TC, Ruth CA, Dart AB. Hydroxyethyl starch (HES) versus other fluid therapies: effects on kidney function. Cochrane Database Syst Rev. 2013(7):Cd007594.
4. Bunn F, Roberts I, Tasker R, Akpa E. Hypertonic versus near isotonic crystalloid for fluid resuscitation in critically ill patients. Cochrane Database Syst Rev. 2004(3):Cd002045.
5. Arrich J, Holzer M, Havel C, Mullner M, Herkner H. Hypothermia for neuroprotection in adults after cardiopulmonary resuscitation. Cochrane Database Syst Rev. 2016;2:Cd004128.
6. Sydenham E, Roberts I, Alderson P. Hypothermia for traumatic head injury. Cochrane Database Syst Rev. 2009(2):Cd001048.
7. Alejandria MM, Lansang MA, Dans LF, Mantaring JB, 3rd. Intravenous immunoglobulin for treating sepsis, severe sepsis and septic shock. Cochrane Database Syst Rev. 2013(9):Cd001090.
8. Szakmany T, Hauser B, Radermacher P. N-acetylcysteine for sepsis and systemic inflammatory response in adults. Cochrane Database Syst Rev. 2012(9):Cd006616.
9. Afshari A, Brok J, Moller AM, Wetterslev J. Inhaled nitric oxide for acute respiratory distress syndrome (ARDS) and acute lung injury in children and adults. Cochrane Database Syst Rev. 2010(7):Cd002787.
10. Adhikari N, Burns KE, Meade MO, Ratnapalan M. Pharmacologic therapies for adults with acute lung injury and acute respiratory distress syndrome. Cochrane Database Syst Rev. 2004(4):Cd004477.
11. Marti-Carvajal AJ, Sola I, Gluud C, Lathyris D, Cardona AF. Human recombinant protein C for severe sepsis and septic shock in adult and paediatric patients. Cochrane Database Syst Rev. 2012;12:Cd004388.
12. Berton DC, Kalil AC, Teixeira PJ. Quantitative versus qualitative cultures of respiratory secretions for clinical outcomes in patients with ventilator-associated pneumonia. Cochrane Database Syst Rev. 2014(10):Cd006482.
13. Brass P, Hellmich M, Ladra A, Ladra J, Wrzosek A. Percutaneous techniques versus surgical techniques for tracheostomy. Cochrane Database Syst Rev. 2016;7:Cd008045.
14. Annane D, Bellissant E, Bollaert PE, Briegel J, Keh D, Kupfer Y. Corticosteroids for treating sepsis. Cochrane Database Syst Rev. 2015(12):Cd002243.
15. Subirana M, Sola I, Benito S. Closed tracheal suction systems versus open tracheal suction systems for mechanically ventilated adult patients. Cochrane Database Syst Rev. 2007(4):Cd004581.
16. Wikkelso A, Wetterslev J, Moller AM, Afshari A. Thromboelastography (TEG) or thromboelastometry (ROTEM) to monitor haemostatic treatment versus usual care in adults or children with bleeding. Cochrane Database Syst Rev. 2016(8):Cd007871.
17. Sud S, Sud M, Friedrich JO, Wunsch H, Meade MO, Ferguson ND, et al. High-frequency oscillatory ventilation versus conventional ventilation for acute respiratory distress syndrome. Cochrane Database Syst Rev. 2016;4:Cd004085.
18. Tian JH, Ma B, Yang K, Liu Y, Tan J, Liu TX. Bicarbonate- versus lactate-buffered solutions for acute continuous haemodiafiltration or haemofiltration. Cochrane Database Syst Rev. 2015(3):Cd006819.
19. Kelly M, Gillies D, Todd DA, Lockwood C. Heated humidification versus heat and moisture exchangers for ventilated adults and children. Cochrane Database Syst Rev. 2010(4):Cd004711.
20. Allingstrup M, Afshari A. Selenium supplementation for critically ill adults. Cochrane Database Syst Rev. 2015(7):Cd003703.

21. Lockhart K, Dunn D, Teo S, Ng JY, Dhillon M, Teo E, et al. Mesh versus non-mesh for inguinal and femoral hernia repair. Cochrane Database Syst Rev. 2018;9:Cd011517.
22. Brand M, Prodehl L, Ede CJ. Surgical portosystemic shunts versus transjugular intrahepatic portosystemic shunt for variceal haemorrhage in people with cirrhosis. Cochrane Database Syst Rev. 2018;10:Cd001023.
23. Saab S, Nieto JM, Lewis SK, Runyon BA. TIPS versus paracentesis for cirrhotic patients with refractory ascites. Cochrane Database Syst Rev. 2006(4):Cd004889.
24. Cipriani A, Santilli C, Furukawa TA, Signoretti A, Nakagawa A, McGuire H, et al. Escitalopram versus other antidepressive agents for depression. Cochrane Database Syst Rev. 2009(2):Cd006532.
25. Theologou T, Bashir M, Rengarajan A, Khan O, Spyt T, Richens D, et al. Preoperative intra aortic balloon pumps in patients undergoing coronary artery bypass grafting. Cochrane Database Syst Rev. 2011(1):Cd004472.
26. Shantsila E, Lip GY. Antiplatelet versus anticoagulation treatment for patients with heart failure in sinus rhythm. Cochrane Database Syst Rev. 2016;9:Cd003333.
27. Andrade-Castellanos CA, Colunga-Lozano LE, Delgado-Figueroa N, Magee K. Heparin versus placebo for non-ST elevation acute coronary syndromes. Cochrane Database Syst Rev. 2014(6):Cd003462.
28. Aboumarzouk OM, Kata SG, Keeley FX, McClinton S, Nabi G. Extracorporeal shock wave lithotripsy (ESWL) versus ureteroscopic management for ureteric calculi. Cochrane Database Syst Rev. 2012(5):Cd006029.
29. Webster J, Osborne S. Preoperative bathing or showering with skin antiseptics to prevent surgical site infection. Cochrane Database Syst Rev. 2015(2):Cd004985.
30. Holme O, Bretthauer M, Fretheim A, Odgaard-Jensen J, Hoff G. Flexible sigmoidoscopy versus faecal occult blood testing for colorectal cancer screening in asymptomatic individuals. Cochrane Database Syst Rev. 2013(9):Cd009259.
31. Pouwer AW, Farquhar C, Kremer JA. Long-acting FSH versus daily FSH for women undergoing assisted reproduction. Cochrane Database Syst Rev. 2015(7):Cd009577.
32. Bhutia S, Wong PF. Once versus twice daily low molecular weight heparin for the initial treatment of venous thromboembolism. Cochrane Database Syst Rev. 2013(7):Cd003074.
33. Derry S, Bell RF, Straube S, Wiffen PJ, Aldington D, Moore RA. Pregabalin for neuropathic pain in adults. Cochrane Database Syst Rev. 2019;1:Cd007076.
34. Dasari BV, Tan CJ, Gurusamy KS, Martin DJ, Kirk G, McKie L, et al. Surgical versus endoscopic treatment of bile duct stones. Cochrane Database Syst Rev. 2013(12):Cd003327.
35. Andras A, Sala Tenna A, Stewart M. Vitamin K antagonists versus low-molecular-weight heparin for the long term treatment of symptomatic venous thromboembolism. Cochrane Database Syst Rev. 2017;7:Cd002001.
36. Wardlaw JM, Murray V, Berge E, del Zoppo GJ. Thrombolysis for acute ischaemic stroke. Cochrane Database Syst Rev. 2014(7):Cd000213.
37. Sandercock PA, Soane T. Corticosteroids for acute ischaemic stroke. Cochrane Database Syst Rev. 2011(9):Cd000064.
38. Hurley KF, Magee K, Green R. Aminophylline for bradyasystolic cardiac arrest in adults. Cochrane Database Syst Rev. 2015(11):Cd006781.
39. Iheozor-Ejiofor Z, Newton K, Dumville JC, Costa ML, Norman G, Bruce J. Negative pressure wound therapy for open traumatic wounds. Cochrane Database Syst Rev. 2018;7:Cd012522.
40. Wilson CH, Sanni A, Rix DA, Soomro NA. Laparoscopic versus open nephrectomy for live kidney donors. Cochrane Database Syst Rev. 2011(11):Cd006124.
41. Rygard SL, Jonsson AB, Madsen MB, Perner A, Holst LB, Johansson PI, et al. Effects of shorter versus longer storage time of transfused red blood cells in adult ICU patients: a systematic review with meta-analysis and Trial Sequential Analysis. Intensive Care Med. 2018;44(2):204-17.
42. Gebistorf F, Karam O, Wetterslev J, Afshari A. Inhaled nitric oxide for acute respiratory distress syndrome (ARDS) in children and adults. Cochrane Database Syst Rev. 2016(6):Cd002787.

43. Lundstrom LH, Duez CHV, Norskov AK, Rosenstock CV, Thomsen JL, Moller AM, et al. Effects of avoidance or use of neuromuscular blocking agents on outcomes in tracheal intubation: a Cochrane systematic review. Br J Anaesth. 2018;120(6):1381-93.
44. Fabritius ML, Geisler A, Petersen PL, Nikolajsen L, Hansen MS, Kontinen V, et al. Gabapentin for post-operative pain management - a systematic review with meta-analyses and trial sequential analyses. Acta Anaesthesiol Scand. 2016;60(9):1188-208.
45. Sethi NJ, Feinberg J, Nielsen EE, Safi S, Gluud C, Jakobsen JC. The effects of rhythm control strategies versus rate control strategies for atrial fibrillation and atrial flutter: A systematic review with meta-analysis and Trial Sequential Analysis. PLoS One. 2017;12(10):e0186856.
46. Barbateskovic M, Marker S, Granholm A, Anthon CT, Krag M, Jakobsen JC, et al. Stress ulcer prophylaxis with proton pump inhibitors or histamin-2 receptor antagonists in adult intensive care patients: a systematic review with meta-analysis and trial sequential analysis. Intensive Care Med. 2019;45(2):143-58.
47. Feinberg J, Nielsen EE, Korang SK, Halberg Engell K, Nielsen MS, Zhang K, et al. Nutrition support in hospitalised adults at nutritional risk. Cochrane Database Syst Rev. 2017;5:Cd011598.
48. Meyhoff T, Co-authors? Title?? (Submitted).
49. Sethi NJ, Nielsen EE, Safi S, Feinberg J, Gluud C, Jakobsen JC. Digoxin for atrial fibrillation and atrial flutter: A systematic review with meta-analysis and trial sequential analysis of randomised clinical trials. PLoS One. 2018;13(3):e0193924.
50. Buggeskov KB, Gronlykke L, Risom EC, Wei ML, Wetterslev J. Pulmonary artery perfusion versus no perfusion during cardiopulmonary bypass for open heart surgery in adults. Cochrane Database Syst Rev. 2018;2:Cd011098.
51. Liang N, Kong Z, Ma SS, Lu CL, Yang M, Feng LD, et al. Radix Sophorae flavescentis versus no intervention or placebo for chronic hepatitis B. Cochrane Database Syst Rev. 2019;4:Cd013089.
52. Petersen MW, Perner A, Ravn F, Sjovall F, Moller MH. Untargeted antifungal therapy in adult patients with complicated intra-abdominal infection: a systematic review. Acta Anaesthesiol Scand. 2018;62(1):6-18.
53. Gareb B, van Bakelen N, Dijkstra P, Vissink A, Bos R, van Minnen B. Biodegradable versus titanium osteosyntheses in maxillofacial traumatology: a systematic review with meta-analysis and trial sequential analysis. Int J Oral Maxillofac Surg (submitted). 2019.
54. Blokzijl F, Dieperink W, Keus F, Reneman MF, Mariani MA, van der Horst IC. Cardiac rehabilitation for patients having cardiac surgery: a systematic review. J Cardiovasc Surg (Torino). 2018;59(6):817-29.
55. French CJ, Glassford NJ, Gantner D, Higgins AM, Cooper DJ, Nichol A, et al. Erythropoiesis-stimulating Agents in Critically Ill Trauma Patients: A Systematic Review and Meta-analysis. Ann Surg. 2017;265(1):54-62.
56. Lauridsen SV, Tonnesen H, Jensen BT, Neuner B, Thind P, Thomsen T. Complications and health-related quality of life after robot-assisted versus open radical cystectomy: a systematic review and meta-analysis of four RCTs. Syst Rev. 2017;6(1):150.
57. Hiemstra B, Koster G, Wetterslev J, Gluud C, Jakobsen JC, Scheeren TWL, et al. Dopamine in critically ill patients with cardiac dysfunction: A systematic review with meta-analysis and trial sequential analysis. Acta Anaesthesiol Scand. 2019;63(4):424-37.
58. Volbeda M, Wetterslev J, Gluud C, Zijlstra JG, van der Horst IC, Keus F. Glucocorticosteroids for sepsis: systematic review with meta-analysis and trial sequential analysis. Intensive Care Med. 2015;41(7):1220-34.
59. Lai BY, Liang N, Cao HJ, Yang GY, Jia LY, Hu RX, et al. Pediatric Tui Na for acute diarrhea in children under 5 years old: A systematic review and meta-analysis of randomized clinical trials. Complement Ther Med. 2018;41:10-22.
60. Koster G, Wetterslev J, Gluud C, Zijlstra JG, Scheeren TW, van der Horst IC, et al. Effects of levosimendan for low cardiac output syndrome in critically ill patients: systematic review with meta-analysis and trial sequential analysis. Intensive Care Med. 2015;41(2):203-21.